

February 2, 2026 - Monday**Statistics at a crossroads: How it can revolutionize artificial intelligence****Rongling Wu(邬荣领)**

Beijing Institute of Mathematical Sciences and Applications(BIMSA)

Artificial intelligence (AI) is profoundly impacting science and society by applying algorithms and machine learning to enable machines to perform humanlike tasks. As a branch of mathematics, statistics lies at the core of AI and data science, but it is facing an unprecedented challenge with the surge of complex, heterogenous data across a variety of platforms. In a real sense, statistics is at a crossroads to leverage its central role in revolutionizing the foundational and fundamental framework of AI. In this talk, I will present several state-of-the-art statistical methods that have been widely used in AI across various fields. I will focus on how to develop statistically principled reasoning and theory to validate the application of AI and enhance its interpretability and sustainability. Our approach builds on statistical mechanics theory and methodology derived from interdisciplinary integration.

Whole- integumentary optical transparency assisted topological analysis for three-dimensional dynamics of acupoints and meridians**Yi Feng(冯昇)**

Fudan University

The branching patterns and spatial arrangement of somatic blood vessels and autonomic nerves underpin their function relevance, with particular importance at acupoints. Despite being recognized as sites enriched in neural terminals and bioactive factors, the fine microarchitecture of acupoints has remained elusive, largely owing to the absence of high-resolution three-dimensional (3D) molecular imaging and robust quantitative frameworks. To address this, we combined tissue clearing with topological data analysis to reconstruct acupoint neurovascular and functional networks in situ and to establish quantitative descriptors of their organization. In parallel, second near-infrared (NIR-II) imaging enabled real-time visualization of functional activity linked to these structural arrangements. From the intrinsic features of branching and neurovascular couplings, we derived four topological indices. In addition, dye-specific labeling of the whole integument was applied to map functional complexes during electroacupuncture. Together, these indices provide a systematic metric for quantifying neurovascular complexity and may serve as sensitive markers of subtle structural remodeling during physiological regulation or disease progression.

On the Topology of Transformations Between Data Structures

Konstantin Sorokin

National Research University “Higher School of Economics” (HSE)

3 Key words: Topological data analysis, structural evolution, persistent cobordism

While persistent homology has proven to be a powerful tool for analyzing the shape of data, a formal topological framework for understanding the dissimilarity between evolving data structures remains an open area of research. This work introduces "persistent cobordism," a novel method that extends the principles of persistence to the study of transformations between data objects such as graphs, simplicial complexes, and point clouds. By analyzing the topology of the "space of changes" between structures, our approach provides a topological characterization of local geometric distortions in data, both in terms of similarity and dissimilarity. We will explore the theoretical foundations of persistent cobordism, contrast it with current methodologies, and discuss its computational feasibility and implications for the future of topological data analysis.

Semiparametric regression, Dirichlet process Gaussian mixture model, Hamiltonian Monte Carlo

Ke Deng(邓柯)

Tsinghua University

3 Key words: A Semiparametric Bayesian Method for Sufficient Dimension Reduction

This work proposes a novel semiparametric Bayesian approach for statistical inference of the central subspace in the problem of sufficient dimension reduction. Unlike conventional Bayesian approaches for sufficient dimension reduction that model the conditional distributions of the response variable given the projected predictive variables, the new approach chooses to model their joint distribution instead via a Dirichlet process Gaussian mixture model, leading to both conceptual simplicity and computational convenience. Posterior consistency of the proposed approach is established under the framework of Schwartz's theorem. A Monte Carlo strategy based on the Gibbs sampler and geodesic Monte Carlo is developed for efficient posterior sampling. Both simulation studies and real data applications confirm the advantages of the proposed approach over existing Bayesian and frequentist methods.

Persistent representation theory

Jian Liu(刘健)

Chongqing University of Technology

3 Key words: Persistence group, persistent representation, decomposition theorem

In this talk, we introduce the notion of persistence representations and show that irreducible subrepresentations exhibit a birth–death behavior analogous to that of generators in persistent homology. As a consequence, irreducible persistence representations supported on intervals naturally correspond to bars. We establish a decomposition theorem stating that any finite persistence

representation decomposes as a direct sum of irreducible persistence representations supported on intervals, thereby generalizing the classical structure theorem for persistence modules. We further interpret persistence representations as graded modules over a polynomial ring induced by the shift operator and analyze the evolution of irreducible components in persistent regular representations.

