香港教育大學
The Education University of Hong Kong

Department of Mathematics and Information Technology

Symposium on

Applied Mathematics and Data Science

18 – 19 July 2023

Programme Schedule

Organized by:
Department of Mathematics and Information Technology,
Faculty of Liberal Arts and Social Sciences,
The Education University of Hong Kong (EdUHK)

Programme Schedule

- **Format:** Hybrid (Face-to-face in EdUHK and online via Zoom)

(All times are displayed in GMT + 8)

Day 1: 18 July 2023 (Tue)		
Time	**Event (Lady Ivy Wu Lecture Theatre, D1-LP-04)**	
09:30 – 09:50	Registration	
09:50 – 10:00	Welcoming Speech and Opening Ceremony	
10:00 – 11:00	Keynote Speech	
	Room A (D2-LP-03)	***Room B (D2-LP-04)***
11:00 – 12:30	Parallel Session #1 *Recent Advances in Fluid Dynamics and Some Related Models*	Parallel Session #2 *Theory and Application of Nonlinear PDEs*
12:30 – 14:30	Lunch Break	
14:30 – 16:00	Parallel Session #3 *Recent Progress on Applied Analysis and its Applications*	Parallel Session #4 *Recent Developments on Nonlinear PDEs in Fluids and Waves*

Day 2: 19 July 2023 (Wed)			
Time	**Event (Lady Ivy Wu Lecture Theatre, D1-LP-04)**		
10:00 – 11:00	Keynote Speech		
	Room A (D2-LP-03)	***Room B (D2-LP-04)***	***Room C (D2-LP-08)***
11:00 – 12:30	Parallel Session #1 *Recent Advances in AI for Texts and Images*	Parallel Session #2 *Recent Advances in Survival and Reliability Analysis*	Parallel Session #3 *Recent Developments on Data Science and its Applications*
12:30 – 14:30	Lunch Break		
14:30 – 16:00	Parallel Session #4 *Statistics, Data Science and their Applications*	Parallel Session #5 *Statistical and Machine Learning for Large Matrices and Tensors*	

Day 1: 18 July 2023 (Tue)	
Time	**Event**
09:30 – 09:50	**Registration**
09:50 – 09:55	**Welcoming Speech** by Prof. Wai Keung LI, Dean and Research Chair Professor of Data Science, EdUHK
10:00 – 11:00	**Keynote Speech** *Some Results on the Blow-up Solutions of the Fast Diffusion Equation* Chairperson: Dr. Chun Kit SUEN Prof. Kin Ming HUI Research Fellow Institute of Mathematics Academia Sinica Taiwan
11:00 – 12:30	**Parallel Session #1** *Recent Advances in Fluid Dynamics and Some Related Models* Chairperson: Dr. Man Wai YUEN • *PDEs with Pseudo-peakons and Non-traveling Wave Peakon Solutions* Prof. Zhijun QIAO • *Ill/well-posedness of Non-diffusive Active Scalar Equations with Physical Applications* Dr. Chun Kit SUEN • *Some Recent Blow-up and Lifespan Estimate Results for Compressible Euler Equations with Damping* Prof. Ningan LAI

Day 1: 18 July 2023 (Tue)	
Time	**Event**
11:00 – 12:30	**Parallel Session #2** *Theory and Application of Nonlinear PDEs* Chairperson: Dr. Tak Kwong WONG • *Well-Posedness of Mildly Dissipative Perturbations of a Family of Active Scalar Equations* Dr. Vincent R MARTINEZ • *Global Existence of Weak Solutions for Compressible Navier–Stokes–Fourier Equations with the Truncated Virial Pressure Law* Dr. Fei WANG • *Multiple Collapsing Blowup Solutions to the 2D Keller-Segel System* Dr. Van Tien NGUYEN
12:30 – 14:30	Lunch Break
14:30 – 16:00	**Parallel Session #3** *Recent Progress on Applied Analysis and its Applications* Chairperson: Dr. Tak Kwong WONG • *Some Recent Progress on Keller-Segel Equation* Dr. Yong YU • *Energy Concentration and Weak Stability in Fluid Dynamics* Dr. Xianpeng HU • *Sobolev Spaces for Wave Equations* Prof. Po Lam YUNG

