Research agenda of Quantum Machine Learning Group (Zbigniew Puchała, ITAI PAS)

I. Motivation and the general goal

With increasing complexity and inter-connectivity in the modern world the ability to solve optimization problems becomes indispensable. However, these problems are intrinsically hard to resolve as they usually require searching over an *astronomically* large spaces of possible solutions [1]. The most promising idea to overcome these difficulties could rely on quantum computers [2], annealers such as the D-Wave 2000Q chip [3], in particular. In principle, such machines could solve variate of (hard) optimization problems (almost) "naturally" by finding low energy eigenstates [4, 5].

Machine learning [6] is term that encompasses a large number of computational techniques whose goal is to create algorithms able to learn from data. In recent years the huge success of machine learning transformed developed societies in an unprecedented way.

Quantum machine learning is a field of research and engineering that focuses of employing quantum information and computation for learning from data. Such intrinsic features of quantum mechanics as linearity, randomness, superposition and entanglement can be naturally employed to process data using quantum resources. Many classical machine learning techniques use linear algebra and probability theory as their foundations therefore there exists a natural mapping between mathematical theory of machine learning and two first features of quantum mechanics. Yet the second two can also form a basis for quantum data processing [7, 8]. Quantum machine learning can be divided into three distinct groups [8]

• quantum processing of classical data, • application of classical machine learning techniques for quantum algorithms development, • quantum processing of quantum data.

In the context of NISQ technology the first approach is the most promising for the technical applications.

In the last two years one can observe large investments in the companies building products using quantum machine learning as core element of their business. These companies usually focus either on providing quantum hardware and basic software or providing software focused in specific business needs of clients. Amongst hardware start-up companies one can point to Rigetti Computing formed in 2013 and aiming at building universal quantum computers. Recently Rigetti obtained additional funding of about 90M USD in 2017[1]. Companies providing services and software for quantum computers include 1QBit, who rised 45M USD in 2017[2].

Several commercial actors are involved in development of quantum hardware [9] and software. In the scope of this project the following companies are the most relevant IBM, Google, Rigetti Computing, Intel, and D-Wave. The first four of them focus on quantum computing based on the gate model while the last one builds quantum annealers. Quantum computers being fragile quantum analog devices are inherently noisy and difficult to control. While quantum error correction will allow for creation of noise-resistant logical qubits in terms of NISQ technologies one has to rely on noisy physical qubits [10]. Four main features describe the quality of quantum hardware: number of qubits, precision of operations acting on qubits, controllability of qubits, and admissible two qubit operations. Precision of quantum operations is the crucial feature of a quantum device, yet very difficult to measure since it depends on device calibration, temperature and other physical factors [11].

The D-Wave quantum annealer 2000Q is composed of 2048 flux qubits and has a limited controllability. By finding low energies of classical Ising models is this device is able to find approximate solutions of some combinatorial optimisation problems. Current generations of D-Wave quantum annealers is based on so called Chimera graph topology but upcoming technology, which should be available in 2019 will be based on Pegasus topology that provides better connectivity between qubits.

A. Translation of business problems to quantum annealing architecture

Quantum computers in both the gate model and in the adiabatic quantum computation model are becoming one of the applicable tools for solving problems in the areas of optimisation, simulation of quantum systems and machine learning. While currently existing quantum computers are probably not able to solve any useful problems future generations might be. Therefore there exists a need to construct new tools such as compilers, hybrid classical-quantum algorithms, and domain specific tools. In particular, because quantum annealers are the most advanced quantum computational systems existing now, there is a major need for creation of tools that will facilitate applications of adiabatic quantum computation model for solving real-life technical problems. **The main goal of this part of the project is to derive a method for compiling probabilistic graphical models to Ising models suitable for quantum annealing.**

B. New schemes for benchmarking and certification

A major obstacle for large-scale quantum technologies is the lack of practical reliable certification tools. The certification and validation of sources producing quantum states and measurement devises, which enables to perform computations is a necessary step of quantum technology, which must follow from a fundamental scientific research. The assertion of quality of the elements of quantum device is a basic element of correctness of operating quantum computer. When dealing with certification usually, one uses additional assumptions, for instance, that the certification of a given device often depends on the fact that other devices are properly calibrated. Some efforts were made in the case of certifications of experimental preparations or arbitrary m-mode pure Gaussian states, pure non-Gaussian states generated by linear-optical circuits with n-boson Fock-basis states as in-

[1] https://www.crunchbase.com/search/funding_rounds/field/organizations/funding_total/rigetti-computing
[2] https://1qbit.com/news/1qbit-raises-45m-series-b-round/

puts [12]. To obtain full information about a quantum device a full process tomography must be performed, but for large system this requires measurement of exponentially many observables. **The main goal of this part of a project is to devise a reliable certification technique, such that the number of measurements scales polynomially with the size of the system and with a high probability the correct answer is obtained for a fixed number of tries.**