Iterated Integrals on the Digraphs

Yunpeng Zi (白云鹏)
Shandong University

3 Key words: Iterated Integral, Digraph, GLMY Homology

Iterated integral is a classical geometric structure on the smooth manifolds. It introduced a interesting hopf algebra which depends on the homotopy type of the manifold and annihilated by some ideal of the group algebra of π_1 . In this research we introduced the iterated integral on the digraphs based on the GLMY theory of G,L,M and Yau. We proved the computation properties of this structure and induced some interting algebraic structures to the digraphs.

Matrix-weighted networks for modeling multidimensional dynamics

Yu Tian (田雨)
Max Planck Institute for the Physics of Complex Systems

3 Key words: Matrix-Weighted Networks, Network Coherence, Multidimensional Consensus and Random walks

Networks are powerful tools for modeling interactions in complex systems. While traditional networks use scalar edge weights, many real-world systems involve multidimensional interactions. For example, in social networks, individuals often have multiple interconnected opinions that can affect different opinions of other individuals, which can be better characterized by matrices. We propose a general framework for modeling such multidimensional interacting dynamics: matrix-weighted networks (MWNs). We present the mathematical foundations of MWNs and examine consensus dynamics and random walks within this context. Our results reveal that the coherence of MWNs gives rise to nontrivial steady states that generalize the notions of communities and structural balance in traditional networks.

From Geometric Hydrodynamics to Periodic Geodesics on Manifolds of Mappings

Levin Maier (Online)
University Heidelberg

In this talk, we begin by recalling Arnold's geometric formulation of hydrodynamics and then extend this framework to a broader class of Hamiltonian systems, incorporating various PDEs arising in mathematical physics. This motivates the study of infinite-dimensional manifolds and, in particular,

half Lie groups: topological groups in which right multiplication is smooth while left multiplication is only continuous. Important examples include groups of (H^s) - or (C^k) -diffeomorphisms of compact manifolds.

Within this setting, we establish several Hopf-Rinow type theorems for right-invariant magnetic systems and for certain Lagrangian systems on half Lie groups, thereby extending recent results of Bauer-Harms-Michor from the case of geodesic flows to this more general context. Finally, we show that any half Lie group with non trivial π_1 equipped with a strong Riemannian metric necessarily admits a periodic geodesic in each homotopy class.

February 3, 2026 - Tuesday

A Spectator's View on the Manifold Hypothesis

Tuschmann Wilderich

Karlsruhe Institute of Technology (KIT)

At its core, the Manifold Hypothesis asserts that real-world, high-dimensional data is not uniformly or randomly distributed throughout its high-dimensional “ambient” space, but concentrated on or near a low-dimensional manifold (or a collection of manifolds) embedded within that high-dimensional ambient space. In my talk, I will discuss reasons and facts that speak for as well as against this hypothesis and also address geometric alternatives.

The theory of Vietoris-Rips complexes

Henry Adams

University of Florida

3 Key words: Persistent homology, Vietoris-Rips complexes

I will survey what is known (and mostly unknown) about Vietoris-Rips complexes of manifolds. We might consider a dataset to be quite nice if it were densely sampled from a manifold. But we do not know what the Vietoris-Rips persistent homology of manifolds looks like, typically! I will describe what is known about Vietoris-Rips complexes of the circle, spheres, ellipsoids, tori, and manifolds, and advertise many open questions.

