DAIKIN International Symposium on Physics of Intelligence

Statistical Mechanics and Machine Learning: A Powerful Combination for Data Analysis

Nov. 6-8, 2024@Koshiba Hall, The University of Tokyo

Poster List

Poster 1

Yohei Saika

National Institute of Technology, Gunma College

Dynamical Properties of Air Conditioning Using Thermal Comfort Based on the Bayesian Inference Using the EAP Estimation

On the basis of the theoretical viewpoint using statistical mechanics of information, we investigate dynamical properties of air conditioning using the thermal comfort in small-scale area based on the Bayesian inference using the expected a posterior (EAP) estimation and a sensor network system composed of a set of data loggers observing both temperature and relative humidity set in the small-scale area.

Then, in order to analytically examine the time evolution of the environmental variables, we here utilize the Bayesian inference using the EAP estimation by making use of the mean-field theory for the full-connected model and the Master equation.

In this method, we use the model of the true prior enhancing both the optimum of the thermal comfort and the similarity of the temperature (relative humidity) between sampling points.

Then, we use the likelihood function expressing the models both of the air cooling and dehumidification at each sampling point.

We find that the thermal comfort converges to the optimal value most smoothly under the Bayes-optimal condition around the Bayes-optimal condition.

Also, we find that the thermal comfort converges to almost its optimum most smoothly, if we use dehumidification and stabilizing the optimum of the thermal comfort.

Poster 2

Ryo Igarashi
OMRON SINIC X Corporation

NeSF and Crystalformer: A New Machine Learning Framework for Crystal Structures

We developed new Crystalformer encoder and Neural Structural Field(NeSF) decoder to express crystal structures in the machine learning framework. Crystalformer is a simple yet effective transformer-based encoder for crystal structures, which handles infinitely-connecting attention network correctly. NeSF is inspired by the concepts of vector fields in physics and implicit neural representations in computer vision, and considers a crystal structure as a continuous field rather than as a discrete set of atoms. We show the promising results of these architectures and the connection to physics.

Poster 3

Kai Nakaishi
The University of Tokyo

Critical Phase Transition in a Large Language Model

The performance of large language models (LLMs) strongly depends on the temperature parameter, which controls the sharpness of the distribution of next tokens. Empirically, at very low temperatures, LLMs generate sentences with clear repetitive structures, while at very high temperatures, the generated sentences are often incomprehensible. In this presentation, we will present the results of extensive numerical analyses of sequences generated by GPT-2, a representative LLM, to demonstrate that this temperature-induced change is a phase transition with critical phenomena, such as a power-law decay of correlation and dynamical slowing down. In addition, our analysis of natural language corpora shows that GPT-2 at the critical point shares various statistical properties with natural languages. We also discuss intriguing phenomena at both high and low temperatures that are not observed in equilibrium statistical mechanical systems.

Poster 4

Yoshihiko Nishikawa
Kitasato University

Fast free-energy estimation using simulated tempering with Wang-Landau-type feedback

Simulated tempering, proposed by Marinari and Parisi in 1992, is an extended-ensemble Monte Carlo algorithm that changes control parameters, e.g. temperature, as well as microscopic degrees of freedom during a run. Although it can efficiently explore the low-temperature states of glassy systems, it needs the free energy of the systems for optimal performance, which is notoriously difficult to estimate. This strongly contrasts with parallel tempering, as parallel tempering runs without any prior knowledge of the system. Here, we propose a new simulated tempering algorithm to estimate the free energy with high precision. Our algorithm iteratively modify the estimate of the free energy with the Wang-Landau-type feedback, eventually converging to the true free energy. We apply the algorithm to the 3d Edwards-Anderson (EA) spin glass model and find that the algorithm quickly converges. We also show that optimal simulated tempering with the nonreversible dynamics outperforms parallel tempering, allowing us to access low-temperature equilibrium states deep inside the spin glass phase of the 3d EA spin glass model, with up to 64^3 spins. Some possible applications of the algorithm to machine learning, soft matter, and biophysics problems will be discussed.