Day 1: 18 July 2023 (Tue)	
Time	**Event**
14:30 – 16:00	**Parallel Session #4** *Recent Developments on Nonlinear PDEs in Fluids and Waves* Chairperson: Dr. Chun Kit SUEN • *Rigorous Derivation of the Michaelis-Menten Kinetic in the Presence of Diffusion for Enzyme Reactions* Dr. Bao Quoc TANG • *Rigidity for Steady Solutions of Navier-Stokes System in High Dimension* Prof. Chunjing XIE • *Peakons and Weak Solutions for the Modified Camassa-Holm Equation* Dr. Yu GAO

Day 2: 19 July 2023 (Wed)	
Time	**Event**
10:00 – 11:00	**Keynote Speech** *PHM of Complex Engineering Systems – Methods and Applications* Chairperson: Dr. Man Ho LING Prof. Yan-Fu LI Department of Industrial Engineering Director of the Institute for Quality & Reliability Tsinghua University China
11:00 – 12:30	**Parallel Session #1** *Recent Advances in AI for Texts and Images* Chairperson: Prof. Leung Ho YU • *How to Reproduce ChatGPT* Dr. Yaohua TANG • *DetCLIP: Dictionary-Enriched and Scalable Pre-training for Open-Vocabulary Object Detection via Word-Region Alignment* Dr. Hang XU • *Development of a Novel Dementia Risk Prediction Model in the General Population: A Large, Longitudinal, Population-based Machine-Learning Study* Dr. Jia YOU

Day 2: 19 July 2023 (Wed)	
Time	**Event**
11:00 – 12:30	**Parallel Session #2** *Recent Advances in Survival and Reliability Analysis* Chairperson: Dr. Man Ho LING • *Imputations in One-Shot Devices Data using Machine Learning Algorithms* Dr. Hon Yiu SO • *On Support Vector Machine-Based Semi-parametric Cure Model* Dr. Suvra PAL • *Application of a General Family of Bivariate Distributions in Modelling Dependent Competing Risks Data with Associated Model Selection* Dr. Debanjan MITRA
11:00 – 12:30	**Parallel Session #3** *Recent Developments on Data Science and its Applications* Chairperson: Dr. Tse Tin CHAN • *Examining the Spillovers of NFTs and Cryptocurrencies* Dr. Kin Hon HO • *On Candidate Selection with Proportional Fairness Constraints* Prof. Chung Keung POON • *A Study of Google Flu Trends* Mr. Jared XIN
12:30 – 14:30	Lunch Break

Day 2: 19 July 2023 (Wed)	
Time	**Event**
14:30 – 16:00	**Parallel Session #4** *Statistics, Data Science and their Applications* Chairperson: Prof. Wai Keung LI • *Asymptotic Inference of the ARMA Model with Time-Functional Variance Noises* Miss Bibi CAI • *A Hybrid Markov Chain Monte Carlo Algorithm for Structural Learning in Bayesian Networks* Mr. Lupe S.H. CHAN • *Estimation for the Generalized Dynamic Panel Model with Fixed Effects* Mr. Bing SU
14:30 – 16:00	**Parallel Session #5** *Statistical and Machine Learning for Large Matrices and Tensors* Chairperson: Prof. Leung Ho YU • *Robust Factored Principal Component Analysis for Matrix-valued Outlier Accommodation and Detection* Prof. Jianhua ZHAO • *Bayesian Robust Tensor Completion via CP Decomposition* Dr. Anita Xiaohang WANG • *Preference Matrix Completion with Multiple Network Views based on Graph Neural Networks* Mr. Yipeng ZHUANG

Titles and Abstracts

Day 1 (Applied Mathematics): July 18, 2023

Keynote Speech - Some Results on the Blow-up Solutions of the Fast Diffusion Equation

Prof. Kin Ming HUI, Institute of Mathematics, Academia Sinica, Taiwan

Venue: Lady Ivy Wu Lecture Theatre (D1-LP-04), EdUHK

Time: 10:00 am – 11:00 am

In this talk I will discuss some recent results on the blow-up solutions of the fast diffusion equation. I will also discuss my recent result on the existence and large time behaviour of the one point and finite points blow-up solutions of the fast diffusion equation. This is joint work with Sunghoon Kim, Soojung Kim and Jinwan Park.

