丘成桐数学科学中心
YAU MATHEMATICAL SCIENCES CENTER

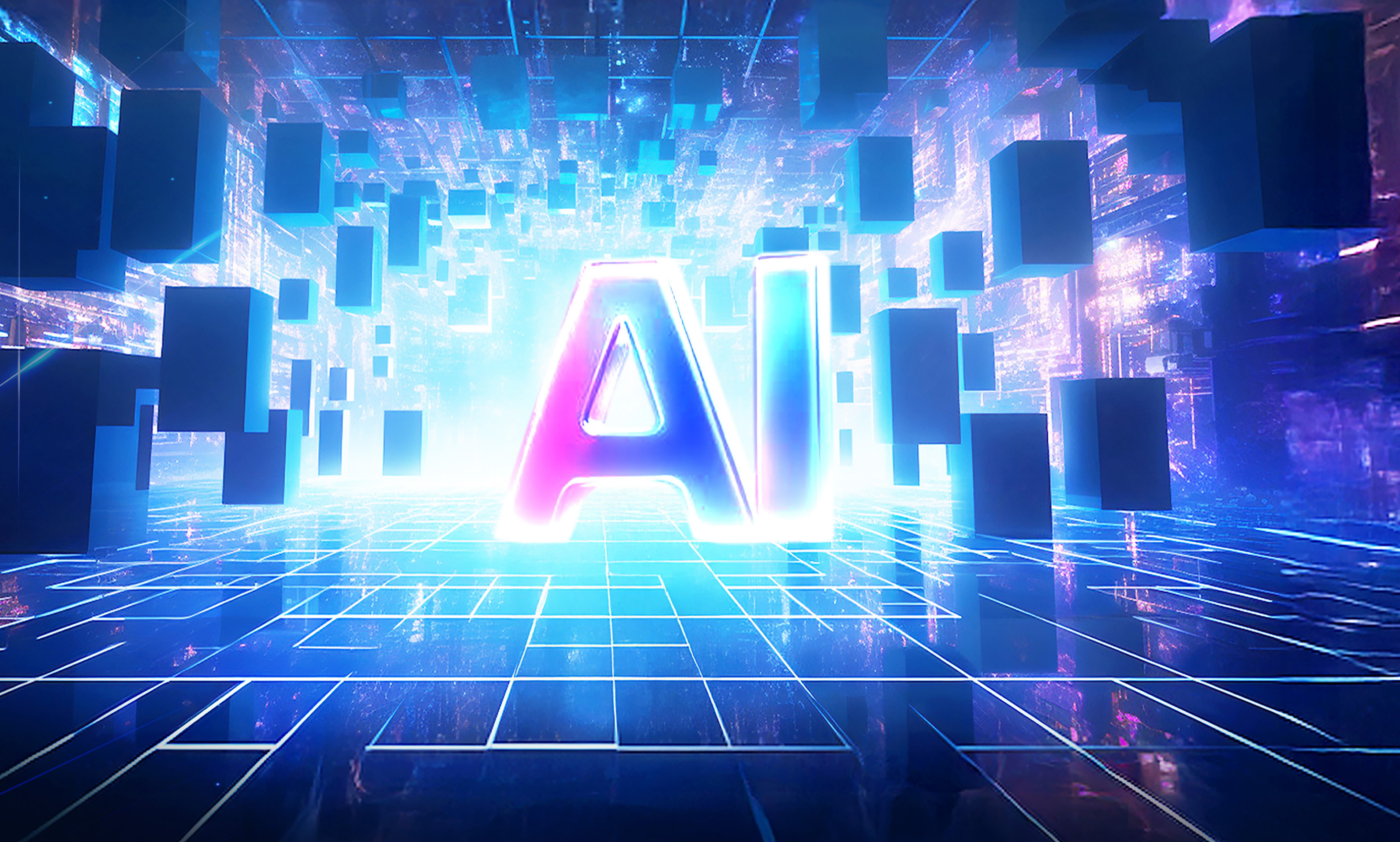
TSIMF
清华三亚国际数学论坛

人工智能时代的时空相依性统计学习

Learning spatial and temporal dependence in the age of AI

January 5-9, 2026

Room A-110, TSIMF



组织者 ORGANIZERS

Oliver Linton, University of Cambridge

Yundong Tu(涂云东), Peking University

Yuhong Yang(杨宇红), Tsinghua University

Qiwei Yao(姚琦伟), London School of Economics and Political Science

Contents

About the conference

人工智能时代的时空相依性统计学习

Learning spatial and temporal dependence in the age of AI

Date

January 5-9, 2026

Venue

Room A-110, TSIMF

Organizers

Oliver Linton, University of Cambridge

Yundong Tu(涂云东), Peking University

Yuhong Yang(杨宇红), Tsinghua University

Qiwei Yao(姚琦伟), London School of Economics and Political Science

Abstract

Data with temporal and spatial dependencies are ubiquitous in every domain. The age of AI has introduced unprecedented challenges, as well as opportunities, for research on the statistical learning for dependent data. While the ideas from machine learning and deep neural networks/AI have demonstrated remarkable empirical performance across a range of tasks, there is huge potential to use these ideas with novel approaches for learning dependence, especially beyond linear correlations, in high-dimensional setting, with complex data structures, or under nonstationary scenarios. On the other hand, there is an urgent need to gain some appreciation when/how/why DNN and various learning techniques work for dependent data, while the results on understanding DNN with independent data start to merge.

This workshop aims to draw energy and inspiration from the recent development in machine learning (ML) including DNN to tackle some important challenges in complex time series analysis and spatial statistics, and reversely from tools of analyzing dependent data to aid ML. It will gather the leading scholars as well as young researchers in this broad area to reflect and to exchange the new advancements in the analysis of complex dependent data, with a balanced focus on methodology, theory, and algorithms. The key topics include (i) modelling complex time series including tensor processes, dynamic networks, and spatial-temporal processes, (ii) nonlinear PCA/ICA for time series, (iii) nonstationary spatial processes and spatial cointegration, (iv) change-point/anomaly detection in time and/or over space, (v) applications of tools of analyzing dependent data in ML/AI.

The workshop will feature oral presentations, poster sessions, and topic/panel discussions.

Description of the aim

The workshop will cover five inter-correlated areas under the umbrella of learning dependence in the age of AI.

1. Modelling complex time series including tensor processes, dynamic networks, and spatial-temporal processes

Modelling complex time series involving tensor processes, dynamic networks, and spatial- temporal processes presents multifaceted challenges. Tensor processes, with their multi- dimensional nature, not only increase computational costs and the risk of overfitting but also pose interpretability issues. Dynamic networks require capturing their constantly evolving structure, yet the driving factors are often complex and unobservable, and scaling models for large-scale data is computationally demanding. Spatial-temporal processes need to account for nonlinear and variable spatial and temporal dependencies, while also contending with data sparsity that can bias results. Integrating these components into a unified framework is arduous, as their complex interactions can lead to excessive model complexity, necessitating a delicate balance between accuracy and computational feasibility.