Poster 5

Yuanbo Wang
The University of Tokyo, Frontier Science

Combined with neurogenesis of contrastive learning

Inspired by biological neural generation, where the continuous generation of neurons enhances the brain's ability to separate similar input patterns, we use contrastive learning to quantify this ability of the artificial neuron network. Overall, we use contrastive learning to measure the network's recognition capability and employ neural generation to influence this capability, ultimately developing a method for generating different networks for various data types.

Poster 6

Pablo Morales

Araya

Higher-Order Interactions effects from The Maximum Entropy Principle of Curved Statistical Manifolds

Introduction of information-theoretic methods to understand physical phenomena has continued to gain traction in the physics community, being promising for enabling new effective methods of analysis and opening new avenues of inquiry. From the discovery of the Ryu-Takayanagi formula for the entanglement entropy (Phys. Rev. Lett. 96, 181602) more than a decade ago to powerful tools to study quantum gravity and quantum field theory. However, many physical systems of interest cannot be described by the standard tools of information theory, and therefore extensions are required. There have been developments in non-Shannonian methods, but the lack of a unifying principle or encompassing framework, renders such approaches — despite their efficacy — heuristic at best.

In this talk, we address this knowledge gap by introducing a method grounded in the maximum entropy principle applied to curved statistical manifolds. One of the consequences of this extension, is the induction of higher-order interactions. This feature opens the door to a plethora of future theoretical and practical investigations ranging from neural-networks to neuron systems. Importantly, it allows for the relatively simple construction of a model that exhibit very complex phenomena, including the appearance of explosive phase transitions.

Poster 7

Suriyanarayanan Vaikuntanathan

University of Chicago

Generative diffusion with active non equilibrium dynamics

The statistical mechanical basis of generative diffusion has been a topic of much interest lately. In our work, we suggest that ideas from non-equilibrium statistical mechanics can be used to further fine tune the performance of generative diffusion models. In particular, using ideas developed in the context of the physics of active matter systems, our work shows how generative diffusion models can be made to learn (and potentially generalize) data distributions more efficiently. The talk will discuss some of the basic theoretical underpinnings of generative diffusion, how they can be expanded using ideas from non-equilibrium statistical mechanics, and results on some model systems

Poster 8

Mikiya Doi

Graduate School of Information Sciences, Tohoku University

Statistical mechanical analysis of signal recovery phase transition in binary compressed sensing based on L1-norm minimization

Compressed sensing is a signal processing scheme that reconstructs high-dimensional sparse signals from a limited number of observations.

Recently, in the field of compressed sensing, various problems involving signals with a finite number of discrete values have attracted attention in contexts such as digital communications.

In particular, binary compressed sensing, which restricts signal elements to binary values {0, 1}, is the most fundamental and straightforward analysis subject in such problem settings. We evaluate the typical performance of noiseless binary compressed sensing based on L1-norm minimization using the replica method, a statistical mechanical approach. We analyze a general setting where the elements of the observation matrix follow a Gaussian distribution, including a non-zero mean. We also compare the results of the numerical simulations with our theoretical analysis. We demonstrate that the biased observation matrix indicates more reconstruction success conditions in binary compressed sensing.

Similar results have already been reported in [S. Keiper 2017, A. Flinth & S. Keiper 2019]. In our poster presentation, we will discuss the derivation of the phase diagrams by the replica method, which differs from their approach, and the behavior of the approximate message passing algorithm and state evolution in our problem settings.

Poster 9

Ryota Kawasumi,
Kobe Gakuin University
Koujin Takeda
Ibaraki University

Effectiveness of automatic hyperparameter tuning algorithm in sparse matrix factorization

In our previous work, an analytical solution for sparse matrix factorization with Laplace prior was obtained by variational Bayesian methods under the approximations of mean-field approximation, expansion of the prior up to the first order, and neglect of covariance. Based on the above solution, a novel numerical method for tuning hyperparameter of Laplace prior is proposed by evaluating the zero point of the normalization factor in the sparse matrix prior. Numerical experiment shows that the performance of the sparse factorization matrix reconstruction is comparable to that of sparse PCA, a widely used sparse matrix factorization algorithm. Additionally, to verify the practicality of the proposed method, an experiment is conducted to extract the dictionary matrix from real images, whose result shows that the proposed method can extract the appropriate matrix.