Topological Zigzag Spaghetti for Diffusion on Graphs

Yulia Gel

Virginia Tech

Diffusion models have recently emerged as a new powerful machinery for generative artificial intelligence on graphs, with applications ranging from drug design to knowledge discovery. However, despite their high potential, most, if not all, currently existing graph diffusion models are limited in their ability to holistically describe the intrinsic *higher-order* topological graph properties, which obstructs model generalizability and adoption for downstream tasks. We propose to address this fundamental challenge and extract the latent salient topological graph descriptors at different resolutions by leveraging zigzag persistence. We develop a new computationally efficient topological summary, zigzag spaghetti (ZS), which delivers the most inherent topological properties *simultaneously over a sequence of graphs at multiple resolutions*. We derive theoretical stability guarantees of ZS and present the first attempt to integrate dynamic topological information into graph diffusion models.

Our extensive experiments on graph classification and prediction tasks suggest that ZS has a high promise not only to enhance performance of graph diffusion models, but also to substantially

booster model robustness under uncertainties. This is a joint work with Yuzhou Chen, UC Riverside.

Topological deep learning for speech recognition

Yifei Zhu(朱一飞)

Southern University of Science and Technology

3 Key words: speech processing, machine learning, persistent homology

Human speech data is an important type of time series data with diverse applications in artificial intelligence. I'll report on progress with incorporating topological methods to machine learning of such data, in three stages: (1) a combination of feature extraction, through time-delay embedding and persistent homology, and basic machine learning algorithms, (2) an enhancement on state-of-the-art neural networks for speech processing, by concatenating topological features with final hidden states, and (3) a topology-informed convolutional kernel for speech recognition tasks. This is joint work with Zeyang Ding, Pingyao Feng, Qingrui Qu, Siheng Yi, Zhiwang Yu, and Haiyu Zhang.

Effective Resistance and Random Walks on Simplicial Complexes

Sarah Beth Percival

University of New Mexico

3 Key words: Effective resistance, simplicial complexes, random walks

The effective resistance originates from electric circuit analysis and becomes an important concept in graph theory due to its connection to random walks. Using previous work that describes electrical circuit laws in terms of chains and cochains, I will provide a generalization of the relationship between commute time in random walks and effective resistance on graphs to simplicial complexes.

Computing singular simplicial homologies of digraphs and quivers

Matthew Burfitt

Beijing Institute of Mathematical Sciences and Applications(BIMSA)

3 Key words: Algebraic topology, Quivers, Computation

Significant success has been obtained in applying the homology of the directed flag complex to study digraphs arising as networks within numerous scientific disciplines. Directed cliques enjoys relative ease of computation when compared to other digraph homologies, making them preferable for use in applications concerning large networks. By extending the ideas of singular simplicial homology to quivers in categories of different morphism types, several new singular simplicial homology theories have recently been constructed by Li, Muranov, Wu and Yau.

Generally, computationally efficient homologies for quivers have not previously been widely considered. I will describe two homotopy theories of quivers necessary to derive functors that

realise isomorphisms between the singular simplicial quiver homologies and the homologies of certain spaces. The simplicial chains of these spaces arise in a conveniently compact form that is at least as convenient as the directed flag complex for computations. Moreover, the constructions are natural with respect to the isomorphisms on homology making them suitable for application in conjunction with persistent homology for practical use. To end the talk I will highlight progress on the construction of efficient algorithms.

A mathematical strategy to map the genetic architecture of a single diplotype towards genome editing

Yangyang Bian(边洋洋)
Beijing Forestry University

Quantitative genetics is essential for genetic dissection of complex traits, yet its population-based theory fails to comprehend genetic architecture owned by a single individual. Here, we develop a diplotype-based mathematical strategy to assemble all genome loci into omnigenic interactome networks. These networks can not only be used to dissect the genetic architecture of a single individual, but also leverage dominance, epistasis, and pleiotropy to be more interpretable bidirectional, signed, and weighted interactions. While traditional approaches can only estimate genetic parameters at the population level, our strategy can model and characterize how each allele acts and interacts with every other allele throughout the whole genome for a single individual, thus facilitating its genome editing and genome engineering. We apply this strategy to analyze transcriptomic data of two diplotyped cultivars of a woody plant, interpreting the genetic mechanisms underlying this species' cold resistance and inter-organ communication. Individualized genetic findings can be readily transformed to precision breeding or precision medicine. Our strategy can be generalized to establish the foundation of individualized quantitative genetics, a new theory that can make genetics more transformable.