2. Nonlinear PCA/ICA for time series

While principal/independent component analysis (PCA/ICA) are effective to transform a multi- or high-dimensional data analysis problem into a number of univariable problems, their application to time series or spatial data are less than ideal, as the dependence across different times or over space are ignored. The progress has been made to extend PCA for time series in the sense that one can derive a contemporaneous linear transformation for a p-variate time series such that the transformed series is segmented into several lower-dimensional subseries, and those subseries are uncorrelated with each other both contemporaneously and serially. Therefore, those lower-dimensional series can be analyzed separately as far as the linear dynamic structure is concerned. However such an approach ignore the dependence beyond linear correlation. The DNN architecture Autoencoder can be viewed a nonlinear generalization of PCA. However it is not suitable for time series PCA/ICA for three reasons: (i) the transformation deduced is not necessarily invertible, (ii) it does not require the `components' to be uncorrelated or independent, and (iii) it ignores the dependence across different lags. The workshop will provide a forum to explore some particular class of invertible DNN (such as Real NVP) for nonlinear PCA/ICA for time series.

3. Nonstationary spatial processes and spatial cointegration

Spatial data analysis has become increasingly relevant in various fields such as economics, environmental sciences, epidemiology, and social sciences. Most real-world spatial processes are nonstationary. Simultaneously, spatial cointegration has emerged as an important concept when dealing with nonstationary spatial processes. Cointegration, in a time series context, refers to a situation where two or more nonstationary series are linked by a long-term equilibrium relationship. When extended to spatial data, spatial cointegration explores whether spatial units (such as geographic regions or neighboring locations) exhibit long-term relationships that account for spatial dependencies and common trends. There are two approaches in dealing nonstationarity: (i) econometric spatial autoregression in which the nonstationarites are reflected in weight matrices, (ii) intrinsic processes which can be viewed as a generalization of differenced time series to spatial processes sampled over irregular grids. Both the inference methods and the associated theory for cointegration under both setting are to be developed.

4. Change-point/anomaly detection in time and/or over space

Ideas from machine learning and AI have demonstrated remarkable empirical performance across a range of tasks. There is huge potential to use these ideas with novel approaches to detect change-points/anomalies, albeit with a simultaneous need to better understand theoretically why and when these methods work reliably at detecting and localizing change-points/anomalies. Deep learning offers an effective framework for understanding the characteristics of "normal" data, creating a baseline for detecting change points. These methods are adaptive to various types of latent structure in the data, thanks to their ability to learn complex patterns and representations from data. Another area which will be explored in the workshop is transfer Learning. This allows for pre-trained models on large datasets to be fine-tuned for specific change-point/anomaly detection tasks. This is particularly suited for some condition monitoring applications where one can learn a model based on data across many similar machines/sites, which then can be refined with much more limited data to be specific for each individual machine/site.

more limited data to be specific for each individual machine/site.

5. Applications of tools of analyzing dependent data in ML/AI

Recent developments in AI intend to take advantage of dependences. Examples include attention weights in transformers capturing dependence between tokens, CNN and RNN leveraging structured dependencies for better modeling performance. In recent studies on AI model protection, dependent perturbations are used to prevent an adversarial to steal an AI model in service. Insight.....

Schedule

<table>
<tr><th>Time&Date</th><th>Monday
(January 5)</th><th>Tuesday
(January 6)</th><th>Wednesday
(January 7)</th><th>Thursday
(January 8)</th><th>Friday
(January 9)</th></tr>
<tr><td>7:30-8:30</td><td colspan="5">Breakfast (60 minutes)</td></tr>
<tr><td>8:30-8:40</td><td>Openning Remark</td><td colspan="4"></td></tr>
<tr><td>Chair</td><td>Yundong Tu</td><td>Qiwei Yao</td><td>Yuhong Yang</td><td>Qianqian Zhu</td><td>Baiqing Wang</td></tr>
<tr><td>8:40-9:20</td><td>Ching-Kang Ing</td><td>Bingyi Jing</td><td>Jian Huang</td><td>Qiying Wang</td><td>Xinyang Yu</td></tr>
<tr><td>9:20-10:00</td><td>Junhui Wang</td><td>Binyan Jiang</td><td>Degui Li</td><td>Yingqian Lin</td><td>Dan Yang</td></tr>
<tr><td>10:00-10:30</td><td colspan="5">Coffee Break (30 minutes)</td></tr>
<tr><td>Chair</td><td>Shaojun Guo</td><td>Rongmao Zhang</td><td>Haihan Tang</td><td>Dan Yang</td><td>Xinling Xie</td></tr>
<tr><td>10:30-11:10</td><td>Qianqian Zhu</td><td>Yang Zu</td><td>Jinyuan Chang</td><td>Ling Zhou</td><td>Han Yan</td></tr>
<tr><td>11:10-11:50</td><td>Zhentao Shi</td><td>Xinling Xie</td><td>Zhou Yu</td><td>Chaohua Dong</td><td>Baiqing Wang</td></tr>
<tr><td>12:00-13:30</td><td colspan="5">Lunch (90 minutes)</td></tr>
<tr><td>Chair</td><td>Bin Guo</td><td rowspan="7">Free Discussion</td><td>Ying Wang</td><td>Han Yan</td><td rowspan="7">Free Discussion</td></tr>
<tr><td>14:00-14:40</td><td>Yuying Sun</td><td>Anna Bykhovskaya</td><td>Weichi Wu</td></tr>
<tr><td>14:40-15:20</td><td>Dalei Yu</td><td>Qing Yang</td><td>Xinghao Qiao</td></tr>
<tr><td>15:20-15:50</td><td>Coffee Break</td><td colspan="2">Coffee Break (30 minutes)</td></tr>
<tr><td>Chair</td><td>Dong Li</td><td>Long Yu</td><td>Feng Li</td></tr>
<tr><td>15:50-16:30</td><td>Siwei Wang</td><td>Xiao Han</td><td>Chenchen Ma</td></tr>
<tr><td>16:30-17:30</td><td>Qian Lin</td><td>Yingxing Li</td><td>Haiqi Li</td></tr>
<tr><td>17:30-19:00</td><td>Dinner</td><td>Banquet
18:00-20:00</td><td colspan="3">Dinner (90 minutes)</td></tr>
</table>

January 5, 2026 - Monday

Time	Details	Chair
7:30-8:30	Breakfast (60 minutes)	
8:30-8:40	Openning Remark	
	Session 1	
8:40-9:20	**Ching-Kang Ing** Covariate Shift in Spatial Autoregressive Models	**Yundong Tu**
9:20-10:00	**Junhui Wang** Learning nonparametric graphical model on heterogeneous network-linked data	**Yundong Tu**
10:00-10:30	Coffee Break (30 minutes)	
	Session 2	
10:30-11:10	**Qianqian Zhu** Improving time series estimation and prediction via transfer learning	**Shaojun Guo**
11:10-11:50	**Zhentao Shi** Zero Variance Portfolio	**Shaojun Guo**
12:00-13:30	Lunch (90 minutes)	
	Session 3	
14:00-14:40	**Yuying Sun** Optimal Parameter-Transfer Learning by Time-Varying Model Averaging	**Bin Guo**
14:40-15:20	**Dalei Yu** Time-varying Model Averaging for General Loss Functions	**Bin Guo**
15:20-15:50	Coffee Break (30 minutes)	
	Session 4	
15:50-16:30	**Siwei Wang** Structural Break-driven Optimal Subsample Forecast Combination	**Dong Li**
16:30-17:30	**Qian Lin** Towards a mathematical understanding of deep learning	**Dong Li**
17:30-19:00	Dinner (90 minutes)	