II. Research objectives

A. Translation of business problems to quantum annealing architecture

1. Task 1: Compiling probabilistic graphical models to Ising models suitable for quantum annealing

One of a very general language that allows for description of wide range of problems is provided by the probabilistic graphical models [13]. Graphical models represent relations between random variables in form graphs whose nodes are associated with these random variables and edges with relations between them. These models have applications in manufacturing, safety analysis, medicine, biology, finance, signal processing, machine learning and many others areas. Graphical models are usually divided into directed and non directed. The first class describes Bayesian relations between variables, the second describes conditional independence of variables. Unfortunately many problems related to graphical models such as: training and inference are NP-hard [14], and therefore approximate methods are employed to solve them.

While quantum computers are not likely to be able to solve NP-complete problems efficiently [15], it is generally believed that quantum annealing is a protocol that is expected to allow for finding quickly approximate solutions of some subclasses of NP-hard problems. Existing quantum annealers such as D-Wave machines are difficult to use due to their limitations imposed by their physical construction. Therefore it is important to be able to translate high-level problems originating from applications into a form that can be executed on quantum annealers. This form of translation should be performed by compilers that would take into account:

• structure of the high-level problem, • amount and structure of available computational classical and quantum computational resources, • quality of the solutions provided by classical and quantum computers.

The compilers mentioned above should allow to use mixed quantum-classical resources in order to perform three main tasks:

• inference, • learning graphical model structure learning from data, and • parameter learning.

Current research on implementation of graphical models focuses on particular models and do not encapsulates the problem in its entirety. The goal of the proposed research is to close this gap using a high-level approach. This goal will be achieved by creating a stack of methods that will facilitate translation of high-level problems encoded in the form of a graphical model into a series of classical and quantum operations that can be executed on classical and quantum computational resources.

Software stack will consists of the following components:

- high level domain specific language allowing for description of graphical models restricted to discrete random variables;
- compiler consisting of lexer, parser, decomposer and optimiser;
- runtime environment able to dispatch computation to classical and quantum computers. The core component of this stack is decomposer and optimiser. Decomposer will perform the following steps:
- if graphical model is defined as a Bayesian Network (BN) convert it to Markov Random Field (MRF),
- transform MRF into Higher Order Ising Model (HOIM),
- transform HOIM into Quadratic or Cubic Ising Model (QCIM),

The role of the optimiser is to

- sparsify QCIM by removing low value edges of the Ising graph,
- find the sets of variables that can be computed classically and therefore treated as constants in the quantum annealing step, and thus forming reduced QIM,
- find the minor embedding of reduce QIM into given hardware-dependent connection topology.

While some research was already done in the area discussed above, extending and compiling already existing results is necessary to provide good quality and useful business tool.

There are three mile stones in this subject to reach:

- Development of Domain Specific Language allowing for description of probabilistic graphical models
- Development of compiler, decomposer and optimizer;
- Creation of Runtime Environment able to dispatch on classical and quantum computers.

2. Task 2: Emulation of quantum annealers and its applications

Quantum annelers can find low energy eigenstates of Hamiltonians by exploting the laws of quantum mechanics but, the low energy spectrum of most *local* and gapped Hamiltonians is known to be only *slightly* entangled [16]. Therefore, it should be within the reach of classical computers. For classical algorithms to find low energy states *effectively* it is often necessary for them be confined to a small region of the entire Hilbert space [17]. This is believed to be the case for sparse graphs. Motivated by the topology of near-term quantum annealers, we focus on a class of quasi-two-dimensional problems encoded via the Ising Hamiltonian [18],

$$H(\boldsymbol{s}) = -\sum_{\langle i,j\rangle\in\mathscr{E}} J_{ij}s_is_j - \sum_{i\in\mathscr{V}} h_is_i, \qquad (1)$$