Interval Multiplicities and the Essential-Cover Technique

Enhao Liu(刘恩豪)
Kyoto University Institute for Advanced Study (KUIAS)

3 Key words: multi-parameter persistence, interval multiplicities, essential cover

Interval modules play a fundamental role in persistent homology, as intervals encode the lifetimes of topological features emerging from data and admit simple characterizations. In this talk, we present an explicit formula for computing interval multiplicities—i.e., the multiplicities of interval modules (summands)—in persistence modules defined over arbitrary finite posets. Our formula thus generalizes the well-known one-parameter persistence formula that relates multiplicities of birth-death pairs to persistent Betti numbers. Next, we introduce the essential-cover technique, which enables efficient computation of interval multiplicities by transforming persistence modules over original posets to modules over simpler, algorithmically tractable posets, such as zigzag posets, where fast decomposition algorithms are available. This technique makes it possible to compute interval multiplicities directly from the filtration level of topological spaces. The content of this talk is mainly based on the joint work with Hideto Asashiba.

February 4, 2026 - Wednesday

The integral Weyl invariants of Lie groups

Haibao Duan(段海豹)

Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences

3 Key words: Lie group; Weyl invariants; Cohomology

Let $\text{Inv}(G)$ denote the ring of integral Weyl invariants of a compact connected Lie group G . Abstracting its common algebraic features, we introduce the notion of an integral domain with multiplicative content. Based on an algebraic classification of such domains, we provide a uniform presentation of the ring $\text{Inv}(G)$, applicable to all Lie groups.

As applications, we offer a canonical description of the mod p invariant ring $\text{Inv}(G; F_p)$, and determine the structure of $\text{Inv}(G)$ for the spinor group $G = \text{Spin}(n)$.

Robust Model Reconstruction Based on the Topological Understanding of Point Clouds Using Persistent Homology

Hongwei Lin(蔺宏伟)

Zhejiang University

3 Key words: Surface reconstruction, Persistent homology, Representative cycle

Reconstructing models from unorganized point clouds presents a significant challenge, especially when the models consist of multiple components represented by their surface point clouds. Such models often involve point clouds with noise that represent multiple closed surfaces with shared regions, making their automatic identification and separation inherently complex. In this paper, we propose an automatic method that uses the topological understanding provided by persistent homology, along with representative 2-cycles of persistent homology groups, to effectively distinguish and separate each closed surface. Furthermore, we employ Loop subdivision and least squares progressive iterative approximation (LSPIA) techniques to generate high-quality final surfaces and achieve complete model reconstruction. Our method is robust to noise in the point cloud, making it suitable for reconstructing models from such data. Experimental results demonstrate the effectiveness of our approach and highlight its potential for practical applications.

Making sense of spatial transcriptomics: from statistical foundations to biological insights

Yuehua Cui(崔跃华)

Michigan State University

3 Key words: Spatial transcriptomics; Cell type deconvolution; Spatially variable genes

Spatial transcriptomics has revolutionized our ability to study gene expression within intact tissues,

but realizing its full potential depends on rigorous statistical and computational modeling. In this talk, I will present recent progress and open challenges in developing statistical frameworks for spatial transcriptomics, focusing on three key areas. First, I will discuss spatial cell type deconvolution, which aims to infer cellular composition from spatially resolved data with multicellular resolution. Second, I will cover methods for detecting cell type-specific spatially variable genes (SVGs), enabling the discovery of context-dependent transcriptional patterns. Third, I will introduce an emerging contrastive learning framework for spatial domain detection that integrates spatial cell type information to enhance domain identification. Together, these advances demonstrate how principled statistical modeling can translate experimental complexity into biological insight. I will conclude by highlighting future opportunities for spatial transcriptomics analysis driven by ongoing innovations in biotechnology.