Parallel Session 1: Recent Advances in Fluid Dynamics and Some Related Models

Prof. Zhijun QIAO, University of Texas Rio Grande Valley, United States

Venue: D2-LP-03, EdUHK

Time: 11:00 am – 11:30 am

Title: *PDEs with Pseudo-peakons and Non-traveling Wave Peakon Solutions*

In this talk, we introduce some higher order models with peakons or pseudo-peakons we proposed recently. We also present new type of peakon solutions of partial differential equations, called rogue peakons provided with a non-traveling wave. Some linear and nonlinear models are taken to illustrate the rogue peakon solutions. Some open problems will be addressed for discussion in the end. Part of work is joint with my student Zhenteng Zeng, Dr. Baoqiang Xia, and Dr. Enrique Reyes.

Dr. Chun Kit SUEN, Department of Mathematics and Information Technology, Education University of Hong Kong, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 11:30 am – 12:00 pm

Title: *Ill/well-posedness of Non-diffusive Active Scalar Equations with Physical Applications*

We consider a general class of non-diffusive active scalar equations with constitutive laws obtained via an operator $\mathbf{T}$ that is singular of order $r_0 \in [0, 2]$. We obtain ill/well-posedness results for various values for r_0. We then apply the results to several physical problems including the magnetogeostrophic equation, the incompressible porous media equation and the singular incompressible porous media equation. This is a joint work with Susan Friedlander and Fei Wang.

Prof. Ningan LAI, College of Mathematics and Science, Zhejiang Normal University, China

Venue: D2-LP-03, EdUHK

Time: 12:00 pm – 12:30 pm

Title: *Some Recent Blow-up and Lifespan Estimate Results for Compressible Euler Equations with Damping*

In this talk, I would like to talk about some recent blow-up results and lifespan estimate for compressible Euler system with damping, the coefficient of which may depend on time or space variable. This talk is based on the joint works with Dr. Jinbo Geng, Nico Michele Schiavone, Manwai Yuen and Jiang Zhou.

Parallel Session 2: Theory and Application of Nonlinear PDEs

Dr. Vincent R MARTINEZ, CUNY Hunter College and Graduate Center, United States

Venue: D2-LP-04, EdUHK

Time: 11:00 am – 11:30 am

Title: *Well-Posedness of Mildly Dissipative Perturbations of a Family of Active Scalar Equations*

We study dissipative perturbations of the 2D generalized surface quasi-geostrophic (gSQG) equations. This family contains the 2D Euler equations in vorticity form at one endpoint, an active scalar equation whose constitutive law relates the velocity with the scalar with a loss of one derivative, and contains the SQG equation at its midpoint. Recent work of Bourgain/Li, Elgindi/Masmoudi, Cordoba/Zoroa-Martinez, and Jeong/Kim have established ill-posedness of this family at critical regularity. This work considers a mild perturbation of the gSQG equation which recovers well-posedness, but instantaneously confers a mild degree of regularity. This work is in contradistinction with strongly dissipative perturbations, which instantaneously confer Gevrey regularity and recover well-posedness at critical regularity (Jolly/Kumar/M 2021), and

inviscid regularization, which do not regularize solutions, but nevertheless recover local well-posedness at critical regularity (Chae/Wu 2010). We show that in this intermediate regime that one may recover local well-posedness at borderline Sobolev regularity, as well as a global existence theory at the 2D Euler endpoint. Moreover, we provide a general existence theory for an entire class of such perturbations that is effectively sharp in light of the recent ill-posedness results. This is joint work with A. Kumar (Florida State University).

Dr. Fei WANG, School of Mathematical Science, Shanghai Jiao Tong University, China

Venue: D2-LP-04, EdUHK

Time: 11:30 am – 12:00 pm

Title: *Global Existence of Weak Solutions for Compresssible Navier–Stokes–Fourier Equations with the Truncated Virial Pressure Law*

This paper concerns the existence of global weak solutions *à la Leray* for compressible Navier–Stokes–Fourier system with periodic boundary conditions and the truncated virial pressure law which is assumed to be thermodynamically unstable. More precisely, the main novelty is that the pressure law is not assumed to be monotone with respect to the density. This provides the first global weak solutions result for the compressible Navier-Stokes-Fourier system with such kind of pressure law which is strongly used as a generalization of the perfect gas law. The paper is based on a new construction of approximate solutions through an iterative scheme and fixed point procedure which could be very helpful to design efficient numerical schemes. Note that our method involves the recent paper by the authors published in Nonlinearity (2021) for the compactness of the density when the temperature is given.

