QI 9: Quantum Machine Learning and Classical Simulability

Time: Tuesday 9:30–13:15 Location: HFT-FT 101

Invited Talk QI 9.1 Tue 9:30 HFT-FT 101

Does provable absence of barren plateaus imply classical simulability? Or, why we might need to rethink variational quantum computing — •Zoe Holmes — EPFL, Lausanne, Switzerland

A large amount of effort has recently been put into understanding the barren plateau phenomenon. Here we face the increasingly loud elephant in the room and ask a question that has been hinted at by many but not explicitly addressed: Can the structure that allows one to avoid barren plateaus also be leveraged to efficiently simulate the loss classically? We present strong evidence that commonly used models with provable absence of barren plateaus are also in a sense classically simulable, provided that one can collect some classical data from quantum devices during an initial data acquisition phase. This follows from the observation that barren plateaus result from a curse of dimensionality, and that current approaches for solving them end up encoding the problem into some small, classically simulable, subspaces. This sheds serious doubt on the non-classicality of the information processing capabilities of parametrized quantum circuits for barren plateau-free landscapes and on the possibility of superpolynomial advantages from running them on quantum hardware. We end by discussing caveats in our arguments, the role of smart initializations, and by highlighting new opportunities that our perspective raises.

QI 9.2 Tue 10:00 HFT-FT 101

Can a neural network fake a Boson Sampler? — •Martina Jung[1], Martin Gärttner[1], and Moritz Reh[1,2] — [1]Friedrich-Schiller-Universität, Jena, Deutschland — [2]Universität Heidelberg, Heidelberg, Deutschland

Originally defined to demonstrate quantum supremacy, Boson Sampling and its simulation have become an own field of research. The simulation of a Boson Sampler is an - per-construction - classically intractable problem due to the computational complexity of its distribution. A statistics-based approach to learn the probability distribution of a sampling process faces issues like sparse data and highly correlated output configurations. These problems are reminiscent of natural language processing (NLP) tasks where a neural network is trained to respond to a query. Indeed, NLP models like a recurrent neural network (RNN) decompose the task by learning to sequentially predict the next word based on the preceding sequence of words. Transferring this concept to the bosonic Fock space, we train a RNN to simulate a Boson Sampler by predicting the conditional probabilities related to input-output configurations. The model's ability to extrapolate is tested on input sequences of lengths beyond the ones seen during the training.

QI 9.3 Tue 10:15 HFT-FT 101

Parametrized Quantum Circuits and their approximation capacities in the context of quantum machine learning — Alberto Manzano[1], •David Dechant[2,3], Jordi Tura[2,3], and Vedran Dunjko[2,3,4] — [1]Department of Mathematics and CITIC, Universidade da Coruña, Campus de Elviña s/n, A Coruña, Spain — [2]Applied Quantum Algorithms Leiden, The Netherlands — [3]Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands — [4]LIACS, Universiteit Leiden, P.O. Box 9512, 2300 RA Leiden, Netherlands

Parametrized quantum circuits (PQC) are used in recent approaches to quantum machine learning to learn various types of data, with an underlying expectation that if the PQC is made sufficiently deep, and the data plentiful, the generalization error will vanish, and the model will capture the essential features of the distribution. While there exist results proving the approximability of square-integrable functions by PQCs under the L^2 distance, the approximation for other function spaces and under other distances has been less explored. In this work we show that PQCs can approximate the space of continuous functions, p-integrable functions and the H^k Sobolev spaces under specific distances. Moreover, we develop generalization bounds that connect different function spaces and distances. These results provide a theoretical basis for different applications of PQCs, for example for solving differential equations. Furthermore, they provide us with new insight on how to design PQCs and loss functions which better suit the specific needs of the users.

QI 9.4 Tue 10:30 HFT-FT 101

Unifying (Quantum) Statistical and Parametrized (Quantum) Algorithms — •Alexander Nietner — FU-Berlin

Kearns SQ oracle lends a unifying perspective for most classical machine learning algorithms. This no longer holds in case of quantum learning and with respect to the SQ or QSQ oracle. In this work we explore the problem of learning from an evaluation oracle, which provides an estimate of function values. We introduce an intuitive framework that yields unconditional lower bounds for learning from evaluation queries and characterizes the query complexity for learning linear function classes. The framework is directly applicable to the QSQ setting and virtually all algorithms based on loss function optimization.