Deep Generative Modeling of Facial Morphology from Genetic Variants

Wenting Zhao(赵雯婷)

Institute of Forensic Science of China (IFSC)

3 Key words: Genetics variants, 3D facial morphology, diffusion model

The variance in facial phenotype not only defines individual uniqueness but is also implicated in numerous genetic syndromes. Unraveling the genetic architecture governing this diversity remains a significant challenge, primarily due to the high dimensionality and intricate complexity of craniofacial morphology. Existing approaches to facial phenotyping vary across disciplines: anthropology and genetics widely utilize 2D landmarks and 3D point clouds, while computer vision employs embeddings for recognition and topology for animation. In this work, we leverage deep neural networks to establish a highly non-linear mapping between genetic variants and 3D facial morphology. Our model enables bi-directional cross-modal prediction, facilitating both the generation of 3D facial morphology from genotype and the prediction of relevant genetic variants from a facial phenotype. Additionally, we employ idopnetwork to conduct in-depth analyses of these genetic variants, identifying key drivers of facial morphology.

February 5, 2026 - Thursday

Presentations of mapping class groups and applications to cluster algebras from surfaces

Fang Li(李方)
Zhejiang University

3 Key words: Surface, mapping class group, cluster algebra

In this talk, we give presentations of the mapping class groups of marked surfaces stabilizing boundaries for any genus. Note that in the existing works, the mapping class groups of marked surfaces were the isotopy classes of homeomorphisms fixing boundaries pointwise. The condition for stabilizing boundaries of mapping class groups makes the requirement for mapping class groups to fix boundaries pointwise to be unnecessary.

As an application of presentations of the mapping class groups of marked surfaces stabilizing boundaries, we obtain the presentation of the cluster automorphism group of a cluster algebra from a feasible surface (S, M) .

Lastly, for the case (1) 4-punctured sphere, the cluster automorphism group of a cluster algebra from the surface is characterized. Since cluster automorphism groups of cluster algebras from those surfaces were given in the cases (2) the once-punctured 4-gon and (3) the twice-punctured digon, we indeed give presentations of cluster automorphism groups of cluster algebras from surfaces which are not feasible.

This is a joint work with Jinlei Dong.

Comparative Analysis of Exact Methods for Testing Equivalence of Prevalences in Bilateral and Unilateral Combined Data with and without Assumptions of Correlation

Changxing Ma(马长兴) (Online)
University at Buffalo

3 Key words: Equivalence of Prevalences, Exact methods, Testing

In clinical studies focusing on paired body parts, diseases can manifest on either both sides (bilateral) or just one side (unilateral) of the organs. Consequently, the data in these studies may consist of records from both bilateral and unilateral cases. There are two different methods of analyzing the data. One of the methods is assuming that the pair of measurements from the same subject are independent, while the other considers the correlation between paired organs. In terms of homogeneity test of proportions, asymptotic methods have been proposed given a moderate size of data. This article extends the existing work by proposing exact methods to deal with the scenarios when the sample size is small and asymptotic methods perform poorly. The impact of the correlation assumption is also explored. Among the proposed methods, calculating p-values by replacing unknown parameters with estimated values while accounting for the correlation is recommended based on its satisfactory type I error controls and statistical powers. The proposed methods are

applied to three real examples for illustration.

Beyond Graphs: Topological Deep Learning for Higher-Order Relational Systems

Mustafa Hajj

University of California, San Francisco

Graphs have become the default language for relational learning. But many of the systems we care about—physical fields, interacting particles, biological processes, scientific simulations, and structured knowledge—are not fundamentally graph-structured. Their interactions are multi-way, constrained, and often governed by conservation laws or geometric structure. For these systems, pairwise edges are not just insufficient—they are the wrong abstraction.

In this talk, I argue that topological deep learning provides a natural and necessary generalization of graph-based learning to higher-order relational systems. By lifting learning from nodes and edges to cells, cochains, and operators on combinatorial complexes, we move from learning correlations to learning structure. Boundary operators, Hodge Laplacians, and sheaf-based constraints are not added for elegance—they encode the grammar of interaction, flow, and consistency that many real systems obey.