Poster 10

Yuma Ichikawa
Graduate School of Arts and Sciences, The University of Tokyo, Fujitsu Limited

Statistical-mechanical analysis of variational autoencoders

Variational autoencoders (VAEs) face a notorious problem wherein the variational posterior often aligns closely with the prior, a phenomenon known as posterior collapse, which hinders the quality of representation learning. To mitigate this problem, an adjustable hyperparameter beta and a strategy for annealing this parameter, called KL annealing, have been proposed. This study presents a theoretical analysis of the learning dynamics in a

minimal VAE. It is rigorously proved that the dynamics converge to a deterministic process within the limit of large input dimensions, thereby enabling a detailed dynamical analysis of the generalization error. Furthermore, the analysis shows that the VAE initially learns entangled representations and gradually acquires disentangled representations. A fixed-point analysis of the deterministic process reveals that when beta exceeds a certain threshold, posterior collapse becomes inevitable regardless of the learning period. Additionally, the superfluous latent variables for the data-generative factors lead to overfitting of the background noise; this adversely affects both generalization and learning convergence. Furthermore, we demonstrate that the stable fixed points of the dynamics coincide with the optimal values obtained through the replica method, showing that global optimal solutions are achievable under specific conditions.

Poster 11

Angelo Giorgio Cavaliere
Cybermedia Center

Learning by message passing in the multilayer perceptron

Deep networks of perceptrons have been recently addressed via the replica method [1][2], which in a dense but not fully-connected regime can be solved exactly, suggesting learning in a deep network happens in a highly heterogeneous manner. In this approach, the network is considered as a physical system with local interactions and fixed boundary conditions at the input and output layers. Message passing equations can be written for this system and solved numerically on real instances of the problem, provided that the permutation symmetry of neurons in the hidden layers is explicitly broken. We apply and discuss these techniques in the simple case of a Bayes optimal, teacher-student setting.

[1] Hajime Yoshino, SciPost Phys. Core 2, 005 (2020)
[2] Hajime Yoshino, arXiv:2302.07419 (2023)

Poster 12

Kaito Takanami
The University of Tokyo

The Role of Soft Labels in Self-knowledge distillation

Knowledge distillation is widely used as a technique for transferring knowledge from a large teacher model to a smaller student model. A variation of this method, known as self-knowledge distillation, where the same model serves as both teacher and student, has been empirically shown to improve generalization performance. However, the theoretical understanding of what occurs internally during self-knowledge distillation remains limited. In this study, we analyze the role of teacher-generated labels in self-distillation by investigating their effects within a classification task using a single-layer neural network. Our aim is to focus on the mechanisms through which these labels influence model performance in self-knowledge distillation settings.

Poster 13

Yoshinori Hara
The University of Tokyo

Stability Analysis of Quantum Spin Glasses

Quantum phase transitions exhibited by mean-field spin glass models have attracted growing attention because they allow analytical characterization of the possibilities and limitations of quantum annealing.

However, so far, detailed analysis has been almost exclusively restricted to the Sherrington-Kirkpatrick model.

In this work, we generalize the analysis to more general systems with random and rotationally invariant coupling matrices.

These quantum systems can be analyzed by mapping them to classical spin systems extended in the imaginary time direction using the Suzuki-Trotter decomposition, which makes the resultant order parameters generally dependent on the imaginary time. However, most of the previous studies assume that the order parameters except for the two-point correlation function are uniform in imaginary time. We term this treatment the quasi-static approximation (qSA). To our knowledge, the validity of qSA has not been thoroughly examined. We obtained the following results for this issue:

•We derived a self-consistent equation for determining the order parameters without using qSA. The solution of qSA represents a special solution of the derived self-consistent equation.