We first apply this formalism to the QSQ setting studying the learnability of unitary and Clifford quantum circuit states at different depth regimes and prove exponential separations of learning stabilizer states from QSQs versus from quantum copy access.

Our second application is to analyze popular QML settings and to develop an intuitive picture that goes beyond that of barren plateaus. This enables us to show how the implications of a barren plateau depend on the particular setting, which gives new and valuable insights into variational algorithms.

QI 9.5 Tue 10:45 HFT-FT 101

Information-theoretic generalization bounds for learning from quantum data — •Matthias C. Caro[1,2], Tom Gur[3], Cambyse Rouzé[4,5], Daniel Stilck França[6], and Sathyawageeswar Subramanian[3,7] — [1]Dahlem Center for Complex Quantum Systems, FU Berlin — [2]IQIM, Caltech — [3]Department of Computer Science and Technology, University of Cambridge — [4]Inria, Télécom Paris - LTCI, Institut Polytechnique de Paris, Palaiseau, France — [5]Zentrum Mathematik, TU München — [6]Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France — [7]Department of Computer Science, University of Warwick

Learning tasks play an increasingly prominent role in quantum information and computation. However, the many directions of quantum learning theory have so far evolved separately. We propose a general mathematical formalism for describing quantum learning by training on classical-quantum data and then testing how well the learned hypothesis generalizes to new data. In this framework, we prove bounds on the expected generalization error of a quantum learner in terms of classical and quantum information-theoretic quantities measuring how strongly the learner's hypothesis depends on the specific data seen during training. To achieve this, we use tools from quantum optimal transport and quantum concentration inequalities. Our framework encompasses and gives intuitively accessible generalization bounds for a variety of quantum learning scenarios. Thereby, our work lays a foundation for a unifying quantum information-theoretic perspective on quantum learning.

QI 9.6 Tue 11:00 HFT-FT 101

Efficient classical surrogate simulation of quantum circuits — •Manuel S. Rudolph[1,5], Enrico Fontana[2,3,4], Ross Duncan[3], Ivan Rungger[4], Zoë Holmes[1], Lukasz Cincio[5], and Cristina Cîrstoiu[3] — [1]EPFL, Lausanne, Schweiz — [2]University of Strathclyde, Glasgow, UK — [3]Quantinuum, Cambridge, UK — [4]National Physical Laboratory, Teddington, UK — [5]Los Alamos National Lab, Los Alamos, USA

Performant classical simulation of quantum systems is crucial for benchmarking quantum algorithms and verifying potential quantum advantages. Here, we provide two results. First, we prove that there exists a polynomial-time algorithm for simulating quantum circuits affected by constant local Pauli noise with bounded average error as the number of qubits or circuit depth increases. This highlights that, on average, there cannot be an exponential quantum-classical separation in observable estimation tasks when the quantum hardware is affected by such noise. Second, we turn our Theorems into a full-fledged high-performance simulation algorithm called "LOWESA" for noisy and noise-free quantum circuits. LOWESA can be understood as a classical surrogate for expectation landscapes with fast re-evaluation at different circuit parameters. We show that we can scale our simulations to the 127-qubit examples presented in Nature 618, 500-505

(2023), where we produce near-exact expectation values and highlight the strengths of LOWESA compared to other established simulation methods.

15 min. break

QI 9.7 Tue 11:30 HFT-FT 101

Exponential concentration in quantum kernel methods — •Supanut Thanasilp[1], Samson Wang[2], Marco Cerezo[3], and Zoe Holmes[1] — [1]EPFL, Lausanne, Switzerland — [2]Imperial college London, London, UK — [3]Los Alamos National Laboratory, New Mexico, US

Kernel methods in Quantum Machine Learning have recently gained significant attention as a candidate for achieving a quantum advantage. Among attractive properties, when training a kernel-based model one is guaranteed to find the optimal models parameters due to the convexity of the landscape. However, this is based on the assumption that the kernel can be efficiently obtained from quantum hardware. In this work we study the performance of quantum kernel models from the perspective of the resources needed to accurately estimate kernel values. We show that, under certain conditions, values of quantum kernels over different input data can be exponentially concentrated (in the number of qubits) towards some fixed value. Thus on training with a polynomial number of measurements, one ends up with a trivial model where the predictions on unseen inputs are independent of the training data. We identify four sources that can lead to concentration including expressivity of data embedding, global measurements, entanglement and noise. For each source, an associated concentration bound of quantum kernels is analytically derived. Lastly, we show that when dealing with classical data, training a parametrized data embedding with a kernel alignment method is also susceptible to exponential concentration.