Dr. Van Tien NGUYEN, Department of Mathematics, National Taiwan University, Taiwan

Venue: D2-LP-04, EdUHK

Time: 12:00 pm – 12:30 pm

Title: *Multiple Collapsing Blowup Solutions to the 2D Keller-Segel System*

The talk presents a constructive approach for the existence of blowup solutions to the two dimensional Keller-Segel system. The constructed solution is formed by a strong interaction/collision of two single solutions with self-similarity and their radiation, which provides *a brand new mechanism* of singularity formation. The analysis relies on spectral approach for multiple-scale problems, renormalization technique and refined energy estimates. The talk is based on a joint project with C. Collot (Paris Cergy), T. Ghoul (NYU Abu Dhabi) and N. Masmoudi (NYU).

Parallel Session 3: Recent Progress on Applied Analysis and its Applications

Dr. Yong YU, Department of Mathematics, Chinese University of Hong Kong, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 2:30 pm – 3:00 pm

Title: *Some Recent Progress on Keller-Segel Equation*

Keller-Segel equation is a dynamical equation arising in the theories of chemotaxis and astrophysics. There are rich phenomena associated with solutions to the Keller-Segel equation. In this talk, I will discuss the long time asymptotics of its global solutions. I will also classify its self-similar blow-up profile under the radial symmetry assumption. The results to be presented will be in any dimensions.

Dr. Xianpeng HU, Department of Mathematics, City University of Hong Kong, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 3:00 pm – 3:30 pm

Title: *Energy Concentration and Weak Stability in Fluid Dynamics*

The weak stability is an important issue in fluid dynamics. We will discuss the mathematical understudying of concentration phenomena in the framework of weak solutions with either critical or subcritical energy. Two typical examples, including two dimensional incompressible Euler equations and compressible Navier-Stokes equations, will be discussed.

Prof. Po Lam YUNG, Mathematical Sciences Institute, Australian National University, Australia

Venue: D2-LP-03, EdUHK

Time: 3:30 pm – 4:00 pm

Title: *Sobolev Spaces for Wave Equations*

We will describe some new function spaces, that can be used to understand regularity of solutions of wave equations in L^p. Built into the construction of the spaces are certain natural wave packet decompositions going back to Seeger, Sogge and Stein, and applications of the spaces include

i) a refinement of some classical Sobolev embedding theorems on Euclidean spaces;

ii) a description of Fourier decoupling inequalities for the unit sphere.

Joint work with Andrew Hassell, Pierre Portal and Jan Rozendaal.

Parallel Session 4: Recent Developments on Nonlinear PDEs in Fluids and Waves

Dr. Bao Quoc TANG, Institute of Mathematics and Scientific Computing, University of Graz, Austria

Venue: D2-LP-04, EdUHK

Time: 2:30 pm – 3:00 pm

Title: *Rigorous Derivation of the Michaelis-Menten Kinetic in the Presence of Diffusion for Enzyme Reactions*

Michaelis-Menten kinetic is one of the most used when modelling enzyme-, or more generally catalytic-, reactions. In the case of homogeneous medium, i.e. the (bio-)chemical concentrations depend solely on time, both formal and rigorous derivations of MM from mass action kinetic have been studied extensively and thoroughly in the last decades. For heterogeneous medium, the modelling should take into account the diffusion of substances, which leads to a system of partial differential equations. In this case, interestingly, only formal derivation of MM from mass action kinetic has been investigated. In this talk, we present, up to our knowledge, the first rigorous derivation of MM in the presence of diffusion. The proof utilises an improved duality technique and a modified energy method. This is based on a joint work with Bao-Ngoc Tran (University of Graz).

Prof. Chunjing XIE, School of Mathematical Science, Shanghai Jiao Tong University, China

Venue: D2-LP-04, EdUHK

Time: 3:00 pm – 3:30 pm

Title: *Rigidity for Steady Solutions of Navier-Stokes System in High Dimension*

The stationary Navier-Stokes equations enjoy a special scaling property thanks to its nonlinear character. Several scaling-invariant classes motivated by the scaling property have proved useful in investigating various properties of a solution. In this talk, we will discuss the rigidity of steady Navier-Stokes system in a more general class, where we do not need any type of self-similarity or smallness of solutions. The proof of the results is based on energy method and is pretty elementary.

Dr. Yu GAO, Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong

Venue: D2-LP-04, EdUHK

Time: 3:30 pm – 4:00 pm

Title: *Peakons and Weak Solutions for the Modified Camassa-Holm Equation*

The modified Camassa-Holm equation is an integrable equation with cubic nonlinearity. It has special solitary wave solutions as weak solutions, which are known as peakons. In this talk, we are going to introduce sticky particle method and particle blob method for N-peakon solutions to the modified Camassa-Holm equation.