January 6, 2026 - Tuesday

Time	Details	Chair
7:30-8:30	Breakfast (60 minutes)	
	Session 5	
8:40-9:20	**Bingyi Jing** Offline Reinforcement Learning: Learning from Datasets Without Interaction	**Qiwei Yao**
9:20-10:00	**Binyan Jiang** Hypergraph Embeddings	
10:00-10:30	Coffee Break (30 minutes)	
	Session 6	
10:30-11:10	**Yang Zu** Ratio-controlled screening for structural break predictive regression	**Rongmao Zhang**
11:10-11:50	**Xinling Xie** A Tale of Structural Instabilities for Sporadic Return Predictability	
12:00-13:30	Lunch	
14:00-14:40	**Free discussion**	
14:40-15:20		
15:50-16:30		
16:30-17:30		
18:00-20:00	Banquet	

January 7, 2026 - Wednesday

Time	Details	Chair
7:30-8:30	Breakfast (60 minutes)	
	Session 7	
8:40-9:20	**Jian Huang** Process-based Models for Learning Probability Distributions	**Yuhong Yang**
9:20-10:00	**Degui Li** Deep Neural Network Estimation of Nonparametric Panel Regression with Latent Group Structure	**Yuhong Yang**
10:00-10:30	Coffee Break (30 minutes)	
	Session 8	
10:30-11:10	**Jinyuan Chang** CP-factorization for high dimensional tensor time series and double projection iterations	**Haihan Tang**
11:10-11:50	**Zhou Yu** Dimension Reduction In the Era of AI	**Haihan Tang**
12:00-13:30	Lunch	
	Session 9	
14:00-14:40	**Anna Bykhovskaya** How Weak are Weak Factors? Uniform Inference for Signal Strength in Signal Plus Noise Models	**Ying Wang**
14:40-15:20	**Qing Yang** Generalized Linear Spectral Statistics of High-dimensional Sample Covariance Matrices and Its Applications	**Ying Wang**
15:20-15:50	Coffee Break (30 minutes)	
	Session 10	
15:50-16:30	**Xiao Han** Quantifying Cross-Domain Knowledge Distillation in the Presence of Domain shift	**Long Yu**
16:30-17:30	**Yingxing Li** Low-Rank and Sparse Network Regression	**Long Yu**
17:30-19:00	Dinner	

January 8, 2026 - Thursday

Time	Details	Chair
7:30-8:30	Breakfast (60 minutes)	
	Session 11	
8:40-9:20	**Qiying Wang** Self-weighted estimation for local unit root autoregression	**Qianqian Zhu**
9:20-10:00	**Yingqian Lin** Spurious Quantile Regressions and Variable Selection in Quantile Cointegrations	
10:00-10:30	Coffee Break (30 minutes)	
	Session 12	
10:30-11:10	**Ling Zhou** Spatial effect detection regression for large-scale spatio-temporal covariates	**Dan Yang**
11:10-11:50	**Chaohua Dong** Robust M-Estimation for Additive Nonparametric Cointegrating Models	
12:00-13:30	Lunch	
	Session 13	
14:00-14:40	**Weichi Wu** A Random Graph-based Autoregressive Model for Networked Time Series	**Han Yan**
14:40-15:20	**Xinghao Qiao** Time Series Gaussian Chain Graph Models	
15:20-15:50	Coffee Break (30 minutes)	
	Session 14	
15:50-16:30	**Chenchen Ma** The Factor Tree: A Data-Driven Approach to Regime Switching in High-Dimensions	**Feng Li**
16:30-17:30	**Haiqi Li** Beyond Principal Components: Likelihood-Based Estimation of Time-Varying Factor Models	
17:30-19:00	Dinner	

January 9, 2026 - Friday

Time	Details	Chair
7:30-8:30	Breakfast (60 minutes)	
	Session 15	
8:40-9:20	**Xinyang Yu** Dynamic Networks with Node Heterogeneity and Homophily	**Baiqing Wang**
9:20-10:00	**Dan Yang** Forecasting Global Economy with SIGMAR: Sparsity-Induced Global Matrix AutoRegressive Model	
10:00-10:30	Coffee Break (30 minutes)	
	Session 16	
10:30-11:10	**Han Yan** Deep Independent Component Analysis for Time Series with Invertible Neural Networks	**Xinling Xie**
11:10-11:50	**Baiqing Wang** Shattering Break Barriers: Inferential Theory for Group Factor Models Subject to Disruptive Breaks	
12:00-13:30	Lunch	
14:00-14:40	**Free discussion**	
14:40-15:20		
15:50-16:30		
16:30-17:30		
17:30-19:00	Dinner	

Titles and Abstracts

Covariate Shift in Spatial Autoregressive Models

Ching-Kang Ing
National Tsing Hua University

Covariate shift-where the covariate distribution di?ers between the training and target environments while the conditional response model remains invari-ant-poses substantial challenges for statistical learning under spatial dependence. This paper investigates covariate shift in the context of spatial autoregressive (SAR) models, a fundamental framework for modeling spatially correlated outcomes. We develop a general theory for importance-weighted estimation and prediction in SAR models under covariate shift, establishing conditions for consistency, asymptotic normality, and convergence rates of the mean-squared prediction error. Simulation results show that conventional estimators can deteriorate markedly under moderate shift, whereas the proposed approach delivers robust estimation and accurate prediction. This work provides one of the first systematic treatments of covariate shift for SAR models and o?ers practical tools for reliable inference in heterogeneous spatial environments.

Learning nonparametric graphical model on heterogeneous network-linked data

Junhui Wang
Chinese University of Hong Kong

Graphical models have been popularly used for capturing conditional independence structure in multivariate data, which are often built upon independent and identically distributed observations, limiting their applicability to complex datasets such as network-linked data. In this talk, we introduce a nonparametric graphical model that addresses these limitations by accommodating heterogeneous graph structures without imposing any specific distributional assumptions. The introduced estimation method effectively integrates network embedding with nonparametric graphical model estimation. It further transforms the graph learning task into solving a finite-dimensional linear equation system by leveraging the properties of vector-valued reproducing kernel Hilbert space. We will also discuss theoretical properties of the proposed method in terms of the estimation consistency and exact recovery of the heterogeneous graph structures. Its effectiveness is also demonstrated through a variety of simulated examples and a real application to the statistician coauthorship dataset.

Improving time series estimation and prediction via transfer learning

Qianqian Zhu
Shanghai University of Finance and Economics

There are many time series in the literature with high dimension yet limited sample sizes, such as macroeconomic variables, and it is almost impossible to obtain efficient estimation and accurate prediction by using the corresponding datasets themselves. This paper fills the gap by introducing a novel representation-based transfer learning framework for vector autoregressive models, and information from related source datasets with rich observations can be leveraged to enhance estimation efficiency through representation learning. A two-stage regularized estimation procedure is proposed with well established non-asymptotic properties, and algorithms with alternating updates are suggested to search for the estimates. Our transfer learning framework can handle time series with varying sample sizes and asynchronous starting and/or ending time points, thereby offering remarkable flexibility in integrating information from diverse datasets. Simulation experiments are conducted to evaluate the finite-sample performance of the proposed methodology, and its usefulness is demonstrated by an empirical analysis on 20 macroeconomic variables from Japan and another nine countries.