•For the solution, we derived the local stability condition against breaking of the imaginary-time uniformity.We also obtained the local stability condition for replica symmetry breaking (RSB).

•The critical mode for RSB is uniform in imaginary time. In addition, unless RSB occurs, the qSA-type solution is locally stable. Further, these can be generalized to an arbitrary step of RSB.

Poster 14

Takahiro Kanazawa

The University of Tokyo

Universality in the dynamical phase transitions of Brownian motion

Dynamical phase transitions (DPTs) arise from qualitative changes in the long-time behavior of stochastic trajectories. Here we discover first-order DPTs in a single Brownian particle without drift when the dimension is higher than four [arXiv:2407.14090, arXiv:2407.18282]. The DPTs accompany temporal phase separations in the trajectories of the particles. We also find second-order DPTs in one-dimensional Brownian motion, corresponding to localization transitions of their trajectories.

Poster 15

Yoshiaki Horiike

Nagoya University

Images with translational and rotational symmetry correspond to the Ising models

As suggested by Hopfield, the image retrieval and prototypical Ising models have a deep correspondence through the underlying interaction between the pixels or spins. A recent approach using principal component analysis (PCA) on states of the Ising model successfully discovers the order parameter, phase, and phase transition. To understand the theory behind its success, we performed the PCA on the two-dimensional Ising model and found that PCA is similar to Fourier expansion. Motivated by these results, we performed the PCA on image data and found similar results but, at the same time, qualitative differences in the explained variance of PCs. By adding image data with translational and rotational symmetry, the results of PCA on the Ising model and image data were agreed upon. We discuss the theoretical background behind this symmetry and PCA.

Poster 16

Takeru Yokota
RIKEN

Physics-informed Neural Networks for Functional Differential Equations

In this poster, we will present the first learning scheme for functional differential equations (FDEs), which we have recently proposed [1]. FDEs play a fundamental role in physics, mathematics, and optimal control. However, the numerical analysis of FDEs has faced challenges due to its unrealistic computational costs and has been a long-standing problem over decades. Thus, numerical approximations of FDEs have been developed, but they often oversimplify the solutions. To tackle these two issues, we propose a hybrid approach combining physics-informed neural networks (PINNs) with the dimensional truncation of the functional space, which transforms FDEs into high-dimensional PDEs. We will present the outline of our formalism and show some numerical demonstration, including an application to the functional renormalization group [2], which is a non-perturbative formalism of the field theory.

[1] T. Miyagawa, T. Yokota, Physics-informed Neural Networks for Functional Differential Equations: Cylindrical Approximation and Its Convergence Guarantees, accepted to The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024)
[2] T. Yokota, Physics-informed neural networks for solving functional renormalization

group on a lattice, Phys. Rev. B 109, 214205 (2024)

Poster 17

Tota Yoshida
Osaka University

Replica analysis on the chaoticity of random multi-layer perceptrons

Chaos in random neural networks has attracted much interest over decades [1] including feedforward-type neural networks used in deep learning [2,3]. The ensemble of such random neural networks is important as it is the basis for developing statistical mechanics approaches to deep learning [4,5]. In this study, we analyzed the chaotic effect in feedforward-type neural networks of a rectangular shape using the replica method considering the zero temperature limit of a spin-glass model put on the same network geometry [5,6]. We also analyzed the effect of ferromagnetic bias in the input data and synaptic coupling to obtain some insights for the cases in which the neural network has some hidden embedded stable patterns. We found that the chaos effect disappears as the bias of the coupling becomes larger than a certain value.

[1] H. Sompolinsky, A. Crisanti, and H. J. Sommers, PRL 61, 259 (1988).
[2]B. Poole, S. Lahiri, M. Raghu, J. Sohl-Dickstein, and S. Ganguli. NIPS 29 (2016).
[3]Yuki Okazaki master's thesis(Osaka Univ., 2020)
[4] H. Yoshino, SciPost Phys. Core 2, 005 (2020)
[5] H. Yoshino, Phys. Rev. Research 5, 033068 (2023).
[6]E. Domany, W. Kinzel and R. Meir, J. Phys. A: Math. Gen. 22(1989) 2081.