QI 9.8 Tue 11:45 HFT-FT 101

On the expressivity of embedding quantum kernels — •Elies Gil-Fuster[1,2], Jens Eisert[1,2,3], and Vedran Dunjko[4,5] — [1]Dahlem Center for Complex Quantum Systems, Freie Universitat Berlin — [2]Fraunhofer Heinrich Hertz Institute, Berlin — [3]Helmholtz-Zentrum Berlin fur Materialien und Energie — [4]Applied Quantum Algorithms, Universiteit Leiden, Netherlands — [5]LIACS, Universiteit Leiden, Netherlands

One of the most natural connections between quantum and classical machine learning has been established in the context of kernel methods. Quantum kernels are typically evaluated by explicitly constructing quantum feature states and then taking their inner product, here called embedding quantum kernels. Since classical kernels are usually evaluated without using the feature vectors explicitly, we wonder how expressive embedding quantum kernels are. In this work, we raise the question: can all quantum kernels be expressed as the inner product of quantum feature states? Our first result is positive: for any kernel function there always exists a corresponding quantum feature map and an embedding quantum kernel. In a second part, we formalize the question of universality of efficient embedding quantum kernels. We show that efficient embedding quantum kernels are universal within a broad class of shift invariant kernels. We then extend this result to a new class of so-called composition kernels, which we show also contains projected quantum kernels introduced in recent works. We finally identify the directions towards new, more exotic, and unexplored quantum kernel families.

QI 9.9 Tue 12:00 HFT-FT 101

A Multi-Excitation Projective Simulation Learning Agent — •Philip LeMaitre, Marius Krumm, and Hans Briegel — Universität Innsbruck, Institut für Theoretische Physik, Innsbruck, Austria

The rapid integration of artificial intelligence (AI) into daily life, driven by advanced large language models such as ChatGPT, highlights a critical question in AI research: how can we comprehend an AI's decision-making process that leads to specific outcomes? To address this question, the field of explainable AI emerges as vital, with the projective simulation reinforcement learning framework being a notable component. An extension of this framework is considered to enable the AI agent to process multiple variables concurrently, enhancing its ability to discern complex correlations within its environment. Additionally, an inductive bias inspired from quantum many-body expansions of the Hamiltonian is introduced. This bias focuses on smaller clusters of memory states during decision-making, balancing the increased complexity inherent in the extended model. The enhanced framework is then applied to two distinct learning scenarios: a simple defence game featuring deceptive strategies by the attacker, and a more complex scenario mimicking computer diagnostics and maintenance tasks. In both contexts, the agent successfully learns optimal policies by leveraging higher-order correlations. Furthermore, a preliminary overview of the quantum variant of the model is provided, offering a more realistic model for future explorations in explainable quantum AI.

QI 9.10 Tue 12:15 HFT-FT 101

On the average-case complexity of learning output distributions of quantum circuits — Alexander Nietner[1], Marios Ioannou[1], Ryan Sweke[1,3], Richard Kueng[2], Jens Eisert[1], •Marcel Hinsche[1], and Jonas Haferkamp[1,4] — [1]FU Berlin — [2]JKU Linz — [3]IBM Quantum — [4]Harvard University

In this work, we show that learning the output distributions of brickwork random quantum circuits is average-case hard in the statistical query model, which models most practical algorithms. Our main results are:

- At super logarithmic circuit depth $d = \omega(\log(n))$, any learning algorithm requires super polynomially many queries to achieve a constant probability of success over the randomly drawn instance.
- There exists a $d = O(n)$, such that any learning algorithm requires $\Omega(2^n)$ queries to achieve a $\Omega(2^{-n})$ probability of success over the randomly drawn instance.
- At infinite circuit depth $d \to \infty$, any learning algorithm requires $2^{2^{\Omega(n)}}$ many queries to achieve a $2^{-2^{O(n)}}$ probability of success over the randomly drawn instance.

Moreover, we confirm a variant of a conjecture by Aaronson and Chen and show that the output distribution of a brickwork random quantum circuit is constantly far from any fixed distribution in total variation distance with probability $1 - O(2^{-n})$.

