Content

Introduction to SPNC-2019……………………………………2
Travel Information……………………………………………3
Hotel Information……………………………………………6
Registration and Conference Venue…………………………7
Campus Maps…………………………………………………..8
Conference Agenda…………………………………………..10
Conference Program…………………………………………13
Abstracts of All Talks………………………………………20
Participant List……………………………………………….. 41
Acknowledgement……………………………………………46
Notes Page……………………………………………………47

Introduction

There are active research interests in the interdisciplinary boundaries across physics, mathematics, machine learning and neural computation, which is mostly stimulated by the success of deep learning techniques in various academic and industrial applications, and also by the inquiry of how the brain works during worldwide brain projects. However, the core topic—-neural networks, still being a black box lacking mechanism explanations, drives the recent efforts of the physics/math community towards understanding the inner workings of these artificial and biological neural networks. Progress along this line should provide keys to open new doors towards novel applications, predictions and even new paradigms. Here, we gather top experts in this field to discuss recent progress, and inspire collaborations.

Scope

- Analytic and numerical techniques to analyze neural networks
- Theory of deep learning
- Dynamical models of neural computation
- Biological neural networks
- Other theoretical topics of neural networks

Organizer

Haiping Huang, School of physics, Sun Yat-sen University

Mobile: +86-188-1839-6276

Email: huanghp7@mail.sysu.edu.cn

WeChat: physhuang

Conference website

https://www.labxing.com/lab/666/news/365
https://huanghp7.wixsite.com/spnn-2019

Travel Information

VISA

As we know, no visa is required for ordinary passport holders from Singapore, Brunei and Japan to visit China for up to 15 days for business, sightseeing, visiting relatives and friends or transit. Other foreign visitors need a visa to enter China. Please check with the Chinese embassy or consulate in your residence area to make sure whether you need a visa.

It is better that you apply for a visa one month before your planned date of entry into China but not earlier than three months.

Most non-Chinese citizens will need a visa to come to China for the conference. The types of visas suitable for the participation of the conference are either "L" or "F". Here is a reference website explaining different types of visas, http://www.china-embassy.org/eng/visas/hrsq/. To apply for a Chinese visa, please check the website above or the homepage of your nearest Chinese consulate for required documents. You will need an invitation letter from us.

To get a proper invitation letter, please fill in the word file or contact huanghp7@mail.sysu.edu.cn. We will then prepare an invitation letter for you to support your visa application.

Transportation to Guangzhou

1. Baiyun airport (CAN), Guangzhou (Recommended)
    - http://www.gbiac.net/en/byhome, ~45km away from the city center .
    - Take a taxi to SYSU (about RMB150). It takes about one hour but not in busy time.
    - Take the Airport Express Bus and continue with a taxi: Take the Airport Express Bus (about RMB30) to the New Pearl River Hotel (Metro Line 6) or to Guangzhou Textile Exchange Park (Metro Line 10), get off and take a taxi (RMB10) to the South Campus of SYSU.
    - Take the metro at Airport South Station to SYSU (RMB ~12): First take Line 3 (Airport S. to Tiyuxi Road) to the Tiyuxi Road Station then trans-

fer to another Line 3 (Tianhe Coach Terminal to Panyu Square) and get to the Kecun Station, then transfer to Line 8 (towards Fenghuang Xincun). Get off at Lujiang Station (for SYSU East Gate) or at Sun Yat-Sen University Station (for SYSU South Gate).

2. Hong Kong airport (HKG), Hong Kong
   - Buses: There are a few limo/bus companies from the airport to Guangzhou. Follow the sign "Transportation to Mainland"and there are several coach companies providing such services. Tickets cost about HKD230-250. You will be taken to your buses by the staffs at the airport but, unfortunately, there is no English information from the website of the bus company and English service after you board the bus is not guaranteed any more. These buses stop at a few major connection points in Guangzhou and they vary from company to company. Not suggested to people who don't speak Chinese but quite convenient for people who do: detailed information is HERE. Most buses would stop at Tianhe (Jinnan University), which might be the most convenient to SYSU.
   - Metro + Trains (transfer at the border): Use the metro system in Hong Kong to Luohu Station first. Then cross the border to Mainland and you are in Shenzhen. You can use the bullet trains connecting Luohu Station Shenzhen and Guangzhou East. It takes ~80-90 minutes and runs every 15-20 minutes. Walk to the Guang-Shen ticket counter (follow the signs) and buy a Guangzhou-Shenzhen express ticket with your passports.
   - Ferry service to Guangzhou Lian Hua Shan from Hong Kong Airport, but the time table is very limited. Please see the link. From Lian Hua Shan port, you can take a taxi/metro system to the University (see below). For the ferry, see also here.
3. Land connections to Guangzhou (within China):
   - Bullet train: Guangzhou South Station
   - China Railway High-speed (CRH): Guangzhou South Station
4. Within Guangzhou:

- Taxi service is very convenient in Guangzhou. An available Taxi displays a red sign of "空车(Empty)" near the front windshield. Please keep the printed invoice when you get off so that you can contact the taxidriver or the taxi company in case of any problems.
- Our address is: "广州市海珠区新港西路135号中山大学" in Chinese and the nearest metro stations are Lujiang or Sun Yat-Sen University Station on Line 8.

Note that for invited speakers, please buy the flight ticket yourself, and then get the reimbursement. Please keep the receipt, E-ticket itinerary and the flight boarding cards for us. When you attend the conference, please give us your passport for a copy, together with bank account information (including swift code for foreigners).

Hotel Information

The hotel rooms for invited and contributed speakers are booked by the organizer. The location is within the campus, i.e., Sun Yat-sen University Kaifeng Hotel near to the North Gate.

Hotel Venue Directions for you and your Drivers:

“请送我到中大凯丰酒店（广州市海珠区滨江东路中山大学南校区北门）。”

“Please send me to Kaifeng Hotel (North Gate of Sun Yat-sen University, Binjiang Dong Road, Haizhu District, Guangzhou, 510275, China).”