I will show how graph neural networks emerge as a special case of a broader topological message-passing framework, and how higher-order architectures enable multi-scale reasoning, physically meaningful inductive biases, and operator learning on structured domains. Through examples from scientific machine learning and physics-informed modeling, I will highlight how topology allows us to design models that respect conservation, locality, and hierarchy—without hard-coding equations.

The central message is simple: if intelligence is relational, then learning must be topological. Moving beyond graphs is not about adding complexity—it is about choosing the right objects to learn on.

LLM-Based Multi-Agent System and Simplicial Self-Supervised Learning Model for Regional Cancer Prevalence Estimation Using Satellite Imagery

Yuzhou Chen(陈宇舟) (Online)

University of California, Riverside

3 Key words: GeoAI, Large Language Models, Precision Population Health

Traditional cancer rate estimations are often limited in spatial resolutions and lack considerations of environmental factors. Satellite imagery has become a vital data source for monitoring diverse urban environments, supporting applications across environmental, socio-demographic, and public health domains. However, while deep learning (DL) tools, particularly convolutional neural networks, have demonstrated strong performance in extracting features from high-resolution imagery, their reliance on local spatial cues often limits their ability to capture complex, non-local, and higher-order structural information. To overcome this limitation, we propose a novel LLM-based multi-agent coordination system for satellite image analysis, which integrates visual and contextual reasoning through a simplicial contrastive learning framework (Agent-SNN). Our Agent-SNN

contains two augmented superpixel-based graphs and maximizes mutual information between their latent simplicial complex representations, thereby enabling the system to learn both local and global topological features. The LLM-based agents generate structured prompts that guide the alignment of these representations across modalities. Experiments with satellite imagery of Los Angeles and San Diego demonstrate that Agent-SNN achieves significant improvements over state-of-the-art baselines in regional cancer prevalence estimation tasks.

Some optimizations in computing persistent Laplacian and beyond

Rui Dong(董瑞)
Vrije Universiteit Amsterdam

3 Key words: Persistent Laplacian, Cheeger inequality, topological data analysis

Clearing method is a well-known algorithmic optimization of computing persistent homology in TDA. In this talk we will introduce the application of clearing into the computation of persistent Laplacian. Besides that, we will show some relationship between the minimal nonzero eigenvalue of up persistent Laplacian and that of up combinatorial Laplacian.

Universal Lie algebra weight systems and graph invariants

Zhuoke Yang(杨卓科)
Beijing Institute of Mathematical Sciences and Applications(BIMSA)

To a finite type knot invariant, a weight system can be associated, which is a function on chord diagrams satisfying so-called 4-term relations. In particular, a weight system can be associated with any metrized Lie algebra. We extend these Lie algebra weight systems to permutations and provide a recursion for its computation. Recently, we showed that the leading term of the universal so weight system also becomes the chromatic polynomial under a specific substitution.

This work is based on the joint work with S. Lando and M. Kazaryan.

Data-Driven Approach to the Coefficients of the sl_2 Weight System

Usanov Sergei
National Research University “Higher School of Economics” (HSE)

3 Key words: Knot Invariants, Machine Learning, Graph Theory

The search for powerful invariants to distinguish knots is a central problem in the knot theory. A key tool in this area is the sl_2 weight system, which transforms knot diagrams into polynomials. While the structure of these polynomials is well-studied, finding explicit combinatorial formulas for their coefficients remains a significant challenge.

In this talk, we address this problem for the next unknown coefficient, a_3 . We demonstrate how techniques from data science and machine learning can be successfully applied to this pure mathematical question. By framing the problem as a regression task on a space of graphs, we use

computational data to derive a precise candidate formula for a_3 as a linear combination of subgraph-counting functionals.

We will discuss the theoretical foundation that justifies this data-driven approach, relying on reconstruction theorems from finite restrictions. This work provides a compelling case study at the intersection of low-dimensional topology and artificial intelligence, showcasing how machine learning can serve as a powerful guide for mathematical discovery.