Zero Variance Portfolio

Zhentao Shi
Chinese University of Hongkong

When the number of assets is larger than the sample size, the minimum variance portfolio interpolates the training data, achieving a pathological zero variance. We show that if the weights of the zero variance portfolio are learned by a novel "Ridgelet" estimator, ?in a new test data the portfolio enjoys out-of-sample generalizability. It exhibits the double descent phenomenon, and attains optimal risk in the overparametrized regime when the number of assets dominates the sample size. In contrast, a "Ridgeless" estimator which invokes the pseudoinverse fails in-sample interpolation and diverges away from out-of-sample optimality. Extensive simulations and empirical studies demonstrate that our method performs well in large portfolio optimization.

Optimal Parameter-Transfer Learning by Time-Varying Model Averaging

Yuying Sun
Chinese Academy of Sciences

The presence of structural changes in economics and related fields poses challenges for out-of-sample forecasting, particularly when the sample size of the target data is limited. This paper proposes a novel optimal parameter transfer learning approach through time-varying model averaging (Transfer-TVMA), which enhances target model predictions by adaptively transferring potentially shared parameter information under structural changes and model uncertainty.A local forward-validation weight choice criterion is developed to select time-varying combination weights for candidate models based on different populations. The asymptotic optimality and convergence properties of the selected time-varying weights are established under regularity conditions. Beyond point forecasts, we further develop a conformal prediction interval algorithm based on Transfer-TVMA and provide its asymptotic validity without the exchangeability assumption. Simulation studies and an empirical application to exchange rate forecasting demonstrate the superior predictive performance of the proposed approach compared to existing competing methods.

Time-varying Model Averaging for General Loss Functions

Dalei Yu
Xi'an Jiaotong University.

We propose a unified time-varying model averaging approach that accommodates general loss functions, including Lin-lin loss and asymmetric squared error loss, to improve prediction performance under structural change. This flexibility enables averaging across diverse candidate models, such as time-varying coefficient quantile regression models. We develop a local forward-validation criterion to determine time-varying combination weights without the standard constraint of summing up to 1 and establish theoretical justifications previously unexplored in the literature. First, when all candidate models are misspecified, the proposed averaging prediction is asymptotically optimal in the sense of achieving the lowest possible prediction risk with a convergence rate. Second, we establish a novel convergence rate for time-varying weight consistency that does not depend on the extent of misspecification among the candidate models. Furthermore, we develop a time-varying sparsity-oriented importance learning procedure that consistently identifies the true predictor set. Monte Carlo simulations and empirical applications demonstrate superior finite-sample performance relative to existing model selection and averaging methods.

Structural Break-driven Optimal Subsample Forecast Combination

Siwei Wang
Hunan University

In the practice of economic and financial time series forecasting, structural breaks are pervasive, and while integrating pre- and post-break data has long been recognized to potentially enhance prediction accuracy compared to relying solely on post-break information, a consensus on effectively leveraging break information remains elusive. This paper addresses this gap by proposing a novel subsample forecast combination scheme: subsamples are constructed based on the identified most recent break, with a subsample tuning parameter governing subsample specifications (length and quantity), candidate forecasts are generated using parameter estimates from each subsample to summarize break-related information (e.g., magnitude, location), and forecast combinations are derived via weights that minimize a forward validation criterion alongside optimal subsample specification selection. Theoretical analysis establishes uniform consistency of estimated coefficients and asymptotic optimality of selected weights and subsample specifications; further, if correctly-specified models exist among candidate subsample forecasts, they are assigned all weights with probability approaching one. Numerical results from simulations and a real-data application to U.S. equity premium forecasting demonstrate the combination strategy's superior practical performance.

Towards a mathematical understanding of deep learning

Qian Lin
Tsinghua University

In our previous work, we introduced the adaptive feature approach to elucidate the superior performance of neural networks. This approach emphasizes the dynamic adjustment of feature representations to optimize model performance. In this study, we present several empirical and theoretical examples that validate the effectiveness of the adaptive feature approach.

Offline Reinforcement Learning: Learning from Datasets Without Interaction

Bingyi Jing
Chinese University of Hong Kong, Shenzhen

Reinforcement learning (RL) has revolutionized how artificial intelligence learns through trial-and-error interaction, achieving superhuman performance in games and simulations. However, this online learning paradigm faces fundamental limitations in real-world applications where exploration is costly, dangerous, or impractical. Offline reinforcement learning emerges as a transformative alternative-enabling agents to learn optimal behavior exclusively from pre-collected datasets, without any environment interaction during training. This talk introduces the principles, challenges, and applications of offline RL. We demonstrate offline RL's potential to leverage historical data for decision-making while respecting safety constraints.

Hypergraph Embeddings

Binyan Jiang
The Hong Kong Polytechnic University

Hypergraphs generalize graphs by allowing each edge, known as a hyperedge, to connect multiple vertices. Despite their significant advantages, hypergraph?embeddings have been underexplored compared to pairwise graphs due to the inherent complexity of the hypergraph topologies. Existing approaches often rely on fixed-dimensional embeddings, where the relative closeness among nodes is fixed, regardless of hyperedge order. This fixed-dimensional setting encourages heredity among hyperedges of different orders and fails to offer a flexible projection to capture the complex relationships among nodes. In this project, we propose a novel increasing dimensional embedding approach that jointly considers sparsity and node heterogeneity, including both degree heterogeneity and node heterogeneity in the latent dependencies among hyperedges of different orders. The proposed framework offers a more flexible approach to capturing diverse features of hypergraphs and could potentially provide new insights in different real applications.

Ratio-controlled screening for structural break predictive regression

Yang Zu
University of Macau

Predictive regression is a crucial tool for exploring return predictability. In this study, we introduce an efficient procedure for selecting and estimating active predictors and change points in high-dimensional structural break predictive regression. Our approach allows the number of change points to increase with the sample size and accommodates sparse active predictors that may be stationary or cointegrated. We begin by identifying the active predictors using a Sure Independence Canonical Screening (SICS) procedure. Next, we estimate the change points through a Ratio-Controlled Regression Screening (RRS) method. Finally, we reduce redundancy by eliminating unnecessary breakpoints and predictors using information criteria (IC). This approach allows for consistent estimation and selection of true breakpoints and active predictors. Our simulations and empirical studies demonstrate that the proposed procedure performs effectively.

A Tale of Structural Instabilities for Sporadic Return Predictability

Xinling Xie
Southwestern University of Finance and Economics

Return predictability has been one of the central research questions in finance for many decades. This talk discusses statistical inference on predictive regressions that account for structural instabilities to capture the sporadic predictive ability of potential predictors for the return series. This allows the return predictability to vary over time or across the state of the economy, respectively. Efficient estimation procedures and robust testing methods are developed to reveal the return predictability in the presence of structural instabilities that could be otherwise disguised. In the realm of estimation, we employ machine learning techniques, including the sequential estimation and adaptive group Lasso, to accurately identify the unknown breakpoints and thresholds. For testing purpose, we propose unified procedures based on the instrumental variable filtering technique to detect possible episodic predictability of a persistent predictor regardless of its persistence level. The asymptotic properties of the proposed methodologies are rigorously established and are further corroborated with their nice finite sample simulation performance. Finally, the empirical applications to U.S. stock returns underscore the practical relevance and effectiveness of our approaches in enhancing the return predictability.