Poster 18

Satoki Ishikawa
Institute for Science Tokyo

Local loss optimization in the infinite width: Stable parameterization of predictive coding networks and target propagation

Local learning, which trains a network through layer-wise local targets and losses, has been studied as an alternative principle to backpropagation (BP) in neural computation. However, its algorithms often become more complicated or require additional hyperparameters due to the locality, making it challenging to identify desirable settings where the algorithm progresses in a stable manner.
To gain theoretical and quantitative insight, we introduce a maximal update parameterization (muP) in the infinite-width limit for two representative designs of local targets: predictive coding (PC) and target propagation (TP). We verify that muP enables hyperparameter transfer across models of different widths.
Furthermore, our analysis of parameterization presents unique and intriguing properties not observed in conventional BP. Leveraged by the analysis of deep linear networks, we find that PC's gradients interpolate between BP and Gauss-Newton-like gradients, depending on the parameterization.
We demonstrate that, in some common settings, PC in the infinite-width limit leads to behavior more similar to BP.
For TP, even with the standard scale of the last layer different classical muP, it prefers feature learning rather than the kernel regime.

Poster 19

Kyosuke Adachi
RIKEN iTHEMS

Prediction of coexistence in condensation of disordered protein sequences

Different protein condensates, such as nucleolus and stress granules, play functional roles in cells. Phase separation caused by protein interactions has been proposed as the underlying mechanism for condensate formation. Though recent numerical studies based on coarse-grained polymer models of disordered proteins have facilitated the investigation of condensates, the theoretical prediction of the properties of condensates from the amino-acid sequence remains incomplete. In this study, we propose a computationally efficient method to estimate the interaction between disordered protein sequences, leveraging statistical and polymer physics as guiding principles. We apply this approach to predict the physical properties, such as the critical temperature for phase separation. We further extend the prediction method to multi-species systems, predicting whether distinct species are mixed or demixed in phase separation, as well as what sequence feature is important for demixing.

Poster 20

Kosei Nozaki
Nagoya University

Analysis of Attention Weights in AlphaFold 2

AlphaFold 2 has achieved remarkable accuracy in protein structure prediction. However, its inner workings remain a black box. In this study, we analyze the attention weights in AlphaFold and identify unique patterns specific to protein language by comparing them with those in language models. Additionally, we provide insights into the sequence order in which AlphaFold interprets amino acid sequences, shedding light on its approach to protein folding.

Poster 21

Matthew Smart
Flatiron Institute

In-context denoising with one-layer transformers and connections to associative memory retrieval

Recently, Ramsauer et al. introduced a dense associative memory (DAM) network, also known as a modern Hopfield network, with continuous states and an update rule that is reminiscent of the attention mechanism used in transformers. While this potential connection has stimulated the statistical physics community to further explore DAMs, the practical impact on conventional transformers including large language models is muted since associative memory networks are not typically competitors of transformers for the same tasks. Towards developing this emerging connection, we consider a transformer-friendly task: in-context denoising. We show that certain restricted denoising problems can be solved even by a one-layer transformer. We then show that one-step updates of corresponding associative memory networks, starting with noisy inputs as the initial state, provide qualitatively similar denoising solutions

Poster 22

Yukihiro Tomita
Osaka University

Glasses in 1+∞ dimensional space: a microscopic theory for spatially heterogeneous glasses

Recently exact replica liquid theory for structural glasses was established in $d=\infty$ dimensional limit [1]. Here we extend the theory into $1+\infty$-dimensional space to analyze spatially inhomogeneous glasses [2]. From a broader perspective, the problem can be related to deep neural networks subjected to random training data [3]. We first analyze hard spheres confined in a cavity bounded by frozen walls. We obtained correlation lengths that diverge at the dynamical transition point (exponent is -1/4) and at the Kauzmann transition point (exponent is -1). The exponents agree with previous works including the inhomogeneous MCT[4].