For other participants, the information about two other hotels on campus is given as follows:

1. Wing Kwong Hall Hotel (中大荣光堂）

广州市海珠区新港西路135号中大南校园 （靠北门）Near to the North gate, Tel：020-84112138

2. Zi-jing-yuan Hotel (中山大学紫荆园宾馆)

中大南校区内，靠南门，Near to the South gate, Tel：020-84111888

For tour information, please see the following links:

- http://www.chinatourguide.com/guangzhou/
- https://www.chinahighlights.com/guangzhou/
- https://www.lonelyplanet.com/china/guangdong/guangzhou

Registration and Conference Venue

Time: 14:00-21:00 Oct 3, 2019

Venue: Sun Yat-sen University Kaifeng Hotel near to the North Gate

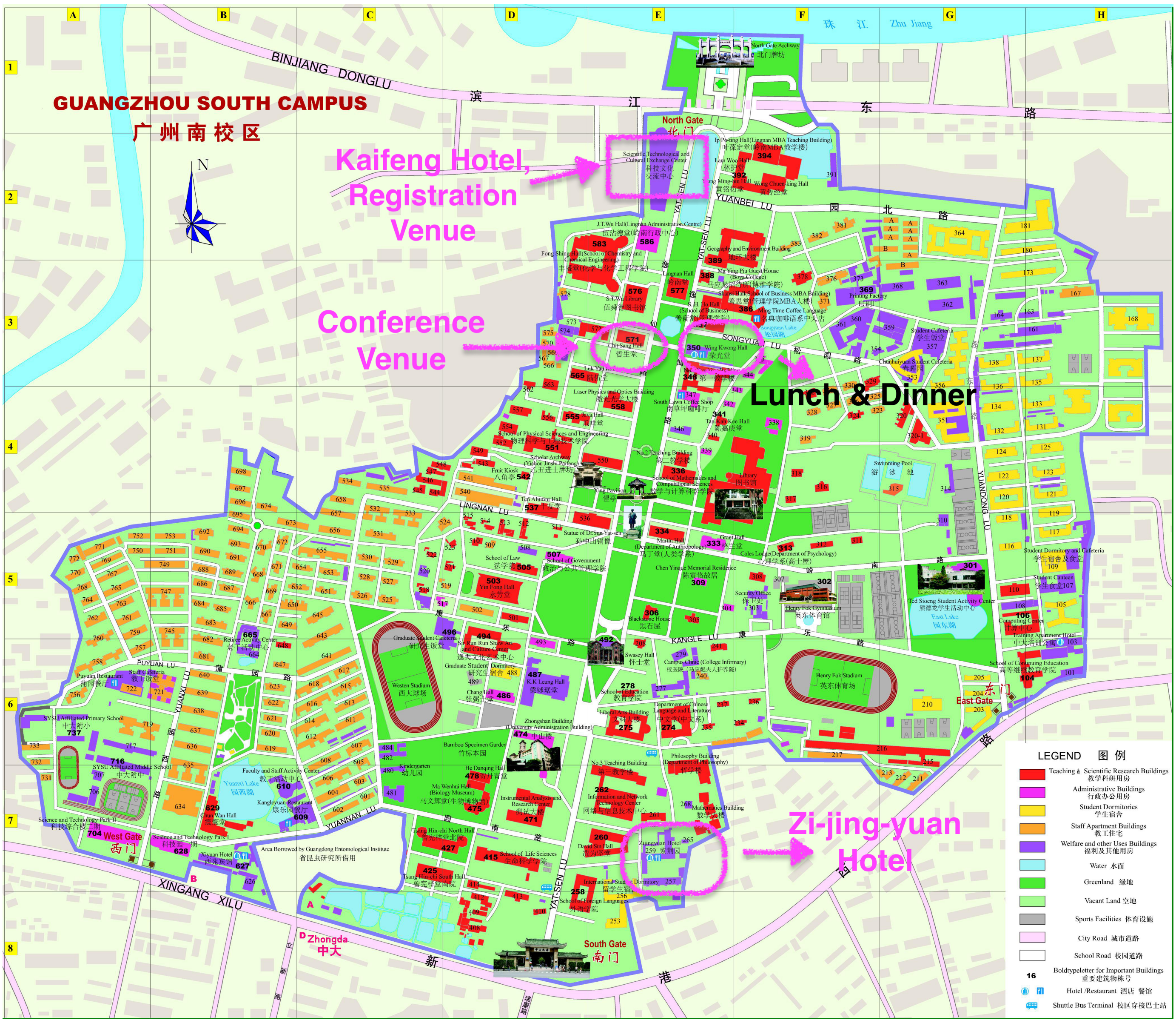
For those missing the chance of registration on Oct 3, please do registration in the conference venue during 8:00-8:30 on Oct 4.

Conference Venue:

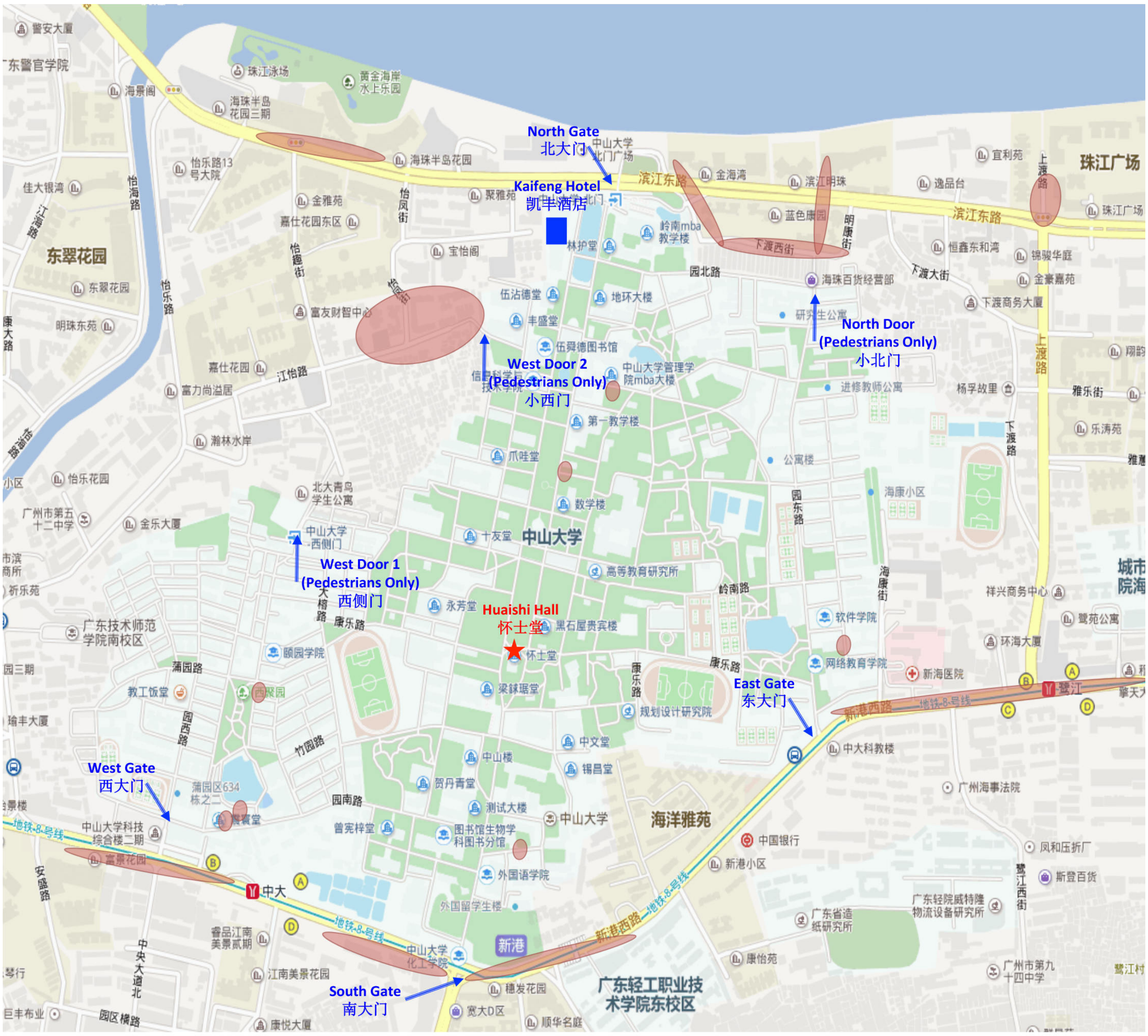
Chit Sang Hall, 哲生堂/Lecture Hall, 3rd floor

For map details, see the campus maps in the following pages.