Process-based Models for Learning Probability Distributions

Jian Huang
The Hong Kong Polytechnic University

In this talk, we present process-based models, diffusion and continuous normalizing flows, for learning probability distributions. We then show how to integrate foundation models to learn rich conditional distributions, either by using them as priors that regularize the generative process or as feature extractors that provide effective data representation. We illustrate these ideas with applications to conditional image synthesis, functional protein sequence design, and synthetic data augmentation.

Deep Neural Network Estimation of Nonparametric Panel Regression with Latent Group Structure

Degui Li
University of Macau

This talk introduces the nonlinear conditional mean function estimation for large panel data via deep neural networks (DNN). A latent group structure is imposed on heterogeneous regression functions, i.e., the regression structure is invariant within a group of subjects but varies over different groups. The main methodology is built on a combination of the DNN estimation method with ReLU activation and an easy-to-implement clustering algorithm, which consistently estimates the latent structure. In addition, we propose a post-grouping DNN estimation, which further improves the nonparametric estimation convergence. The large-sample properties for the developed methods are derived under some regularity conditions and numerical studies are given to examine the finite-sample performance.

CP-factorization for high dimensional tensor time series and double projection iterations

Jinyuan Chang
Southwestern University of Finance and Economics

We adopt the canonical polyadic (CP) decomposition to model highdimensional tensor time series. Our primary goal is to identify and estimate the factor loadings in the CP decomposition. We propose a one-pass estimation procedure through standard eigen-analysis for a matrix constructed based on the serial dependence structure of the data. The asymptotic properties of the proposed estimator are established under a general setting as long as the factor loading vectors are algebraically linear independent, allowing the factors to be correlated and the factor loading vectors to be not nearly orthogonal. The procedure adapts to the sparsity of the factor loading vectors, accommodates weak factors, and demonstrates strong performance across a wide range of scenarios. A tractable limiting representation of the estimator is derived, which plays a key role in the related inference problems. To further reduce estimation errors, we also introduce an iterative algorithm based on a novel double projection approach. We theoretically justify the improved convergence rate of the iterative estimator, and also provide the associated limiting distribution. All results are validated through extensive simulations and a real data application.

Dimension Reduction In the Era of AI

Zhou Yu
East China Normal University

From classical statistical Principal Component Analysis to modern large-scale deep neural network-based autoencoders, dimensionality reduction and feature extraction have consistently remained a focal point in the intersection of statistics and artificial intelligence. We aim to uncover the intrinsic connections between statistical and AI perspectives on dimensionality reduction and representation learning, and to systematically integrate statistical viewpoints with modern artificial intelligence methodologies. This will establish a theoretical and methodological framework for supervised nonlinear dimensionality reduction and representation learning, including comprehensive algorithmic procedures and error analysis. Finally, the framework will be validated through real-world applications.

How Weak are Weak Factors? Uniform Inference for Signal Strength in Signal Plus Noise Models

Anna Bykhovskaya
Duke University

In high-dimensional data analysis separating meaningful structure from noise is a central challenge. I will discuss four classical signal-plus-noise models factor models, spiked covariance matrices, Wigner matrices with low-rank perturbations, and canonical correlation analysis - with a focus on measuring the strength of the signals. Traditional Gaussian approximations provide reliable inference only when signals are strong enough, but they break down near the critical threshold where signal and noise become indistinguishable. I will present a new framework for constructing confidence intervals that remain valid uniformly across strong, weak, and critical regimes. The method is based on a universal transitional distribution and offers a practical tool for applied work. I will illustrate its performance with applications to macroeconomic and financial data.

Generalized Linear Spectral Statistics of High-dimensional Sample Covariance Matrices and Its Applications

Qing Yang
University of Science and Technology of China.

In this paper, we introduce the Generalized Linear Spectral Statistics (GLSS) of a high-dimensional sample covariance matrix S_n, denoted as $\operatorname{tr} f\left(S_n\right) B_n$, which effectively captures distinct spectral properties of S_n by incorporating an ancillary matrix B_n and a test function f. The joint asymptotic normality of GLSS associated with different test functions is established under mild assumptions on B_n and the underlying distribution, when the dimension n and sample size N are comparable. The convergence rate of GLSS is determined by $\sqrt{N/\operatorname{rank}\left(B_n\right)}$. Subsequently, we propose a novel functional projection approach based on GLSS for hypothesis testing on eigenspaces of "populationspiked" covariance matrices, showcasing a universality phenomenon in the magnitude of the spikes. The theoretical accuracy of our results established for GLSS and the advantages of the newly suggested testing procedure are demonstrated through various numerical studies.

Quantifying Cross-Domain Knowledge Distillation in the Presence of Domain shift

Xiao Han
University of Science and Technology of China

Abstract: Cross-domain knowledge distillation often suffers from domain shift. Although domain adaptation methods have shown strong empirical success in addressing this issue, their theoretical foundations remain underdeveloped. In this paper, we study knowledge distillation in a teacher--student framework for regularized linear regression and derive high-dimensional asymptotic excess risk for the student estimator, accounting for both covariate shift and model shift. This asymptotic analysis enables a precise characterization of the performance gain in cross-domain knowledge distillation. Our results demonstrate that, even under substantial shifts between the source and target domains, it remains feasible to identify an imitation parameter for which the student model outperforms the student-only baseline. Moreover, we show that the student's generalization performance exhibits the double descent phenomenon.

Low-Rank and Sparse Network Regression

Yingxing Li
Sun Yat-sen University

This paper analyzes spillover effects in spatial (network) models when measurement noise might contaminate the neighborhood (i.e. adjacency) matrix. We propose to adopt a low-rank and sparse structure to capture stylized network patterns in empirical datasets. We develop a flexible estimation framework via regularization techniques: a Least Absolute Shrinkage and Selection Operator (LASSO) penalty for the sparse component and a nuclear norm penalty for the low-rank component. We propose two estimators: (1) A two-stage procedure that first de-noises the adjacency matrix via regularization and subsequently integrates the purified network into a regression analysis, and (2) A single-step supervised Generalized Method of Moments (GMM) estimator that jointly estimates the regression parameters and refines the network structure. Simulation evidence underscores the superiority of our approach relative to conventional estimation protocols. In scenarios with noisy networks, our method reduces the root mean squared error (RMSE) of the estimate of spillover effects by around 20% compared to conventional GMM. This advantage is more significant when measurement errors are correlated with the observed outcomes and network contamination is econometrically endogenous. We apply our framework to the dataset in Besley and Case (1995) and demonstrate its practical utility. The decomposition not only improves estimation reliability but also generates granular insights for policy design.