Titles and Abstracts

Day 2 (Data Science): July 19, 2023

Keynote Speech - PHM of Complex Engineering Systems – Methods and Applications

Prof. Yan-Fu LI, Department of Industrial Engineering, Tsinghua University, China

Venue: Lady Ivy Wu Lecture Theatre (D1-LP-04), EdUHK

Time: 10:00 am – 11:00 am

Prognostics and Health Management (PHM) is an important research direction in reliability engineering. With the large-scale deployment of sensors and the maturity of big data technology, PHM has been more and more widely used in complex engineering systems, and it has produced significant social and economic benefits. PHM generally includes the key tasks such as anomaly detection, fault diagnosis, health assessment, and remaining useful life (RUL) prediction. The new generation of machine learning methods represented by deep learning has played a central role in promoting the development of PHM in the big data environment. This report takes high-speed train and aircraft key components and subsystems as examples, and presents the latest developments in our laboratory. These new methods have a certain degree of versatility, and can be extended to the key components of other types of engineering systems.

Parallel Session 1: Recent Advances in AI for Texts and Images

Dr. Yaohua TANG, Moore Threads, China

Venue: D2-LP-03,EdUHK

Time: 11:00 am – 11:30 am

Title: *How to Reproduce ChatGPT*

The emergence of ChatGPT has aroused widespread attention. Its intelligence level is so amazing that some people call it the Fourth Industrial Revolution. In this talk, we provide a detailed guide on how to reproduce the ChatGPT model, starting from data collecting to model training and evaluation. We first discuss the data sources and preprocessing steps, including codes, instructions and human feedback data for reinforcement learning. Next, we describe the four key steps of model training. Finally, we present the evaluation tools and experimental results,

showing the pipeline of the evaluation of ChatGPT-like models. Our aim is to provide a comprehensive and reproducible guide that can help researchers and practitioners to build their own competitive ChatGPT-like models using the same methodology as ChatGPT.

Dr. Hang XU, Computer Vision Team, Huawei Noah Ark Lab, China

Venue: D2-LP-03, EdUHK

Time: 11:30 am – 12:00 pm

Title: *DetCLIP: Dictionary-Enriched and Scalable Pre-training for Open-Vocabulary Object Detection via Word-Region Alignment*

In this talk, we will introduce DetCLIP a pre-training method for open-world object detection and open-vocabulary object detection (OVD). DetCLIP is a dictionary-enriched visual-concept pre-training method that utilizes a designed concept dictionary to provide prior knowledge for each concept. DetCLIP achieves better learning efficiency by extracting concepts separately to better utilize heterogeneous datasets and provides sufficient negative concepts for the construction of the word-region alignment loss. DetCLIP demonstrates strong zero-shot detection performances, outperforming previous works on the LVIS dataset. DetCLIPv2, on the other hand, is an efficient and scalable training framework that incorporates large-scale image-text pairs to achieve OVD. DetCLIPv2 directly learns the fine-grained word-region alignment from massive image-text pairs in an end-to-end manner by employing a maximum word-region similarity between region proposals and textual words to guide the contrastive objective. DetCLIPv2 is trained with a hybrid supervision from detection, grounding, and image-text pair data under a unified data formulation. By jointly training with an alternating scheme and adopting low-resolution input for image-text pairs, DetCLIPv2 exploits image-text pair data efficiently and effectively. DetCLIPv2 demonstrates superior open-vocabulary detection performance, outperforming previous works on the LVIS benchmark, including DetCLIP, and even beats its fully-supervised counterpart by a large margin.

Dr. Jia YOU, Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, China

Venue: D2-LP-03, EdUHK

Time: 12:00 pm – 12:30 pm

Title: *Development of a Novel Dementia Risk Prediction Model in the General Population: A Large, Longitudinal, Population-based Machine-Learning Study*

Background The existing dementia risk models are limited to empirical risk factors; thus we aimed to employ machine learning (ML) in large health data to develop a novel dementia prediction model.

Methods In this longitudinal population-based cohort of UK Biobank, 425,159 non-demented participants were enrolled from 22 recruitment centers around the UK from 2006 to 2010. A data-driven strategy was implemented to identify predictors from 366 candidate variables covering a comprehensive range of genetic and environmental factors and developed the ML model to predict incident dementia and Alzheimer's Disease (AD).