We further extended the theory to the state following setup. We analyzed the confined hard spheres under uniformly applied shear stress. In the linear regime, the shear modulus exhibits a spatial profile that reflects the length scale diverging at the dynamical transition point. In the nonlinear regime, we obtained an inhomogeneous shear strain profile with a large deformation at the center of the cavity.

[1]G. Parisi, P. Urbani, and F. Zamponi, "Theory of Simple Glasses: Exact Solutions in Infinite Dimensions" (Cambridge University Press, 2020).
[2] Y. Tomita and H. Yoshino, in preparation
[3] H. Yoshino, SciPost Phys. Core 2, 005 (2020).
[4] G. Biroli, J.-P. Bouchaud, K. Miyazaki and Reichman, D. R., Phys. Rev. Lett., 97 (2006) 195701.

Poster 23

Kanta Masuki
The University of Tokyo

Generative Diffusion Model with Inverse Renormalization Group Flows

Diffusion models represent a class of generative models that produce data by denoising a sample corrupted by white noise. Despite the success of diffusion models in computer vision, audio synthesis, and point cloud generation, so far they overlook inherent multiscale structures in data and have a slow generation process due to many iteration steps. In physics, the renormalization group offers a fundamental framework for linking different scales and giving an accurate coarse-grained model.

Here we introduce a renormalization group-based diffusion model that leverages multiscale nature of data distributions for realizing a fast and efficient data generation. In the spirit of renormalization group procedures, we define a flow equation that progressively erases data information from fine-scale details to coarse-grained structures. Through reversing the renormalization group flows, our model is able to generate high-quality samples in a coarse-to-fine manner. We validate the versatility of the model through applications to protein structure prediction and image generation. Our model outperforms conventional diffusion models in image generation and exhibits comparable performance in protein structure prediction. The proposed method removes the need for data-dependent tuning of hyperparameters, showing promise for systematically enhancing sample efficiency based on multiscale features

Poster 24

Ryota Nasu
Shizuoka University

Quantum error correction realized by the renormalization group in scalar field theories

We demonstrate that quantum error correction is realized by the renormalization group in scalar field theories. We construct q-level states by using coherent states in the IR region. By acting on them the inverse of the unitary operator U that describes the renormalization group flow of the ground state, we encode them into states in the UV region. We find the situations in which the Knill–Laflamme condition is satisfied for operators that create coherent states. We verify this to the first order in the perturbation theory. This result suggests a general relationship between the renormalization group and quantum error correction and should give insights into understanding the role played by them in the gauge/gravity correspondence.

Poster 25

Reo Otsuki
Department of Human and Engineered Environmental Studies, The University of Tokyo

Predicting stimulus response of neuronal population using parallel reservoir computing

Prediction of spatiotemporal neural activity holds significant potential for early seizure detection and for advancing our understanding of cognitive brain processes. Parallel reservoir computing, a technique in which spatiotemporal data are divided into regions, with each reservoir responsible for predicting its respective region, has been proposed as an efficient and accurate method for handling large-scale spatiotemporal data. However, no studies have examined the capability of parallel reservoir computing to predict the neural responses of large-scale neural populations. In this study, we extend parallel reservoir computing to the prediction of local field potentials in spiking neural networks when external input is applied.

Poster 26

Takashi Takahashi
The University of Tokyo

A replica analysis of under-bagging

Under-bagging (UB), which combines under-sampling and bagging, is a popular ensemble learning method for training classifiers on imbalanced data. In this study, we derive a sharp asymptotics of UB and use it to compare with several other popular methods for learning

from imbalanced data, in the scenario where a linear classifier is trained from a two-component mixture data. The methods compared include the under-sampling (US) method, which trains a model using a single realization of the under-sampled data, and the simple weighting (SW) method, which trains a model with a weighted loss on the entire data. It is shown that the bagging procedure effectively reduces the variance of the prediction and improves the generalization performance, but the performance of UB is almost the same as the SW with the optimal weighting coefficients, indicating that the implicit regularization of bagging is similar to the combination of reweighting and $¥ell_2$ regularization.
The content of this presentation is based on the following paper:
T. Takahashi: Transactions on Machine Learning Research (2024).