LLM-Empowered Representation Learning for Emerging Item Recommendation

Ziying Zhang(张子颖)

Tsinghua University

In this work, we tackle the challenge of recommending emerging items, whose interactions gradually accumulate over time. Existing methods often overlook this dynamic process, typically assuming that emerging items have few or even no historical interactions. Such an assumption oversimplifies the problem, as a good model must preserve the uniqueness of emerging items while leveraging their shared patterns with established ones. To address this challenge, we propose EmerFlow, a novel LLM-empowered representation learning framework that generates distinctive embeddings for emerging items. It first enriches the raw features of emerging items through LLM reasoning, then aligns these representations with the embedding space of the existing recommendation model. Finally, new interactions are incorporated through a meta-learning procedure to refine the embeddings. This enables EmerFlow to learn expressive embeddings for emerging items from only limited interactions. Extensive experiments across diverse domains, including movies and pharmaceuticals, show that EmerFlow consistently outperforms existing methods.

February 6, 2026 - Friday**Topologically Interpretable Graph Learning via Persistent Rationale Filtration**

Jie Gao (Online)
Rutgers University

Many applications of graph neural networks benefit from generating an interpretation alongside a prediction. I'll present recent work on interpretable graph learning that leverages persistent homology to identify persistent rationale subgraphs. The main idea is to learn a filtration ordering of input graphs that separates the rational subgraphs from the remaining graph elements. For this purpose, we introduce a self-adjusted topological constraint, termed topological discrepancy, to enforce a persistent topological distinction between rationale subgraphs and irrelevant counterparts. We provide theoretical guarantees that, under specific conditions, our loss function is uniquely optimized by the ground truth. Experimental results show that our approach improves upon state-of-the-art methods in both predictive accuracy and interpretability.

This is joint work with Cheng Xin from Rutgers and Fan Xu, Xin Ding, Jiaxin Ding from Shanghai Jiaotong University.

Primitive Path Homology: An Algebraic Topology Approach for the Quantitative Characterization of Graph Pangenomes toward Population Genetic Analysis, Demonstrated in *Saccharomyces cerevisiae*

Qi Wu(吴琦)
Beijing Institute of Mathematical Sciences and Applications(BIMSA)

A central task in population genetics is to identify genetic diversity in a population containing a number of individuals.

In recent years, with the development of the third generation sequencing (TGS) technology, pan-genome research has become a hot topic. Although graphical representation has been a popular way to represent the pangenome, few works have attempted to describe it in a more mathematical way.

In this paper, we used 79 high-quality assembly data of third-generation sequencing in *Saccharomyces cerevisiae* to construct the graph pangenome, and introduced the Primitive Path Homology in algebraic topology to quantitatively represent the pan-genome.

We further made an interesting attempt to conduct a population genetic analysis of this resulting dataset of topological features of the graph pangenome.

We found that there was good agreement between the obtained results and the biological context.

In fact, this study has developed a method for population genetic analysis of the genetic diversity of genome structural variation.

An introduction to knot data analysis

Fengchun Lei(雷锋春)

Beijing Institute of Mathematical Sciences and Applications(BIMSA)

Topological data analysis (TDA), which has flourished in recent years, has wide applications in various fields such as biomedicine, materials science, artificial intelligence, and industrial manufacturing. Knot theory is an important branch of topology and has many applications in the total structure analysis of data, but it is rarely seen in the analysis of local structures. The main reason is that knot theory focuses more on the topological properties of knots; knots are equivalent before and after deformation, and the topological properties of knots are insensitive to local changes and continuous deformations. However, small local changes in proteins (such as α -helices and β -sheets) might cause changes in protein structure, thus affecting the expression the function of proteins.

In the talk, I will explain the idea of a novel method of data analysis, "multiscale Gaussian linking integral (mGli)". Its significant advantage lies in its ability to accurately characterize the entanglement of a set of non-intersecting curves (potentially including open branches) in space at different scales. It provides efficient local and quantitative analysis, recovering the local and global topological properties of knots and links at sufficiently large scales. This overcomes the limitations of traditional knot theory in providing precise quantitative analysis of complex biological systems (such as DNA, RNA, and proteins), opening a new direction: knot data analysis. This is a joint work with Li Shen, Hongsong Feng, Fengling Li, Fengchun Lei, Jie Wu and Guo-Wei Wei.