QI 9.11 Tue 12:30 HFT-FT 101

Understanding quantum machine learning also requires rethinking generalization — •Elies Gil-Fuster[1,2], Jens Eisert[1,2,3], and Carlos Bravo-Prieto[1] — [1]Dahlem Center for Complex Quantum Systems, Freie Universitat Berlin — [2]Fraunhofer Heinrich Hertz Institute, Berlin — [3]Helmholtz-Zentrum Berlin fur Materialien und Energie

Quantum machine learning models have shown successful generalization performance even when trained with few data. In this work, through systematic randomization experiments, we show that traditional approaches to understanding generalization fail to explain the behavior of such quantum models. Our experiments reveal that state-of-the-art quantum neural networks accurately fit random states and random labeling of training data. This ability to memorize random data defies current notions of small generalization error, problematizing approaches that build on complexity measures such as the VC dimension, the Rademacher complexity, and all their uniform relatives. We complement our empirical results with a theoretical construction showing that quantum neural networks can fit arbitrary labels to quantum states, hinting at their memorization ability. Our results do not preclude the possibility of good generalization with few training data but rather rule out any possible guarantees based only on the properties of the model family. These findings expose a fundamental challenge in the conventional understanding of generalization in quantum machine learning and highlight the need for a paradigm shift in the design of quantum models for machine learning tasks.

QI 9.12 Tue 12:45 HFT-FT 101

More efficient exchange-only quantum gates via reinforcement learning — •Violeta N. Ivanova-Rohling[1,2,3], Niklas Rohling[1], and Guido Burkard[1] — [1]Department of Physics, University of Konstanz — [2]Zukunftskolleg, University of Konstanz — [3]Department of Mathematical Foundations of Computer Sciences, IMI, Bulgarian Academy of Sciences

There has recently been rapid progress in the research of spin qubits [1], including the realization of exchange-only qubits [2,3]. Here, we use reinforcement learning to optimize the efficiency of exchange-based pulse sequences that encode the universal two-qubit gates CNOT and CZ with nearest-neighbor interaction for quantum dot arrangements in a chain and in a 2 by 3 grid. We improve on gate sequences currently known in the literature. Specifically, with our reinforcement learning framework, we manage to find a gate sequence encoding CNOT with a shorter total time than the Fong-Wandzura sequence [4] which is currently state of the art. Moreover, the flexibility of our approach makes it applicable for gate-sequence optimization for a variety of desired quantum gates and a variety of different connection topologies.

[1] Burkard, Ladd, Pan, Nichol, Petta, Rev. Mod. Phys. **95**, 025003 (2023)
[2] DiVincenzo, Bacon, Kempe, Burkard, Whaley, Nature **408**, 339 (2000)
[3] Weinstein et al., Nature **615**, 817 (2023)
[4] Fong, Wandzura, Quantum Info. Comput. **11**, 1003 (2011)

QI 9.13 Tue 13:00 HFT-FT 101

The Mean King's Problem as a learning task — •Niklas Rohling — Department of Physics, University of Konstanz

The Mean King's Problem [1-5] is an early example of an advantage due to the availability of additional quantum resources. This original version of the problem is a single-shot measurement where Alice has to determine correctly the outcome of a measurement which was performed previously by the king's men. The difficulty comes from the fact that the measurement basis used by the king's men is revealed only after Alice has completed her measurement. The striking result is that Alice can find the correct answer with certainty if she is allowed to entangle the state initially with an additional quantum system. Here, we formulate the Mean King's Problem as a learning task where several copies of the state after the king's men's measurement, sorted by their outcome, are available. We investigate how the number of copies required to determine the measurement outcome of the king's men within desired error bounds ε and success probability $1-\delta$ scales with system size when additional quantum resources are (or are not) allowed to be used. We compare to the exponential advantage of quantum-enhanced learning found recently for measurements in product bases [6].

[1] Vaidman, Aharonov, Albert, PRL **58**, 1385 (1987)
[2] Aharonov, Englert, Z. Naturforsch. **56a**, 16 (2001)
[3] Englert, Aharonov, Physics Letters A **84**, 1 (2001)
[4] Aravind, Z. Naturforsch. **58a**, 682 (2003)
[5] Durt, Int. J. Mod. Phys. B **20**, 1742 (2006)
[6] Huang et al., Science **376**, 1182 (2022)