Poster 27

Kaichi Miyashita
Keio University

Task Difficulty Evaluation Using a Restricted Boltzmann Machine

In recent years, AI has advanced rapidly, but the inner workings of trained models often remain a black box, sometimes causing unexpected issues. For AI to be effectively implemented in society, it is crucial to understand what these models have learned and how. Deep neural networks, while widely used, are complex and difficult to analyze. To address this, the study focuses on the simpler structure of the Restricted Boltzmann Machine (RBM) to analyze the internal mechanics of a trained model. The study involved training an RBM with the MNIST (Mixed National Institute of Standards and Technology) handwritten digits dataset and comparing task difficulty through weight analysis. RBM consists of visible and hidden layers connected by weighted edges. The difficulty perceived by the RBM was defined by the information content of these weights. Singular value decomposition was performed on the weight matrix, and the singular value distribution was detected. By evaluating the bias of this distribution using the Gini coefficient, the difficulty of the learning task was quantitatively assessed. Additionally, t-SNE (t-distributed Stochastic Neighbor Embedding) was used to visualize the features encoded by the hidden layer, offering another approach to evaluate the learning difficulty of RBM tasks.

Poster 28

Cancelled

Poster 29

Kota Mitsumoto

Graduate Schoole of Arts and Sciences, The University of Tokyo

Supercooled liquid dynamics in a frustrated spin model

Frustration suppresses an ordinary long-range order, leading to various intriguing states such as glasses and supercooled liquids. Supercooled liquids are often observed in off-lattice systems, while few examples of lattice models are known. We have found a frustrated spin model that behaves as a good glass former [Kota Mitsumoto, Chisa Hotta, and Hajime Yoshino Phys. Rev. Research. 4, 033157 (2022)]. The model includes no quenched disorder and has a unique ground state. However, the system cannot reach the ground state with even a slow cooling rate. At low temperatures, the system shows glassy dynamics. The auto-correlation function exhibits a two-step relaxation and the functional form is stretched exponential. The typical relaxation time obeys Arhenius's low, i.e., a strong glass. We have found that the glassy dynamics is governed by a characteristic defect whose energy is very close to the local ground-state energy. Our model provides a key concept of supercooling and glass transition without quenched disorder in lattice systems.

Poster 30

Tomoei Takahashi

The University of Tokyo

Alpha helices are more evolutionarily robust to environmental perturbations than beta sheets: Statistical mechanics of protein evolution

According to Darwinian evolution, natural selection and genetic mutations are the primary drivers of biological evolution. However, the concept of “robustness of the phenotype to environmental perturbations across successive generations," which seems crucial from the perspective of natural selection, has not been formalized or analyzed. In this study, through Bayesian learning and statistical mechanics we formalize the stability of the free energy in the space of amino acid sequences that can design particular protein structure against perturbations of the chemical potential of water surrounding a protein as such robustness. Consequently, in a two-dimensional square lattice protein model composed of 36 residues, we found that as we increase the stability of the free energy against perturbations in environmental conditions, the structural space shows a steep step-like reduction. Furthermore, lattice protein structures with higher stability against perturbations in environmental conditions tend to have a higher proportion of α-helices and a lower proportion of β-sheets.

Poster 31

Toshiaki Omori
Kobe University

Data-driven State Space Modeling Approach for Simultaneously Estimating and Controlling Nonlinear Dynamical Systems

Estimating and controlling dynamical systems from observable time-series data is crucial for understanding and manipulating nonlinear dynamics. We propose a probabilistic framework that simultaneously estimates and controls nonlinear dynamics under partial observation processes. Our method utilizes the sequential Monte Carlo approach, serving not only as a state estimator and a prior estimator for system dynamics but also as a controller. To demonstrate the effectiveness of our framework, we apply it to two distinct types of nonlinear dynamical systems: a chaotic system and a neuronal system. The results show that our framework can effectively estimate and control complex nonlinear dynamical systems under partial observations.