Self-weighted estimation for local unit root autoregression

Qiying Wang
The University of Sydney

A new self-weighted least squares (LS) estimator is developed for local unit root (LUR) autoregression with heteroskedasticity. The proposed estimator has mixed Gaussian limit theory and the corresponding studentized statistic converges to a standard normal distribution free of the unknown localizing coefficient which is not consistently estimable. The estimator is super consistent with a convergence rate slightly below the $O(n)$ rate of LS estimation. The asymptotic theory relies on a new framework of convergence to the local time of a Gaussian process, allowing for the sample moments generated from martingales and many other integrated dependent sequences. A new unit root (UR) test in augmented autoregression is developed using self-weighted estimation and the methods are employed in predictive regression, providing an alternative approach to IVX regression. This is a joint work with Hu, Liu and Phillips.

Spurious Quantile Regressions and Variable Selection in Quantile Cointegrations

Yingqian Lin
Shanghai University of Finance and Economics

Quantile regression is an effective tool in modeling conditional distribution of economic and financial time series. This paper first investigates the spurious quantile regression phenomenon involving processes moderately deviated from a unit root (PMDURs) through numerical experiments. The main findings include large quantile correlation coefficient estimates, divergent significance test statistics, high R^2 and highly autocorrelated residuals indicated by very low Durbin-Watson statistics, complementing those discovered in spurious mean regressions (Lin and Tu, 2020). Unlike the recently proposed balanced regression (Ren et al., 2019; Lin and Tu, 2020) for the mean regression, this paper pioneers the use of the quantile partial correlation (Li et al., 2015, QPC) as a simple-to-implement robust inference measure for spurious quantile regressions. This is benefited from our finding that the induced QPC estimator is asymptotically normal and free of nuisance parameters. For the correlated but non-cointegrated quantile regressions and quantile cointegrations, the QPC estimator converges to a nonzero value in probability. Consequently, the proposed inference procedure can serve to perform variable selection in quantile regressions/cointegrations with PMDURs. Finally, the finite sample properties of the robust method are demonstrated through both Monte Carlo and real data examples.

Spatial effect detection regression for large-scale spatio-temporal covariates

Ling Zhou
Southwestern University of Finance and Economics

We develop a Spatial Effect Detection Regression (SEDR) model to capture the nonlinear and irregular effects of high-dimensional spatio-temporal predictors on a scalar outcome. Specifically, we assume that both the component and the coefficient functions in the SEDR are unknown smooth functions of location and time. This allows us to leverage spatially and temporally correlated information, transforming the curse of dimensionality into a blessing, as confirmed by our theoretical and numerical results. Moreover, we introduce a set of 0-1 regression coefficients to automatically identify the boundaries of the spatial effect, implemented via a novel penalty. A simple iterative algorithm, with explicit forms at each update step, is developed, and we demonstrate that it converges from the initial values given in the paper. Furthermore, we establish the convergence rate and selection consistency of the proposed estimator under various scenarios involving dimensionality and the effect space. Through simulation studies, we thoroughly evaluate the superior performance of our method in terms of bias and empirical efficiency. Finally, we apply the method to analyse and forecast data from environmental monitoring and Alzheimer's Disease Neuroimaging Initiative study, revealing interesting findings and achieving smaller out-of-sample prediction errors compared to existing methods.

Robust M-Estimation for Additive Nonparametric Cointegrating Models

Chaohua Dong
Zhongnan University of Economics and Law

This paper focuses on robust estimation for additive nonparametric models featuring a deterministic time trend and a combination of stationary and nonstationary explanatory variables. Robust estimators typically rely on loss functions, such as LAD, quantile loss, and Huberj's loss, whereas the nonsmoothness of these losses poses significant analytical challenges. To address this, we employ an innovative method proposed by Dong et al. (2025), approximating nonsmooth objectives via regular sequences in generalized functions. However, until now, it has only been applied to parametric models. Extending this framework to nonparametric settings is not travail; we derive the asymptotic distributions and convergence rates of the proposed estimators. Monte Carlo simulations confirm the methods validity and verify the limit distribution in theoretical part. An empirical application to Consumer Price Index (CPI) forecasting demonstrates its practical usefulness.

A Random Graph-based Autoregressive Model for Networked Time Series

Weichi Wu
Tsinghua University

Contemporary time series data often feature objects connected by a social network, which naturally induces temporal dependence among connected neighbors. The network vector autoregressive model is useful for describing the influence of linked neighbors, while its recent generalizations aim to separate influence and homophily. Existing approaches, however, require either correct specification of a time series model, accurate estimation of a network model, or both, and rely exclusively on least squares for parameter estimation. This paper proposes a new autoregressive model incorporating a flexible form for latent variables used to depict homophily. We develop a first-order differencing method for the estimation of influence, requiring only the influence part of the model to be correctly specified. When the homophily part is correctly specified, admitting a semiparametric form, we leverage and generalize the recent notion of neighbor smoothing for parameter estimation, bypassing the need to specify the generative mechanism of the network. We

develop new theory to show that all the estimated parameters are consistent and asymptotically normal. The efficacy of our approach is confirmed via extensive simulations and an analysis of a social media dataset.

Time Series Gaussian Chain Graph Models

Xinghao Qiao
The University of Hong Kong

We develop a novel time series Gaussian chain graph model that jointly captures symmetric conditional dependencies through undirected edges and asymmetric causal relations via directed edges among multivariate time series. In this paper, we first establish new identifiability conditions between undirected and directed edges by exploiting a group sparse plus group low-rank decomposition of the inverse spectral density matrix. This framework further motivates a frequency-domain learning algorithm that leverages the Whittle likelihood to fully recover the time series chain graph structure. Under mild regularity conditions, we demonstrate that the proposed method achieves asymptotic identifiability, consistency, and exact causal ordering and edge recovery. Extensive simulation studies confirm the effectiveness of the proposed approach. The practical usefulness is further illustrated through an application to U.S. macroeconomic data, revealing key monetary policy transmission mechanisms.

The Factor Tree: A Data-Driven Approach to Regime Switching in High-Dimensions

Chenchen Ma
Chinese Academy of Sciences

Threshold factor models are pivotal for capturing rapid regime-switching dynamics in high-dimensional time series, yet existing frameworks relying on a single pre-specified threshold variable often suffer from model misspecification and unreliable inferences. This paper introduces a novel factor tree model that integrates classification and regression tree (CART) principles with high-dimensional factor analysis to address structural instabilities driven by multiple threshold variables. The factor tree is constructed via a recursive sample splitting procedure that maximizes reductions in a loss function derived from the second moments of estimated pseudo linear factors. This procedure terminates when a data-driven information criterion signals no further improvement. To mitigate overfitting, a node merging algorithm further consolidates leaf nodes with identical factor representations. Theoretical analysis establishes consistency in threshold variable selection, threshold estimation, and factor space recovery, supported by extensive Monte Carlo simulations. An empirical application to U.S. financial data demonstrates the factor tree's effectiveness in capturing regime-dependent dynamics.