Campus Maps


GUANGZHOU SOUTH CAMPUS
广州南校区
N
Kaifeng Hotel,
Registration
Venue
Conference
Venue
Lunch & Dinner
Zi-jing-yuan
Hotel
BINJIANG DONGLU
North Gate
北门
YUANBEI LU
SONGYUAN LU
LINGNAN LU
KANGLE LU
PUYUAN LU
YUANXI LU
YUANNAN LU
XINGANG XILU
YAT-SEN LU
YUANDONG LU
West Gate
西门
East Gate
东门
South Gate
南门
Zhongda
中大
Zhu Jiang
LEGEND 图例
Teaching & Scientific Research Buildings 教学科研用房
Administrative Buildings 行政办公用房
Student Dormitories 学生宿舍
Staff Apartment Buildings 教工住宅
Welfare and other Uses Buildings 福利及其他用房
Water 水面
Greenland 绿地
Vacant Land 空地
Sports Facilities 体育设施
City Road 城市道路
School Road 校园道路
Boldtypeletter for Important Buildings 重要建筑物栋号
Hotel /Restaurant 酒店 餐馆
Shuttle Bus Terminal 校区穿梭巴士站

Food places around SYSU



Food Spots&Zones (Open to Public)

Conference Agenda

Date	Time	Task	Venue
Oct 3	14:00-21:00	Registration	Kaifeng Hotel
Oct 4	8:00-8:30	Registration	Chit Sang Hall/哲生堂3rd floor
	8:30-12:15	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor
	12:15-14:00	Lunch	Wing Kwong Hall Hotel/荣光堂酒店, only for invited and contributed speakers
	14:00-17:40	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor
	19:00-21:00	Banquet	Bing-Sheng Pin-Wei (Haiyin Headquarter), only for invited and contributed speakers
Oct 5	8:30-12:00	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor

Date	Time	Task	Venue
	12:00-14:00	Lunch	Wing Kwong Hall Hotel/荣光堂酒店, only for invited and contributed speakers
	14:00-17:40	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor
	18:30-20:00	Dinner	Wing Kwong Hall Hotel/荣光堂酒店, only for invited and contributed speakers
	20:00-21:30	Optional Plan（cost about RMB 100)	Pearl River Cruise from Zhong-Da port. Gather at the lobby of Kaifeng Hotel on 20:00
Oct 6	8:30-12:20	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor
	12:20-14:00	Lunch	Wing Kwong Hall Hotel/荣光堂酒店, only for invited and contributed speakers

Date	Time	Task	Venue
	14:00-17:50	Talks	Chit Sang Hall/哲生堂3rd floor, Tea time on the 4th floor
	18:30-20:00	Dinner	Wing Kwong Hall Hotel/荣光堂酒店, only for invited and contributed speakers

Conference Program

This workshop follows the "**DEIF**" principle. Please do not highlight the journal's impact factor in your talk and slides. We love stories more than others. Thank you for cooperation!

Note: the number behind the speaker indicates the abstract order in the following section.

Oct 4, 2019

Section A Chair : **Haiping Huang**

08:30-08:45	Daoxin Yao (Vice Dean of physics school)	Sun Yat-sen University	Introduction to School of Physics
08:45-09:35	David Saad 1	Aston University, UK	Function Space Entropy in Deep-Learning Networks
09:35-10:25	Adriano Barra 2	Università del Salento, Italy	Guerra's interpolation techniques in Neural Networks and Machine Learning
	Break/Photo	10:25-10:55	

Section B Chair: **Jonathan Kadmon**

10:55-11:45	Alexis Dubreuil 3	ENS, France	Disentangling the roles of dimensionality and cell classes in neural Computations

11:45-12:15	Lukasz Kusmierz 4	RIKEN, Japan	Dynamical mean field theory of neural networks with quenched power-law disorder
	Lunch	12:15-14:00	

Section C Chair: **Si Wu**

14:00-14:50	Pulin Gong 5	Sydney University	An integrative model interrelating dynamical and computational properties of cortical circuits
14:50-15:40	Taro Toyoizumi 6	RIKEN, Japan	Exploring the learning principle in the brain
	Break	15:40-16:10	

Section D Chair: **David Saad**

16:10-16:40	Chihiro Noguchi 7	Tokyo Tech, Japan	Approximate matrix completion based on cavity method
16:40-17:10	Yingying Xu 8	Tokyo Tech, Japan	Capacity of Dimension Selection with Simple Perceptron

17:10-17:40	Zedong Bi 9	HKBU	Understanding the computational difficulty of binary-weight perceptron and the advantage of input sparseness
	Banquet	19:00-21:00	Bing-Sheng Pin-Wei (Haiyin Headquarter)

Oct 5, 2019

Section A	Chair: **Changsong Zhou**		
08:30-09:20	Si Wu 10	Peking University	Push-pull Feedback Implements Rough-to-fine Information Processing
09:20-10:10	KY Michael Wong 11	HKUST	Integrating and segregating information from different channels in the Brain
	Break	10:10-10:40	
Section B Chair: **Alexis Dubreuil**			
10:40-11:30	Jonathan Kadmon 12	Stanford University	Deep networks with random representations

11:30-12:00	Jun-hao Liang 13	HKBU	A macroscopic description of excitation-inhibition balanced network with synaptic kinetics
	Lunch	12:00-14:00	
Section C Chair: **Adriano Barra**			
14:00-14:50	Yoshiyuki Kabashima 14	Tokyo Tech	On Replica-BP correspondence in rotationally invariant SG models
14:50-15:40	Pan Zhang 15	ITP, CAS	Solving Statistical Mechanics using Variational Autoregressive Networks
	Break	15:40-16:10	
Section D Chair: **K Y Michael Wong**			
16:10-16:40	Yu Terada 16	RIKEN, Japan	Objective and efficient inference for couplings in neuronal networks
16:40-17:10	Tomoyuki Obuchi 17	Tokyo Tech	Cross validation in sparse linear regression with piecewise continuous nonconvex penalties and its acceleration

17:10-17:40	Jun-tao Wang 18	HKUST	Generalized Belief Propagation in Gaussian Markov Random Fields
	Dinner	18:30-20:00	