Findings During a median follow-up of 11.9 years, 5,287 and 2,416 participants developed dementia and AD, respectively. A novel UKB dementia risk prediction (UKB-DRP) model comprising ten predictors including age, ApoE 4, pairs matching time, leg fat percentage, number of medications taken, reaction time, peak expiratory flow, mother's age at death, long-standing illness, and mean corpuscular volume were established. The proposed model was internally evaluated based on five-fold cross-validation on discrimination and calibration, and it was further compared with existing prediction scales. The UKB-DRP model achieved high discriminative accuracy in dementia (AUC 0.848±0.007) and even better in AD (AUC 0.862±0.015). The model was well-calibrated (Hosmer-Lemeshow goodness-of-fit p-value = 0.92), and the predictive power was solid in different incidence time groups. More importantly, our model presented an apparent superiority over existing models like Cardiovascular Risk Factors, Aging, and Incidence of Dementia Risk Score (AUC 0.705±0.008), the Dementia Risk Score (AUC 0.752±0.007), and the Australian National University Alzheimer's Disease Risk Index (AUC 0.584±0.017). The model was internally validated in the general population of European ancestry and White ethnicity; thus, further validation with independent datasets is necessary to confirm these findings.

Interpretation Our ML-based UKB-DRP model incorporated ten easily accessible predictors with solid predictive power for incident dementia and AD within five, ten, and much longer years, which can be used to identify individuals at high risk of dementia and AD in the general population.

Parallel Session 2: Recent Advances in Survival and Reliability Analysis

Dr. Hon Yiu SO, Department of Mathematics and Statistic, Oakland University, USA

Venue: D2-LP-04, EdUHK

Time: 11:00 am – 11:30 am

Title: *Imputations in One-Shot Devices Data using Machine Learning Algorithms*

One-shot devices are products that will be destroyed immediately after use. Most of them have multiple components. Malfunctioning any one of the components will result in the device's failure. We often test the one-shot devices under constant stress accelerated life-test or collect data from users or surveys to assess such devices. A link function relating to stress levels and lifetime is then applied to extrapolate the lifetimes of units from accelerated conditions to normal operating conditions. However, we often have missing data during the data collection, and imputation is a popular way to analyze this data. This study will explore imputation performance using machine learning algorithms on one-shot datasets and compare them to traditional imputation methods.

Dr. Suvra PAL, Department of Mathematics, University of Texas at Arlington, USA

Venue: D2-LP-04, EdUHK

Time: 11:30 am – 12:00 pm

Title: *On Support Vector Machine-Based Semi-parametric Cure Model*

In this talk, I will present a new semi-parametric promotion time cure model (PCM) that uses the support vector machine (SVM) to model the incidence part. The proposed model inherits the features of the SVM and provides flexibility in capturing non-linearity in the data. For the estimation of model parameters, I will discuss the steps of an expectation maximization algorithm where I will make use of the sequential minimal optimization technique together with the Platt scaling method. Next, I will present the results of a detailed simulation study and show that the proposed model outperforms the existing PCM models, specifically when the true classification boundary is non-linear. I will also show that the proposed model's ability to capture complex classification boundaries can improve the estimation results related to the latency part. Finally, I will analyze a data from leukemia cancer study and show that the proposed model results in improved predictive accuracy.

Dr. Debanjan MITRA, Indian Institute of Management Udaipur, India

Venue: D2-LP-04, EdUHK

Time: 12:00 pm – 12:30 pm

Title: *Application of a General Family of Bivariate Distributions in Modelling Dependent Competing Risks Data with Associated Model Selection*

In lifetime data analysis, the scenario of competing risks arises when there are multiple risk factors that can cause occurrence of the event of interest for subjects under study, and the occurrence of the event from any one of the risk factors prevents the occurrence of the event from any other relevant risk factors. The risk factors, quite often, influence each other. Competing risks models are widely studied in various domains such as medical sciences, engineering, and finance. In this talk, an approach to modelling competing risks data with dependent risk factors will be presented. The general structure of competing risks data considered here can include ties. In this approach, a general family of bivariate distributions that can capture the dependence between the risk factors in presence of ties is used for the modelling purpose. The proposed model, and a comprehensive inferential framework for the proposed model including maximum likelihood estimation, and construction of confidence intervals will be presented in this talk. Selection of an appropriate model for a given data is an important aspect of parametric statistical analysis. In this presentation, model selection within the bivariate family of distributions for a given dependent competing risks data will also be discussed. A detailed Monte Carlo simulation study is conducted to examine the methods of inference for the proposed model, as well as to assess the accuracy of model selection method within the bivariate family. From the results of the simulation study, we will show that the inferential methods

provide quite reasonable results. Finally, analysis of a real data from the Diabetic Retinopathy Study will be presented as an illustrative example.