Beyond Principal Components: Likelihood-Based Estimation of Time-Varying Factor Models

Haiqi Li
Hunan University

This study develops a local maximum likelihood estimation (MLE) for high-dimensional factor models with time-varying factor loadings. We first propose five distinct identification conditions that enable asymptotically rotation-free estimation of both factors and their loadings. Under heteroskedastic errors, we establish consistency, derive convergence rates, and characterize the limiting distributions across different identification schemes. The relative efficiency of the local MLE compared to local principal component analysis is systematically examined. We further develop a nonparametric EM algorithm incorporating the reflection method for estimating time-varying factor loadings. The methodology is extended to a time-varying factor-augmented framework designed for binary outcome forecasting, where the resulting estimators are shown to be consistent and asymptotically normal. Monte Carlo simulations demonstrate that both the local MLE and the time-varying binary factor-augmented model exhibit strong finite-sample performance. In an empirical application to US financial markets, the proposed method demonstrates superior predictive accuracy in forecasting stock price directions compared to existing alternatives.

Dynamic Networks with Node Heterogeneity and Homophily

Xinyang Yu
London School of Economics and Political Science

Statistical modeling of network data is an important topic in various areas. Although many real networks are dynamic in nature, most existing statistical models and related inferences for network data are confined to static networks, and the development of the foundation for dynamic network models is still in its infancy. In particular, to the best of our knowledge, no attempts have been made to jointly address node heterogeneity and link homophily among dynamic networks. Being able to capture these network features simultaneously will not only bring new insights on understanding how networks were formed, but also provide more sophisticated tools for the prediction of a future network with statistical guarantees. In particular, our model accounts for link homophily associated with both observed traits and latent traits of the nodes. A novel normalized least squared loss based framework is constructed to generate stable estimations for the high dimensional parameters. The promising performance of the proposed model is further illustrated by various simulation and real data studies.

Forecasting Global Economy with SIGMAR: Sparsity-Induced Global Matrix AutoRegressive Model

Dan Yang
The University of Hong Kong

Jointly modeling and forecasting economic and financial variables across a large set of countries has long been a significant challenge. Two primary approaches have been utilized to address this issue: the vector autoregressive model with exogenous variables (VARX) and the matrix autoregression (MAR). The VARX model captures domestic dependencies but treats variables exogenous to represent global factors driven by international trade. In contrast, the MAR model simultaneously considers variables from multiple countries but ignores the trade network. In this paper, we propose an extension of the MAR model that achieves these two aims at once, i.e., studying both international dependencies and the impact of the trade network on the global economy. Additionally, we introduce a sparse component to the model to differentiate between systematic and idiosyncratic cross-predictability. To estimate the model parameters, we propose both a likelihood estimation method and a bias-corrected alternating minimization version. We provide theoretical and empirical analyses of the model's properties, alongside presenting intriguing economic insights derived from our findings.

Deep Independent Component Analysis for Time Series with Invertible Neural Networks

Han Yan
London School of Economics and Political Science

Independent Component Analysis (ICA) is a critical tool for identifying latent independent signals within multivariate observations, and has wide applications in fields such as financial time series and biomedical engineering. However, conventional ICA methods are primarily limited to linear transformations and assume independent observations, often failing to capture the complex dependence structures and time-dependent data. To address this, we develop a nonlinear ICA method for time series data. Our approach leverages an energy-distance-based loss function for indepdence measurement, and employs invertible neural networks (INN) to model nonlinear invertible mappings. We further derive theoretical properties of the INN function class, including its approximation capacity and pseodu-dimension, which are essential to establish the convergence rate of the nonlinear ICA estimator.

Shattering Break Barriers: Inferential Theory for Group Factor Models Subject to Disruptive Breaks

Baiqing Wang
Peking University

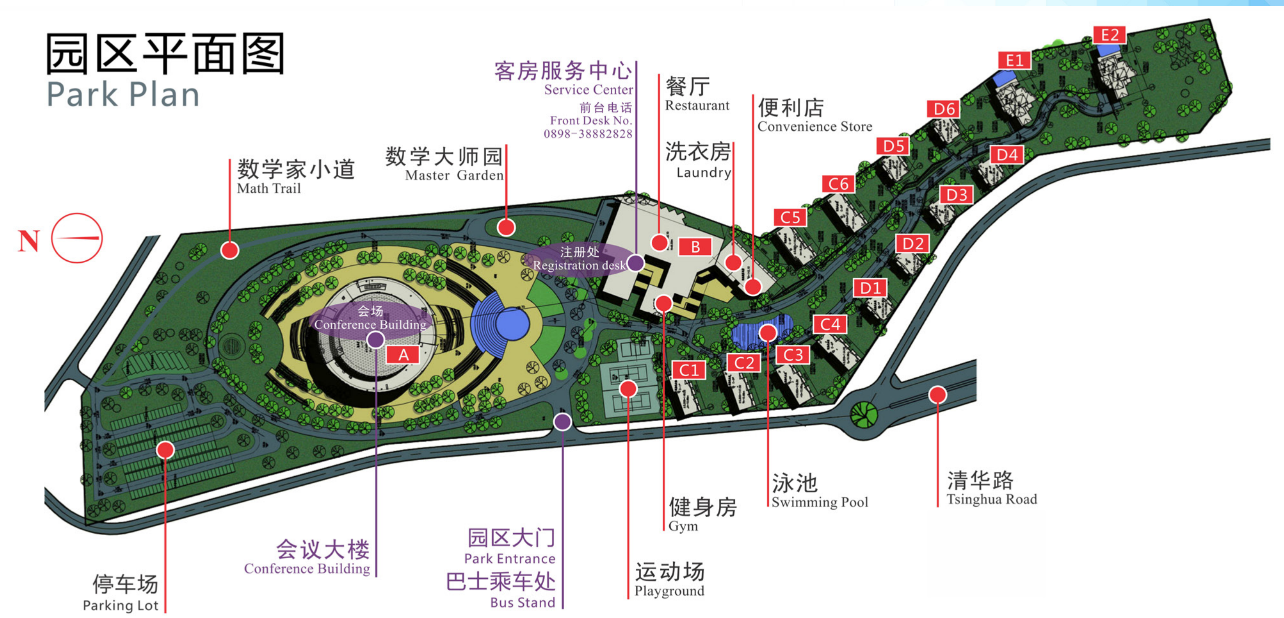
The homogeneous group structure in factor analysis is widely recognized for streamlining the clustering of high-dimensional time series, a property that directly enhances estimation precision. This study focuses on a grouped factor model where two critical components-latent group memberships and factor loadings-experience abrupt structural changes often triggered by real-world disruptions, including technological advancements, economic or financial crises, and policy adjustments. A critical finding of this research is that neglecting structural instability leads to three severe inferential issues: inconsistent recovery of the factor space, overestimation of the number of factors, and overestimation of the number of groups. To address these challenges and "shatter the break barriers," we propose a four-step procedure tailored to achieve consistent and efficient estimation. We first identify a "pseudo group structure" and estimate corresponding pseudo factors within a pseudo linear factor model-one that temporarily sets aside structural breaks. We next detect structural break points using shrinkage techniques applied to the pseudo factor regression. We then sequentially merge pseudo groups in an iterative manner, halting only when no further improvement is achieved according to an information criterion. Finally, we recover the regime-specific group factor space conditional on the estimated break points and regime-specific group structure. Theoretical properties of the proposed procedure, including consistency in break point estimation, accurate recovery of the true group structure, and reliable inference for the factor space, are established. These results are corroborated with both Monte Carlo simulations and an application to studying U.S. monthly macroeconomic dataset that demonstrate the superior performance of our procedure over alternative approaches in high-dimensional time series analysis.