Oct 6, 2019

Section A	Chair: **Taro Toyoizumi**		
08:30-09:20	Carlo Baldassi 19	Bocconi University	A large deviations statistical physics approach to finding wide flat minima in neural networks landscapes
09:20-10:10	Masafumi Oizumi 20	Tokyo University	Towards Identifying Places and Boundaries of Consciousness
	Break	10:10-10:30	
Section B	Chair: **Gang Yan**		

10:30-11:20	Chang-song Zhou 21	HKBU	Hierarchical Connectome Modes and Critical State Jointly Maximize Human Brain Functional Diversity
11:20-11:50	Min Yan 22	HKUST	The Retina Predicts Information in Inertial Stochastic Dynamics
11:50-12:20	Xiang-yu Ma 23	HKUST	Bayesian Model for Multisensory Perception
	Lunch	12:20-14:00	

Section C Chair: **Yoshiyuki Kabashima**

14:00-14:50	Gang Yan 24	Tongji University	Geometric embedding and machine learning reveal causal connectome from neuronal activities
14:50-15:40	Haiping Huang 25	Sun Yat-sen University	Physical laws of unsupervised learning in neural networks
	Break	15:40-16:10	

Section D Chair: **Pan Zhang**

16:10-16:40	Takashi Takahashi 26	Tokyo Tech	Replicated vector approximate message passing for resampling problem
16:40-17:10	Chuang Wang 27	IOA, CAS	A Solvable High-Dimensional Model of GAN
17:10-17:40	Yansong Chua 28	Huawei, 2012 Lab	An introduction to neuromorphic com puting and its cur rent state-of-art: a learning pe rspective
17:40-17:50	Haiping Huang	SYSU	Closing Remarks
	Dinner	18:30-20:00	

Abstracts of all talks

1. Function Space Entropy in Deep-Learning Networks

David Saad
Aston University

Recent engineering achievements of deep-learning machines have both impressed and intrigued the scientific community due to our limited theoretical understanding of the underlying reasons for their success. This work provides a general principled framework to investigate the function space of different types of deep-learning machines, based on the generating functional analysis. The framework facilitates studying the number of solutions of a given error around a reference multi-layer network. Exploring the function landscape of densely-connected networks, we uncover a general layer-by-layer learning behaviour, while the study of sparsely-connected networks indicates the advantage in having more layers for increasing generalization ability in such models. This framework accommodates other network architectures and computing elements, including networks with correlated weights, convolutional networks and discretised variables. It offers a useful tool for better understanding of the properties of deep-learning machines.

2. Guerra's interpolation techniques in Neural Networks and Machine Learning

Adriano Barra
Università del Salento, Italy

Starting with shallow networks, in this talk, I will at first show that what is learnt in a Restricted Boltzmann Machine (e.g. via contrastive divergence) can then be retrieved in a Hebbian fashion as in Hopfield neural networks. Then I will extend this duality discussing the multitasking capabilities of these networks and how their performances can be improved via “sleeping".

Finally I will enlarge such a scheme toward deep Boltzmann machines and dense associative networks, particularly focusing on the impressive signal-to-noise ratio that these machines are able to achieve in suitable operational regimes of pattern recognition. The underlying mathematical techniques will heavily rely on Guerra's interpolation schemes, in particular, on the transport equations lying behind the standard routes in the statistical mechanics of spin glasses.

3. Disentangling the roles of dimensionality and cell classes in neural computations

Alexis Dubreuil
École normale supérieure, France

The description of neural computations in the field of neuroscience relies on two competing views: (i) a classical single-cell view that relates the activity of individual neurons to sensory or behavioural variables, and focuses on how different cell classes map onto computations; (ii) a more recent population view that instead characterises computations in terms of collective neural trajectories, and focuses on the dimensionality of these trajectories as animals perform tasks. How the two key concepts of cell classes and low-dimensional trajectories interact to shape neural computations is however currently not understood. Here we address this question by combining machine-learning tools for training RNNs with reverse-engineering and theoretical analyses of network dynamics. We introduce a novel class of theoretically tractable recurrent networks: low-rank, mixture of Gaussian RNNs. In these networks, the rank of the connectivity controls the dimensionality of the dynamics, while the number of components in the Gaussian mixture corresponds to the number of cell classes. Using back-propagation, we determine the minimum rank and number of cell classes needed to implement neuroscience tasks of increasing complexity. We then exploit mean-field theory to reverse-engineer the obtained solutions and identify the respective roles of dimensionality and cell classes. We show that the rank determines the phase-space available for dynamics that implement input-output mappings, while having multiple cell classes allows networks to flexibly switch between different types of dynamics in the available phase-space. Our results have implications for the analysis of neuroscience experiments and the development of explainable AI.

4. Dynamical mean field theory of neural networks with quenched power-law disorder

Łukasz Kusmierz
RIKEN, Japan

Transitions to chaos have been previously extensively studied in different setups of randomly connected networks. The prevailing assumption is that, due to the central limit theorem, synaptic input can be modeled as a Gaussian random variable. In this scenario, a continuous transition has been found in rate models with smooth activation functions. However, these models do not take into account that neurons feature thresholds that cut off small inputs. With such thresholds, the transition to chaos in Gaussian networks becomes discontinuous, making it impossible for the network to stay close to the edge of chaos and to reproduce biologically relevant low activity states. Here we introduce a model with biologically motivated, heavy-tailed distribution of synaptic weights and analytically show that it exhibits a continuous transition to chaos. Notably, in this model the edge of chaos is associated with well-known avalanches. We validate our predictions in simulations of networks of binary as well as leaky integrate and fire neurons. Our results uncover an important functional role of non-Gaussian distributions of synaptic efficacy and suggest that their heavy tails may form a weak sparsity prior that can be useful in biological and artificial adaptive systems.

5. An integrative model interrelating dynamical and computational properties of cortical circuits

Pulin Gong,
University of Sydney

Cortical circuits constitute the fundamental computational units in the brain. In this talk, I will first present a new model of cortical circuits constrained by recent anatomical data. Based on this circuit model, I will then illustrate a dynamical working regime of the cortex, in which propagating activity patterns with complex spatiotemporal dynamics emerge. Such propagating patterns provide a mechanistic explanation for a great variety of neurophysiological data measured at different neural levels, including variable spiking dynamics and high-frequency oscillations with bursting properties. This dynamical working regime is analyzed by developing a method of fractional diffusion approximation. In addition, I will show that propagating activity patterns with complex dynamics provide a framework for understanding the neurophysiological mechanisms of cognitive functions such as attention.