Parallel Session 3: Recent Developments on Data Science and its Applications

Dr. Kin Hon HO, Department of Business Administration, Hong Kong Shue Yan University, Hong Kong

Venue: D2-LP-08, EdUHK

Time: 11:00 am – 11:30 am

Title: *Examining the Spillovers of NFTs and Cryptocurrencies*

This work investigates spillovers between non-fungible token (NFT) assets. More specially, we conduct a spillover analyses for three scenarios where shocks between the assets may generally assumed to exist: (1) Crypto-NFT: NFTs and their official cryptocurrencies; (2) intra-NFT: different assets belonging to the same NFT, and (3) inter-NFT: assets of the same type belonging to different NFTs. Our preliminary results find that each NFT asset is distinct from the others, so spillover effects are limited in all the scenarios. This finding suggests that investors can realize the benefits of portfolio diversification from NFT assets.

Prof. Chung Keung POON, Department of Computing, Hang Seng University of Hong Kong, Hong Kong

Venue: D2-LP-08, EdUHK

Time: 11:30 am – 12:00 pm

Title: *On Candidate Selection with Proportional Fairness Constraints*

Selecting a subset of candidates with various attributes under fairness constraints has been attracting considerable attention from the AI community, with applications ranging from school admissions to committee selections. In this talk, I will introduce the proportional candidate selection problem, in which the goal is to select a subset of candidates with maximum cardinality while meeting certain proportional fairness constraints. I will then show the intrinsic hardness of the problem, and describe some algorithms that solve certain easier versions of the problem.

Mr. Jared XIN, Sha Tin College, Hong Kong

Venue: D2-LP-08, EdUHK

Time: 12:00 pm – 12:30 pm

Title: *A Study of Google Flu Trends*

In this presentation, I will briefly discuss machine learning and specifically discuss a well known example of big data's failures - google flu trends. I will analyse the reasons behind its failure and inaccuracy, look at the factors for the aforementioned problems and propose solutions to the project to make it more successful.

Parallel Session 4: Statistics, Data Science and their Applications

Miss Bibi CAI, The Hong Kong University of Science and Technology, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 2:30 pm – 3:00 pm

Title: *Asymptotic Inference of the ARMA Model with Time-Functional Variance Noises*

This talk will consider the autoregressive and moving average (ARMA) model with time- functional variance (TFV) noises, called the ARMA-TFV model. We first establish the consistency and asymptotic normality of its least squares estimator (LSE). Based on the theory, the Wald test and a portmanteau test are constructed for variable selection and model checking, respectively. A simulation study is carried to assess the performance of our approach in finite samples and two real examples are given. It should be mentioned that the process from the ARMA-TFV model is not stationary and the technique in this paper is non-standard and may provide some insights for future research in this direction. This talk is based on a joint work with Shiqing Ling and Enwen Zhu.

Mr. Lupe S.H. CHAN, The Hong Kong University of Science and Technology, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 3:00 pm – 3:30 pm

Title: *A Hybrid Markov Chain Monte Carlo Algorithm for Structural Learning in Bayesian Networks*

Bayesian networks are models to represent dependence structures among variables through a directed acyclic graph (DAG). A challenge in structural learning is that the number of possible DAG grows super-exponentially as the number of variables increases. Most existing works discover structures over either the DAG space or the topological order space. We propose an algorithm that uses Markov chain Monte Carlo (MCMC) to learn Bayesian networks from data, making use of both the DAG space and the topological order space. We first partition the variables of similar topological orders into blocks. The structure search is then conducted over the DAG space within each block to promote more targeted edge moves. Across-block search is

also done to ensure the ergodicity of the Markov chain. Both simulation and empirical results suggest that the proposed algorithm enhances efficiency in structural learning.