Poster 32

Chikara Furusawa

Universal Biology Institute, The University of Tokyo

Analysis of evolutionary constraints using bacterial experimental evolution

Biological systems change their state in order to adapt and evolve to changing environmental conditions. However, despite the recognition of the importance of clarifying the adaptive and evolutionary capabilities of organisms, research on the evolvability and plasticity of organisms remains at a qualitative level. To clarify how evolutionary processes are constrained in high-dimensional phenotypic and genotypic space, we performed laboratory evolution under various (>100) stress environments and analyzed phenotypic and genomic sequence changes. These comprehensive analyses revealed that changes in expression are restricted to low-dimensional dynamics, while diverse genomic changes contribute to similar phenotypic changes. To further analyze the nature of the evolutionary constraints, we performed computer simulations of adaptive evolution using a simple cellular model. Again, we found that changes in cell state in adaptation and evolution are generally restricted to low-dimensional dynamics. Based on these results, we would like to discuss the nature of phenotypic plasticity and constraints in bacterial evolution and possible strategies to predict and control evolutionary dynamics.

Poster 33

Keiichi Tamai

The University of Tokyo

Universal Scaling Laws of Absorbing Phase Transitions in Artificial Deep Neural Networks

Critical phenomena at second-order phase transitions have been hypothesized to be the key to the intelligence of living systems. The rapid progress in deep learning applications naturally motivates us to ask whether the hypothesis is also the case for artificial deep neural networks: if so, we may expect unified understanding on the physical principles behind the intelligence. In this presentation, we demonstrate that classical artificial deep neural networks around the phase boundary of the signal propagation dynamics (also known as the edge of chaos) exhibit universal scaling laws of absorbing phase transitions in non-equilibrium statistical mechanics. Numerical evidence suggests that the multilayer perceptrons and the convolutional neural networks respectively belong to the mean-field and the directed percolation universality class. Also, the finite-size scaling can be successfully applied, which may be useful for deeper insight into the depth-width trade-off in deep learning. Furthermore, our analysis on the training dynamics under gradient descent indicates that setting the hyperparameters to the phase boundary is necessary but not sufficient for achieving good generalization with the deep networks. Our work suggests that the notion of criticality is useful for analyzing the behavior of artificial deep neural networks and provides some hints to understand a subtle relationship between criticality and the intelligence.

Poster 34

Jun Takahashi
ISSP, The University of Tokyo

NP-hardness of Curing the Negative-Sign Problem with Clifford Circuits

The negative-sign problem is the difficulty one faces when trying to use Monte Carlo sampling to study thermal states of quantum systems. It not only is a practical obstacle for studying quantum systems but also provides a theoretical framework for considering an interesting middle ground of "stoquastic" systems where a quantum Hamiltonian is not entirely classical but nevertheless admits Monte Carlo sampling without a sign problem. Since the Monte Carlo sign problem is a basis dependent property, it has remained rather mysterious when a quantum system can and cannot be made sign-problem free i.e. stoquastic with some basis rotation. In this poster I will consider curing the sign problem by using Clifford circuits, which is a framework that is "barely classically simulable" i.e., adding just one other type of gate makes the circuit universal in terms of quantum computation. We show that this class can cure the sign problems of many interesting classes including free-fermionic systems and the J1-J2 chain which is traditionally considered to be frustrated. We finally prove computational NP-hardness of the problem of curing the sign-problem using Clifford circuits, which is the first result for a global unitary transformation.

Poster 35

Koki Okajima
The University of Tokyo

Asymptotic Dynamics of Alternating Minimization for Bilinear Regression

This study investigates the asymptotic dynamics of alternating minimization applied to optimize a bilinear non-convex function with normally distributed covariates. This is achieved by employing the replica method to a multi-temperature glassy system which unfolds the algorithm's time evolution. Our results show that the dynamics can be described effectively by a two-dimensional discrete stochastic process, where each step depends on all previous time steps, revealing the structure of the memory dependence in the evolution of alternating minimization. The theoretical framework developed in this work can be applied to the analysis of various iterative algorithms, extending beyond the scope of alternating minimization.