6. Exploring the learning principle in the brain

Taro Toyoizumi
RIKEN, Japan

Animals adapt to the environment for survival. Synaptic plasticity is considered a major mechanism underlying this process. However, the best-known form of synaptic plasticity, i.e., Hebbian plasticity that depends on pre- and post-synaptic activity, can surge coincident activity in model neurons beyond a physiological range. Our lab has explored how neural circuits learn about the environment by synaptic plasticity. The instability of Hebbian plasticity could be mitigated by a global factor that modulates its outcome. For example, TNF-alpha that mediates homeostatic synaptic scaling is released by glia, reflecting the activity level of surrounding neurons. I show that a specific interaction of Hebbian plasticity with this global factor accounts for the time course of adaptation to the altered environment (Toyoizumi et al. 2015). At a more theoretical level, I ask what is the optimal synaptic plasticity rule for achieving an efficient representation of the environment. A solution is the error-gated Hebbian rule, whose update is proportional to the product of Hebbian change and a specific global factor. I show that this rule, suitable also in neuromorphic devices, robustly extracts hidden independent sources in the environment (Isomura and Toyoizumi 2016, 2018, 2019). Finally, I introduce that synapses change by intrinsic spine dynamics, even in the absence of synaptic plasticity. I show that physiological spine-volume distribution and stable cell assemblies are both achieved when intrinsic spine dynamics are augmented in a model (Humble et al.2019).

7. Approximate matrix completion based on cavity method

Chihiro Noguchi
Tokyo Tech, Japan

In order to solve large matrix completion problems with practical computational cost, an approximate approach based on matrix factorization has been widely used. Alternating least squares (ALS) and stochastic gradient descent (SGD) are two major algorithms to this end. In this study, we propose a new algorithm, namely cavity-based matrix factorization (CBMF) and approximate cavity-based matrix factorization (ACBMF), which are developed based on the cavity method from statistical mechanics. ALS yields solutions with less iterations when compared to those of SGD. This is because its update rules are described in a closed form although it entails higher computational cost. CBMF can also write its update rules in a closed form, and its computational cost is lower than that of ALS. ACBMF is proposed to compensate a disadvantage of CBMF in terms of relatively high memory cost. We experimentally

illustrate that the proposed methods outperform the two existing algorithms in terms of convergence speed per iteration, and it can work under the condition where observed entries are relatively fewer. Additionally, in contrast to SGD, (A)CBMF does not require scheduling of the learning rate.

8. Capacity of Dimension Selection with Simple Perceptron

Yingying Xu
Tokyo Tech, Japan

In classification problem, it is nontrivial which elements of the pattern vectors crucially associate with the classification. We show analysis on the optimal selection of the elements of the pattern vectors by assessing the classification problem with the simple perceptron. To validate our analysis, we demonstrate the numerical experiments following a greedy algorithm to select the relevant element for selection. As a consequence, we attain a larger capacity of the so-called Gardner-Cover.

9. Understanding the computational difficulty of binary-weight perceptron and the advantage of input sparseness

Zedong Bi
HKBU

Limited precision of synaptic weights is a key aspect of both biological and hardware implementation of neural networks. To assign low-precise weights during learning is a non-trivial task, but may benefit from representing to-be-learned items using sparse code. However, the computational difficulty resulting from low weight precision and the advantage of sparse coding remain not fully understood. Here, we study a perceptron model, which associates binary (0 or 1) input patterns with desired outputs using binary (0 or 1) weights, modeling a single neuron receiving excitatory inputs. Two efficient perceptron solvers (SBPI and rBP) usually find solutions in dense solution region. We studied this dense-region-approaching process through a decimation algorithm, where every time step, marginal probabilities of unfixed weights were evaluated, then the most polarized weight was fixed at its preferred value. We compared decimation-fixing order of weights with the dynamics of SBPI and rBP, and studied the structure of solution subspace $ \mathcal{S}$ where early-decimation-fixed weights take their fixed values. We found that in SBPI and rBP, most time steps are spent on determining values of late-decimation-fixed weights. This algorithmic difficult point may result from strong cross-correlation between late-decimation-fixed weights in $\mathcal{S}$, and is related to solution condensation in $\mathcal{S}$ during decimation. Input sparseness reduces time steps that SBPI and rBP need to find solutions, due to reduction of cross-correlation between late-decimation-fixed weights. When constraint density increases across a transition point, the portion of weights whose values are difficult to determine sharply increases, signaling difficult accessibility of dense solution region. We propose that computational difficulty of binary-weight perceptron is related to a solution condensation process during approaching dense solution region. Our work highlights the heterogeneity of learning dynamics of weights, which may help understand axonal pruning in brain development, and inspire more efficient algorithms to train neural networks.

10. Push-pull Feedback Implements Rough-to-fine Information Processing

Si Wu
Peking University

Experimental data has revealed that in addition to feedforward connections, there exist abundant feedback connections in a hierarchical neural pathway. Although the importance of feedback in neural information processing has been widely recognized in the field, the detailed mechanism of how it works remains largely unknown. Here, we investigate the role of feedback in hierarchical memory retrieval. Specifically, we consider a multi-layer network which stores hierarchical memory patterns, and each layer of the network behaves as an associative memory of the corresponding hierarchy. We find that to achieve good retrieval performance, the feedback needs to be dynamical: at the early phase, the feedback is positive (push), which suppresses inter-class noises between memory patterns; at the late phase, the feedback is negative (pull), which suppresses intra-class noises between memory patterns. Overall, memory retrieval in the network progresses from rough to fine. Our model agrees with the push-pull phenomenon observed in neural data and sheds light on our understanding of the role of feedback in neural information processing.

11. Integrating and segregating information from different channels in the Brain

K Y Michael Wong
HKUST

Neural systems gather information from different channels resulting in enhanced reliability. The optimal estimate is given by Bayes' rule, and remarkably, experiments on macaques showed that the brain can achieve this optimum. It is therefore interesting to consider the neural architecture and mechanism underlying this feat. Nevertheless, this leads to some interesting information issues to be addressed. For channel inputs that are correlated in their prior distribution, where and how is the information stored? Furthermore, when the channel inputs are partially correlated, how can the system perform partial integration? Besides performing integration when the channel inputs are correlated, how can the system perform segregation when they are not? Finally, since input integration implies that the system extracts the information common to the channels, can the information of the individual channels be recovered? To answer these questions, I will discuss the results of our study on a decentralized network architecture where same-channel and cross-channel information are processed in parallel. I will provide an explanation of the role played by opposite neurons in the VIP and MSTd areas of the brain, and report a striking discovery that the direct and indirect cross-channel pathways are opposite to each other -- an apparently redundant architecture.

12. Deep networks with random representations

Jonathan Kadmon
Stanford University

Recent studies of deep networks have used mean-field methods to describe the propagation of signals through large networks with random connectivity. However, in trained networks, or neural circuits with local plasticity, synaptic weights are not expected to be independent. In this talk, I will present a model of a feedforward network with dense, structured connectivity. Here, local plasticity rules associate between internal representations, which are chosen randomly. Importantly, the patterns and weights do not depend on the data or labels. Defining a simple input model, I will show that the performance depends on three parameters: The typical number of neurons active per pattern; the number of internal patterns per neuron; and the depth of the network. Notably, for sparse networks, the performance is agnostic to the underlying plasticity rule. Finally, I will show that for given input statistics, there is an optimal depth that solves the problem with the minimal number of neurons. I will discuss the relevance of our model to biology and as a kernel method.