Mr. Bing SU, Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong

Venue: D2-LP-03, EdUHK

Time: 3:30 pm – 4:00 pm

Title: *Estimation for the Generalized Dynamic Panel Model with Fixed Effects*

The serial correlation and heteroskedasticity are two important features for the disturbance of the panel model. Many inferential methods have taken these two features into account in the random effect panel models, however, due to the incidental parameter problem, only few attempts have been made toward this goal in the fixed effect panel models. In this paper, we propose a generalized dynamic panel (GDP) model with fixed effects. The GDP model not only accounts for the high-order serial correlation, but also characterizes the heteroskedasticity and parameter heterogeneity. For the GDP model including individual or time effects, we provide the quasi maximum likelihood estimator with its bias correction version. The relevant asymptotic properties are studied by applying some novel asymptotic tools for linear-quadratic vector-form variables. The usefulness of the GDP model with its estimation is illustrated by simulation results and an empirical example on stock returns.

Parallel Session 5: Statistical and Machine Learning for Large Matrices and Tensors

Prof. Jianhua ZHAO, School of Statistics and Mathematics, Yunnan University of Finance and Economics, China

Venue: D2-LP-04, EdUHK

Time: 2:30 pm – 3:00 pm

Title: *Robust Factored Principal Component Analysis for Matrix-valued Outlier Accommodation and Detection*

Principal component analysis (PCA) is a popular dimension reduction technique for vector data. Factored PCA (FPCA) is a probabilistic extension of PCA for matrix data, which can substantially reduce the number of parameters in PCA while yield satisfactory performance. However, FPCA is based on the Gaussian assumption and thereby susceptible to outliers. Although the multivariate t distribution as a robust modeling tool for vector data has a very long history, its application to matrix data is very limited. The main reason is that the dimension of the

vectorized matrix data is often very high, and the higher the dimension, the lower the breakdown point that measures the robustness. To solve the robustness problem suffered by FPCA and make it applicable to matrix data, in this paper a robust extension of FPCA (RFPCA) is proposed, which is built upon a t-type distribution called matrix-variate t distribution. Like the multivariate t distribution, the matrix-variate t distribution can adaptively down-weight outliers and yield robust estimates. A fast EM-type algorithm for parameter estimation is developed. Experiments on synthetic and real-world datasets reveal that (i) RFPCA is compared favorably with several closely related methods. Importantly, RFPCA has a significantly higher breakdown point than its vector-based cousin multivariate t PCA (tPCA), which makes RFPCA more applicable to matrix data; (ii) the expected latent weights of RFPCA can be readily used for outlier detection, and they are much more reliable than those by tPCA. Such detection is rarely available with existing matrix-based methods, especially for gross matrix-valued outliers. Joint work with Xuan Ma, Yue Wang, Changchun Shang and Fen Jiang.

Dr. Anita Xiaohang WANG, College of Big Data and Internet, Shenzhen Technology University, China

Venue: D2-LP-04, EdUHK

Time: 3:00 pm – 3:30 pm

Title: *Bayesian Robust Tensor Completion via CP Decomposition*

The real-world tensor data are inevitably missing and corrupted with noise. Some models of the low rank tensor factorization (LRTF) add an L1 norm or L2 norm to deal with the sparse or Gaussian noise. However, the real noises are usually complex. We propose a robust Bayesian tensor completion method, called MoG BTC-CP, which could impute the missing data and remove the complex noise simultaneously. The observed tensor is assumed to be the summation of a low-rank tensor and the noise. CP decomposition is proposed to extract the low-rank structure of the tensor. We assume that the noise follows a Mixture of Gaussian (MoG) distribution. A full Bayesian framework together with a Gibbs sampling algorithm is designed to estimate the model. Extensive experiments including synthetic data and real-life applications show that MoG BTC-CP outperforms the recently published leading tensor completion and denoising methods.

Mr. Yipeng ZHUANG, Department of Mathematics and Information Technology, Education University of Hong Kong, Hong Kong

Venue: D2-LP-04, EdUHK

Time: 3:30 pm – 4:00 pm

Title: *Preference Matrix Completion with Multiple Network Views based on Graph Neural Networks*

In our digital age, people are connected to many networks. Their preferences (such as ratings and rankings) for all the items (such as movies and products) would not complete and be affected by some of these networks. To predict the missing preferences, we will propose a novel graph neural network model for preference matrix completion with multiple network views and side information. Attention mechanism was applied to measure the influences from multi-view networks. New objective functions are proposed to evaluate the preference prediction performance. Finally, we will apply our proposed model to some real-world movie recommendation datasets. Empirical results demonstrate that our model achieves significant improvements in preference prediction over other existing models. Joint work with Philip L.H. Yu.