Welcome to TSIMF

The facilities of TSIMF are built on a 23-acre land surrounded by pristine environment at Phoenix Hill of Phoenix Township. The total square footage of all the facilities is over 29,000 square meter that includes state-of-the-art conference facilities (over 10,000 square meter) to hold many international workshops simultaneously, two reading rooms of library, a guest house (over 10,000 square meter) and the associated catering facilities, a large swimming pool, gym and sports court and other recreational facilities.

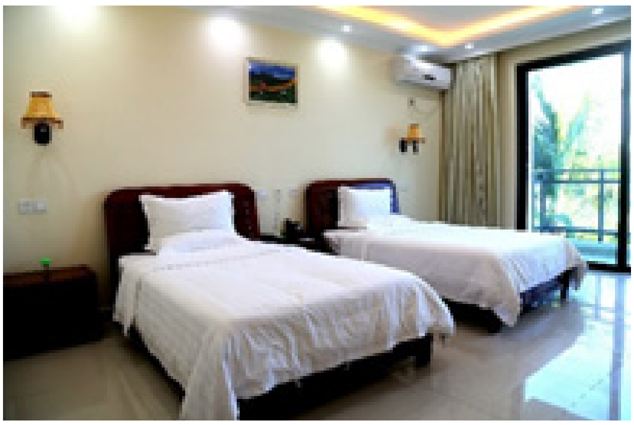
Management Center of Tsinghua Sanya International Forum is responsible for the construction, operation, management and service of TSIMF. The mission of TSIMF is to become a base for scientific innovations, and for nurturing of innovative human resource; through the interaction between leading mathematicians and core research groups in pure mathematics, applied mathematics, statistics, theoretical physics, applied physics, theoretical biology and other relating disciplines, TSIMF will provide a platform for exploring new directions, developing new methods, nurturing mathematical talents, and working to raise the level of mathematical research in China.

About Facilities



Registration

Conference booklets, room keys and name badges for all participants will be distributed at the front desk. Please take good care of your name badge. It is also your meal card and entrance ticket for all events.

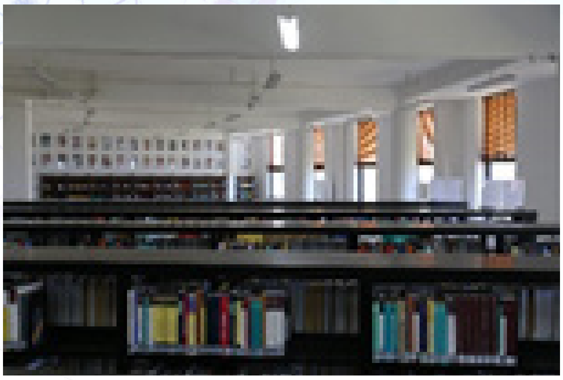


Guest Room

All the rooms are equipped with: free Wi-Fi (Password:tsimf123), TV, air conditioning and other utilities.

Family rooms are also equipped with kitchen and refrigerator.

Library



Opening Hours: 09:00am-22:00pm

TSIMF library is available during the conference and can be accessed by using your room card. There is no need to sign out books but we ask that you kindly return any borrowed books to the book cart in library before your departure.

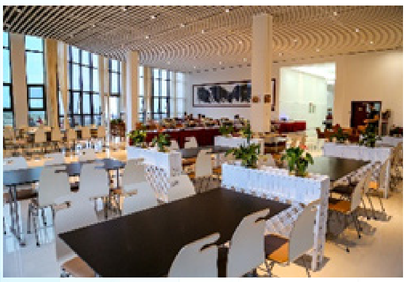
In order to give readers a better understanding of the contributions made by the Fields Medalists, the library of Tsinghua Sanya International Mathematics Forum (TSIMF) instituted the Special Collection of Fields Medalists as permanent collection of the library to serve the mathematical researchers and readers.

So far, there are 271 books from 49 authors in the Special Collection of Fields Medalists of TSIMF library. They are on display in room A220. The participants are welcome to visit.

Restaurant

All the meals are provided in the restaurant (Building B1) according to the time schedule.

Breakfast	07:30-08:30
Lunch	12:00-13:30
Dinner	17:30-19:00



Laundry

Opening Hours: 24 hours

The self-service laundry room is located in the Building(B1).

Gym

Opening Hours: 24 hours

The gym is located in the Building 1 (B1), opposite to the reception hall. The gym provides various fitness equipment, as well as pool tables, tennis tables etc.

Playground

Playground is located on the east of the central gate. There you can play basketball, tennis and badminton. Meanwhile, you can borrow table tennis, basketball, tennis balls and badminton at the reception desk.

Swimming Pool

Please enter the pool during the open hours, swimming attire and swim caps are required, if you feel unwell while swimming, please stop swimming immediately and get out of the pool. The depth of the pool is 1.2M-1.8M.

Opening Hours: 13:00-14:00 18:00-21:00

Free Shuttle Bus Service at TSIMF

We provide free shuttle bus for participants and you are always welcome to take our shuttle bus, all you need to do is wave your hands to stop the bus.

Destinations: Conference Building, Reception Room, Restaurant, Swimming Pool, Hotel etc.

Contact Information of Administration Staff

Location of Conference Affairs Office: Room 104, Building A

Tel: 0086-898-38263896

Conference Affairs :

Sarah 陈媛姗

Tel/Wechat:0086-130-2983-0780

Email: tsimf@tsinghua.edu.cn

Shouxi He 何守喜

Tel/Wechat:0086-186-8980-2225

Email: heshouxi@tsinghua.edu.cn

Location of Accommodation Affairs Office: Room 200, Building B1

Tel: 0086-898-38882828

Accommodation Manager: Ms. Li YE 叶莉

Tel/Wechat: 0086-139-7679-8300

Email: yel@tsinghua.edu.cn

*Reception duty hours: 7:00-23:00, chamber service please call: 0086-38882828 (exterior line) 80000 (internal line)

*Room maintainer night duty hours: 23:00-7:00, if you need maintenance services, please call: 0086-38263909 (exterior line) 30162 (internal line)

IT

Yuanhang Zhou 周远航

Tel/Wechat: 0086-133-6898-0169

Email: 13368980169@163.com

Director Assistant of TSIMF

Kai CUI 崔凯

Tel/Wechat: 0086- 136-1120-7077

Email :cuik@tsinghua.edu.cn

Director of TSIMF

Junpeng Zhu 朱俊鹏

Tel: 0086-136-1113-2615

Email: zjp@tsinghua.edu.cn

清华大学三亚国际论坛管理中心 | 三亚清华数学论坛管理中心
Tsinghua Sanya International Mathematics Forum (TSIMF)
0086-898-38882828
0086-898-38883896
0086-898-38883895
tsimf@tsinghua.edu.cn
572000
https://tsimf.tsinghua.edu.cn
海南省三亚市天涯区清华路 100 号，清华三亚国际数学论坛
No.100, Tsinghua Road, Tianya District, Sanya, Hainan, P. R. China.