13. A macroscopic description of excitation-inhibition balanced network with synaptic kinetics

Jun-hao Liang
SYSU

Neuronal network models in the framework of excitation-inhibition balance have been widely used in studying diverse topics in computational neuroscience, as they can capture the spiking irregularity feature of cortical neurons. However, analytical tools to study the dynamics of E-I balanced network are still very limited. In this talk, I will first introduce the classical mean-field theory of E-I balance of binary neuronal networks and show that it fails to predict the network dynamics if realistic synaptic filtering kinetic is included. Next, I will introduce the first-spiking time theory on analyzing the firing rate of integrate-and-fire neurons with synaptic kinetics and point out the difficulties for generalizing it to E-I networks. Finally, I will introduce a novel mean-field framework we proposed to study the macroscopic dynamics of E-I balanced networks and show that it successfully predicts the critical transition from asynchronous state to synchronous state induced by synaptic kinetics. I then end up with some discussion of its potential applications.

14. On Replica-BP correspondence in rotationally invariant SG models

Yoshiyuki Kabashima
Tokyo Tech, Japan

In Sherrington-Kirkpatrick model, it is known that (1) the macroscopic dynamics of belief propagation (BP), which is sometimes referred to as ``state evolution'' of ``approximate message passing'' these days, is described by the naive iterative substitution of the replica symmetric (RS) saddle point equation, and (2) the instability condition of BP's fixed point accords with the de Almeida-Thouless condition of the RS solution. (1) and (2) also hold when the coupling matrix is given by the Gram matrix of random matrices composed of independent entries from identical distributions. We discuss whether these correspondeces are further generalized or not for spin glass models characterized by rotationally invariant matrix ensembles.

15. Solving Statistical Mechanics using Variational Autoregressive Networks

Pan Zhang
ITP, CAS

Computing free energy, estimating physical quantities, and generating uncorrelated samples are fundamental problems in statistical mechanics. In this talk I will introduce a new framework for solving the statistical mechanics problems for systems with a finite size. The approach extends the celebrated variational mean-field approaches using autoregressive networks, a neural network model which supports direct sampling and exact calculation of normalized probability of configurations. Training of the network employs the policy gradient approach in reinforcement learning, which unbiasedly estimates the gradient of variational parameters. We apply our approach to several classic systems, including 2-d Ising models, Hopfield model, Sherrington--Kirkpatrick spin glasses, and the inverse Ising model, for demonstrating its advantages over exiting variational mean-field methods.

16. Objective and efficient inference for couplings in neuronal networks

Yu Terada
RIKEN, Japan

Inferring directional couplings from the spike data of networks is desired in various scientific fields such as neuroscience. Here, we apply a recently proposed objective procedure to the spike data obtained from the Hodgkin-Huxley type models and in vitro neuronal networks cultured in a circular structure. As a result, we succeed in reconstructing synaptic connections accurately from the evoked activity as well as the spontaneous one. To obtain the results, we invent an analytic formula approximately implementing a method of screening relevant couplings. This significantly reduces the computational cost of the screening method employed in the proposed objective procedure, making it possible to treat large-size systems as in this study.

17. Cross validation in sparse linear regression with piecewise continuous non-convex penalties and its acceleration

Tomoyuki Obuchi
Tokyo Tech, Japan

We investigate the signal reconstruction performance of sparse linear regression in the presence of noise when piecewise continuous nonconvex penalties are used. Among such penalties, we focus on the smoothly clipped absolute deviation (SCAD) penalty. The contributions of this study are three-fold: We first present a theoretical analysis of a typical reconstruction performance, using the replica method, under the assumption that each component of the design matrix is given as an independent and identically distributed (i.i.d.) Gaussian variable. This clarifies the superiority of the SCAD estimator compared with ℓ_1 in a wide parameter range, although the nonconvex nature of the penalty tends to lead to solution multiplicity in certain regions. This multiplicity is shown to be connected to replica symmetry breaking in the spin-glass theory, and associated phase diagrams are given. We also show that the global minimum of the mean square error between the

estimator and the true signal is located in the replica symmetric phase. Second, we develop an approximate formula efficiently computing the cross-validation error without actually conducting the cross-validation, which is also applicable to the non-i.i.d. design matrices. It is shown that this formula is only applicable to the unique solution region and tends to be unstable in the multiple solution region. We implement instability detection procedures, which allows the approximate formula to stand alone and resultantly enables us to draw phase diagrams for any specific dataset. Third, we propose an annealing procedure, called nonconvexity annealing, to obtain the solution path efficiently. Numerical simulations are conducted on simulated datasets to examine these results to verify the consistency of the theoretical results and the efficiency of the approximate formula and nonconvexity annealing.

18. Generalized Belief Propagation in Gaussian Markov Random Fields

Jun-tao Wang
HKUST

We propose the Gaussian Generalized Belief Propagation (GaGBP) for Gaussian Markov Random Fields (GMRFs) of arbitrary topology that can be applied to solve graph inference problems. It is a local iterative algorithm and gives accurate means and variances estimations for marginal distributions. The formalism is derived based on the Cluster Variation Method (CVM), where the GMRF is divided into regions, which are subsets of random variables, and approximated by a collection of variational local beliefs of the regions. We also introduce an original way to construct the set of regions for GMRFs of arbitrary topology based on the minimum cycle basis, so that it is also applicable to irregular graphs. We demonstrated that empirically our formalism provides more accurate approximations for the marginal distributions of GMRFs of arbitrary topology than the conventional Gaussian Belief Propagation (GaBP) and requires much fewer iterations.

19. A large deviations statistical physics approach to finding wide flat minima in neural networks landscapes

Carlo Baldassi
Bocconi University

Learning in Deep Neural Networks takes place by minimizing a non-convex high-dimensional loss function. In current practice, the learning process is often observed to somehow avoid getting stuck in local critical points, and to get to good minima that avoid overfitting. How these two features can be kept under control in nonlinear devices composed of millions of tunable connections is a profound and far reaching open question. In recent years, we have developed a framework based on a large-deviation statistical physics analysis of simplified models (and corroborated by numerical simulations on more realistic ones) that suggests some answers in terms of the geometrical properties of the optimization landscape. We have found that, in a surprisingly wide variety of models, "bad" minima coexist with rare "wide flat minima" regions (or "high local entropy" regions) with good generalization properties. We have developed a series of algorithmic schemes that make such regions accessible (ranging from MCMC, to message-passing, to local search algorithms, to quantum annealing). We have also found that some of the heuristic techniques used in current neural network practice bias the search towards such wide flat minima. Finally, we have observed that targeting such regions can significantly boost automatic feature extraction in unsupervised models (autoencoders).

20. Towards Identifying Places and Boundaries of Consciousness

Masafumi Oizumi
Tokyo University

It has been a long-standing question “where” consciousness resides in the brain. The problem of identifying the place of consciousness can be restated

as the problem of identifying the boundary of consciousness, i.e., the problem of drawing the boundary in a complex neural network, which determines the place of consciousness. Although a lot of experimental findings have been accumulated over the last decades, the boundary problem of consciousness has not yet been resolved. In this talk, I will discuss the information theoretical approach to the boundary problem of consciousness based on Integrated Information Theory (IIT). IIT is an attempt to mathematically quantify consciousness from the viewpoint of information and integration, which are considered to be the essential properties of consciousness. IIT hypothesizes that the place of consciousness corresponds to the locally most "integrated" subnetwork in the brain where the amount of integrated information is locally maximum. Finding the most integrated subnetwork (called a complex or an information core) in a large network is extremely difficult because it involves optimization problems which require an exponentially large amount of computational time. To resolve the difficulty, we have developed efficient algorithms that reduce the computational time to polynomial order, which enables us to find a complex within a reasonable amount of time. I will introduce several applications of the proposed algorithms to real neural data and discuss how we can test the theoretical predictions about the places and boundaries of consciousness by experiments.

21. Hierarchical Connectome Modes and Critical State Jointly Maximize Human Brain Functional Diversity

Chang-song Zhou
HKBU

The brain requires diverse segregated and integrated processing to perform normal functions in terms of anatomical structure and self-organized dynamics with critical features, but the fundamental relationships between the complex structural connectome, critical state and functional diversity remain unknown. Herein, we extend eigenmode analysis to investigate the joint contribution of hierarchical modular structural organization and critical state to brain functional diversity. We show that the structural modes inherent to the hierarchical modular structural connectome allow a nested functional segregation and integration across multiple spatiotemporal scales. The real brain hierarchical modular organization provides large structural capacity for diverse functional interactions, which are generated by sequentially activating and recruiting the hierarchical connectome modes, and the critical state can best explore the capacity to maximize the functional diversity. Our results reveal structural and dynamical mechanisms that jointly support a balanced segregated and integrated brain processing with diverse functional interactions, and they also shed light on dysfunctional segregation and integration in neurodegenerative diseases and neuropsychiatric disorders.

22. The Retina Predicts Information in Inertial Stochastic Dynamics

Min Yan
HKUST

Vision is one of the most important sensory modalities in animals. Experiments showed that the retina of salamander and rabbit[1] not only receive visual signals, but also process the information they received by making predictions. We performed experiments on the retina of bullfrog and found that this ‘anticipative coding’ phenomenon depends on the dynamics of the input. When the bullfrog retina is shown a moving bar, whose positions are calculated from the Hidden Markov Model (HMM), the mutual information between the responses of the bullfrog retina and the input positions as a function of their time differences reveal that the responses of the retina have correlations with subsequent visual signals. However, when the moving bar positions are calculated from Ornstein–Uhlenbeck (OU) process, the correlations disappear. The ‘predictive’ and ‘non-predictive’ behaviors of the bullfrog retina show that when the inputs are inertial (HMM), the retina is able to anticipate the subsequent signals. Otherwise, when the inputs do not contain inertia or momentum information (OU Process), the retina is not able to make predictions about forthcoming signals. We propose a neural network model to simulate this kind of anticipative behaviors, involving neural populations of ganglion cells gated by amacrine cells. Our simulations show that the neural network model can achieve anticipation for HMM inputs. Besides, we found that individual neurons in our neural network model can also realize anticipation, in accordance with experimental results.

23. Bayesian Model for Multisensory Perception

Xiangyu Ma
HKUST

Diverse modalities are able to provide the neural system with complementary information. Experimental data show that in the dorsal medial superior temporal (MSTd) area and the ventral intraparietal (VIP) area, there exist two types of neurons, congruent and opposite cells. Here, we consider a recently proposed model in which the congruent and opposite neurons play a role in multisensory integration and segregation respectively. Multisensory integration and segregation are important for the survival of animals. Experimental data indicated that the brain processes information in an optimal way according to Bayes rule. The model suggested by Zhang et al. has demonstrated that two different neuron groups are related to the multisensory integration and segregation simultaneously. However, the model is suboptimal for intermediate disparity of the inputs. In our recent work, we developed a novel method for revealing the dynamics behind this neural network and discuss the conditions for our new model to achieve Bayesian inference. Furthermore, we also suggest that an output-dependent noise and a network structure accommodating the independent prior are necessary for the whole story.

24. Geometric embedding and machine learning reveal causal connectome from neuronal activities

Gang Yan
Tongji University

The connections between neurons or brain regions represent an indispensable foundation for neurobiological research, forming an emerging field of network neuroscience. Despite recent advances in structural analyses of various brain networks and in probing the relation between structure and functions, most studies rely on functional networks where each edge represents the correlation between two activity series. Such correlation networks lack the predictive power for causality and information flow in brain. In this talk we will present a method based on geometric embedding and machine learning for reconstructing causal connectome from neuronal activities, which is validated in a wide range of dynamical network models. By applying this method to real datasets we revealed novel patterns of information flow in brains.

25. Physical laws of unsupervised learning in neural networks

Haiping Huang
SYSU

We provide a complete picture about mechanisms of unsupervised learning, i.e., in a simple neural network system, unsupervised learning can be interpreted as breaking a series of symmetries, driven by increasing observations. First, it is the spontaneous symmetry breaking starting the concept-formation, then the permutation symmetry among hidden units is spontaneously broken, including in sequence two types: the student side is the first, and the teacher (planted truth) is the second. In this paper, we also analytically prove that the learning threshold in a simple neural network that triggers a spontaneous symmetry breaking (concept-formation) does not depend on the number of hidden neurons (here two for a minimal model) once this number is finite, for the correlation-free case. The underlying physics is the partition function factorization for which we give a proof. Moreover, our analytical result reveals that the weak correlation among receptive fields of hidden neurons significantly reduces the learning threshold, which is consistent with the non-redundant weight assumption popular in system neuroscience and machine learning! By studying the minimal model, we are excited to reveal the inner workings of unsupervised learning, a fundamental process governing artificial and biological intelligence. We expect this work may open doors towards physical laws governing neural learning.

26. Replicated vector approximate message passing for resampling problem

Takashi Takahashi
Tokyo Tech

Resampling techniques are widely used in statistical inference and ensemble learning, in which estimators' statistical properties are essential. However, existing methods are computationally demanding, because repetitions of estimation/learning via numerical optimization/integral for each resampled data are required. In this study, we introduce a computationally efficient method to resolve such problem: replicated vector approximate message passing. This is based on a combination of the replica method of statistical physics and an accurate approximate inference algorithm, namely the vector approximate message passing of information theory. The method provides tractable densities without repeating estimation/learning, and the densities approximately offer an arbitrary degree of the estimators' moment in practical time. In the experiment, we apply the proposed method to the stability selection method, which is commonly used in variable selection problems. The numerical results show its fast convergence and high approximation accuracy for problems involving both synthetic and real-world datasets.

27. A Solvable High-Dimensional Model of GAN

Chuang Wang
IOA, CAS

We present a theoretical analysis of the training process for a shallow GAN model fed by high-dimensional input data. The training dynamics of the proposed model at both microscopic and macroscopic scales can be exactly analyzed in the high-dimensional limit. In particular, we prove that the macroscopic quantities measuring the quality of the training process converge to a deterministic process characterized by an ordinary differential equation (ODE), whereas the microscopic states containing all the detailed weights remain stochastic, whose dynamics can be described by a stochastic differential equation (SDE). This analysis provide a new perspective different from recent analyses in the limit of small learning rate, where the microscopic state is always considered deterministic, and the contribution of noise is ignored. From our analysis, we show that the level of the background noise is essential to the convergence of the training process: setting the noise level too strong leads to failure of feature recovery, whereas setting the noise too weak causes oscillation. Although this work focuses on a simple copy model of GAN, we believe the analysis methods and insights developed here would prove useful in the theoretical understanding of other variants of GANs with more advanced training algorithms.

28. An introduction to neuromorphic computing and its current state-of-art: a learning perspective

Chua Yansong
Huawei, 2012 Lab

Neuromorphic computing has received much hype in recent years. With the advent of neuromorphic chips such as IBM True North, Intel Loihi, and Tsinghua Tianjic chip, the field has chosen to first develop specialized hard-

ware while neuromorphic computing algorithms and applications are still work in progress. Neural networks for neuromorphic computing typically refers to event-based spiking neural networks (SNNs), even though it need not be so. In this talk, I would discuss some of the state-of-art approaches in the field, such as conversion of pre-trained artificial neural networks (ANNs) to SNNs, using back propagation in time (BPTT) to train deep SNNs, and also some of the progress made in local learning, which is more hardware-friendly. Finally I would present some of our recent works, which hopes to address some of the areas in the field, namely, a neuromorphic dataset, event-based encoding of speech data, and training of a SNN for classifying Imagenet with comparable accuracy but much better inference speed and energy efficiency.

Participant List

Adriano Barra	Università del Salento	
Carlo Baldassi	Bocconi University	
Alexis Dubreuil	ENS, France	
Pulin Gong/ 龚璞林	Sydney University	
Haiping Huang/黄海平	SYSU	
Jonathan Kadmon	Stanford University	
Yoshiyuki Kabashima	Tokyo Tech	
Masafumi Oizumi	Tokyo University	
David Saad	Aston University	
Taro Toyoizumi	RIKEN	
K Y Michael Wong/王国彝	HKUST	
Si Wu/吴思	Peking University	
Gang Yan/严钢	Tongji University	
Changsong Zhou/周昌松	HKBU	
Pan Zhang/张潘	ITP, CAS	
Zedong Bi/毕则栋	HKBU	
Yansong Chua/蔡炎松	2012 Lab, Hua Wei	
Chihiro Noguchi	Tokyo Tech	
Łukasz Kusmierz	RIKEN	
Junhao Liang/梁俊豪	SYSU	
Xiangyu Ma/马翔宇	HKUST	
Tomoyuki Obuchi	Tokyo Tech	
Yu Terada	RIKEN	
Takashi Takahashi	Tokyo Tech	

Juntao Wang/王均涛	HKUST	
Chuang Wang/王闯	IOA，CAS	
Yingying Xu/许滢滢	Tokyo Tech	
Min Yan/严敏	HKUST	
Youfang Yan/闫友房	Shanxi Normal University	
Geng Li/李耿	ITP, CAS	
Shangnan Wang/王尚楠	ITP, CAS	
Kwan Tung Li/李军彤	HKBU	
Daniel Kristanto	HKBU	
Yuqi Liang/梁雨琪	HKBU	
Xinran Ma/马欣然	Beijing Normal University	
Tingting Gao/高婷婷	Tongji University	
Xinya Zhang/张昕亚	Tongji University	
Xinjie Zhang/张鑫洁	Tongji University	
Weifeng Yang/杨玮枫	Shantou University	
Jie Li/李洁	Shantou University	
Xiwang Liu/刘希望	Shantou University	
Yan Zhang/张琰	Guangdong University of Foreign Studies	
Jack Murdoch Moore	Tongji University	
Wenkuo Cui/崔文阔	Tongji University	
Qi Zhang/张奇	Zhejiang University	
Zhe Yuan/袁喆	Beijing Normal University	
Lu Yang/杨璐	Hua Wei, Hong Kong	

Wenting Sheng/盛文婷	Central China Normal University	
Qing Yu/余青	Central China Normal University	
Zhihang Liu/刘智航	Central China Normal University	
Yuxiang Yang/杨寓翔	Central China Normal University	
Fei Ma/马飞	Central China Normal University	
Tao Zhang/张陶	Central China Normal University	
Jianmin Shen/申建民	Central China Normal University	
Hongrui Fan/范竑锐	Central China Normal University	
Jinyu Li/李晋豫		
Bo Li/黎勃	Aston University	
Jinhua Zhao/赵金华	South China Normal University	
Ning Liu/刘宁	Beijing Normal University	
Yiming Liu/刘一鸣	Beijing Normal University	
Jian Zhang/张健	Beijing Normal University	
Lu Chen/陈璐	Beijing Normal University	
Yifan Sun/孙怡帆	Renmin University of China	
Chunyang Wang/王春阳	Ludong University	
Rui Zhang/章瑞	Hua Wei, Hong Kong	
Zhengqi He/	RIKEN	
Tianyi Wu/伍天一	Peking University	

Fangyan Ouyang/欧阳昉艳	Communication University of Zhejiang	
Kun Yan/严崐	SYSU	
Long Xiong/熊龙	Zhejiang University	
Wenyong Li/李文湧	SYSU	
Yuanyuan Mi/弭元元	Chongqing University	
Mingyang Bi/毕铭阳	Xiamen University	
Zhen Li/李震	Hua Wei, Hong Kong	
Kaiyuan Cao/曹凯源	Nanjing Normal University	
Nancheng zheng/郑南城	SYSU	
Chan Li/李婵	SYSU	
Dongjie Zhou/周东劼	SYSU	
Wenxuan Zou/邹文轩	SYSU	
Jiangwen Zhou/周健文	SYSU	
Zijian Jiang/蒋子健	SYSU	
Tianqi Hou/侯天齐	HKUST	
Xiaohui Deng/邓晓辉	HKUST	
Zhengye Huang/黄振业	SYSU	
Junbin Qiu/邱俊斌	SYSU	
Leyi Li/李乐仪	SYSU	
Yating Deng/邓雅婷	SYSU	
Xiaorui Zhu/朱晓睿	Beijing Normal University	
Xiaoming Wang/王小明	South China Normal University	

Jiazheng Ma/马嘉政	SYSU	
Yajun Zhang/张亚君	Beijing Normal University	
Jie Sun/孙杰	Hua Wei, Hong Kong	
Mengmeng Jin/金蒙蒙	Hangzhou Dianzi University	

Acknowledgement

Financial supports from NSFC, SYSU and two companies of high performance computing.

扫一扫关注五舟微信

Notes Pages