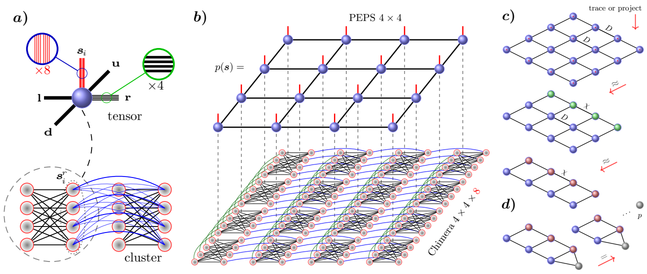

FIG. 1. **Tensor network formalism to solve classical optimization problems on chimera-like graphs.** ***a***) A mapping explaining how PEPS (Projected Entangled Pair States) tensors are assigned to "clusters" of the chimera-like graph. Each tensor is comprised of four virtual and one physical bond of sizes $D = D_0^{\min(m,n)}$ and d^8, respectively. Here, $D_0 = d = 2$ while m is the number of qubits in one cluster interacting with n of those in the neighboring cluster. For the chimera graph $n = m = 4$. ***b***) The resulting PEPS tensor network allowing one to represent the probability distribution $p(\boldsymbol{s})$ for the entire graph. ***c***) Contraction of the network to calculate the marginal probability $p = p(\boldsymbol{s}_A)$ for a given partial configuration $\boldsymbol{s}_A$. First, physical degrees of freedom are projected on $\boldsymbol{s}_A$, and the remaining ones are traced out. Next, an approximate scheme is invoked to collapse the network in a top-bottom fashion until only two rows remain. Finally, the remaining tensors can be exactly contracted as shown in ***d***) to retrieve the probability p.

where the objective is to find a particular configuration $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)$ of variables $s_i = \pm 1$ (or $\uparrow\downarrow$) that minimizes the cost function (energy) H. Input parameters J_{ij} and h_i can be defined on a *generic* chimera-like graph $\mathscr{G} = (\mathscr{E}, \mathscr{V})$ (or beyond) rendering this problem NP-hard [19].

Our idea is to represent the probability distribution $p(\boldsymbol{s}) \sim \exp[-\beta H(\boldsymbol{s})]$ as the Projected Entangled Pair States (PEPS) tensor network [20] as depicted in Fig. 1**a**)- 1**b**). Essentially, this allows one to efficiently calculate the probabilities for any configuration, including the marginal ones. This is executed by approximately contracting the PEPS tensor network as shown in Fig. 1**c**)- 1**d**). As a result, variaty of algorithms based on well known optimization strategies (*e.g.* branch and bound [21]) can be applied to reveal the structure of low energy spectrum of $H(\boldsymbol{s})$. Similarly, this makes sampling from the distribution $p(\boldsymbol{s})$ possible – feature which is crucial for developing new generation of hybrid classical-quantum algorithms [22].

It is needless to say that approaches based on PEPS tensor networks have their limitations [23]. However, when apply to solve classical optimization problems they could potentially outperform current state of the art methods.

There are three mile stones in this subject we would like to reach:

1) Efficient implementation of the outlined heuristic algorithm for the current and near-future D-Wave architectures. That includes chimera graphs (DW2X, 2000Q) and pegasus topology.

2) Devising an error correcting algorithm based on machine learning techniques. Such software should remove (or at least minimize) the noise in D-Wave quantum annealers and improve the quality of the generated solutions.

3) Implementing quantum phase transition experiments (e.g. Kibble-Zurek theory) on the resulting "noise-free" D-Wave machine as a benchmark for error-correcting mechanism.

Essentially, an efficient algorithm as describe in 1) should be able to generate high quality solutions for typical instances of the control parameters (J_{ij}, h_i) for large graphs, $N \sim 10^3$. Those solutions could be then use to train an appropriate neural network [3] to identify the noise (δJ_{ij}, δh_i) present in D-Wave quantum annealers. This knowledge, could be then apply to general cases to improve overall performance of D-Wave chips.

Simulating physics, quantum phase transition in particular, on D-Wave quantum annealers is hard. Many experimental results are far from theoretical predictions even for the simplest models [24]. Hopefully, with the errors correcting software one will be able to conduct such experiments and obtain meaningful results. Furthermore, having an efficient sampling algorithm for large systems opens new possibilities for hybrid classical-quantum algorithms (cf. Ref. [22]). The latter could in principle tackle real optimization challenges (e.g. traffic optimization and control). Tensor networks based techniques could provide profound insight into the structure of large low-dimension spin-glass problems, with ramifications both for noisy intermediate-scale quantum devices and machine learning.

There are quantum computing companies, including "startups" based in Poland [4], that would most likely be interested in commercializing results of the current research proposal.

B. New schemes for benchmarking and certification

1. Task 1: Discrimination of quantum devices

The first stage is the identification of families of quantum operations which benefit from adaptive discrimination. This task focuses on major extension of the results presented in our previous work [25]. First, we will formulate conditions for perfect parallel discrimination of quantum measurements based on the approach presented in [26]. Next, we will study the sequential approach. The starting point of this research is the work [27]. This will allow us to establish which types of measurements can be perfectly distinguished in parallel and which require additional processing after each usage of the black box. In the cases when perfect discrimination cannot be achieved, we will derive the probability of optimal discrimination utilizing the diamond norm. The next stage is a characterization of operations needed in the intermediate stages for the procedure.

2. Task 2: Learning of quantum operations

The first stage is an analysis of quantum learning procedures of von Neumann measurements. Learning of quantum

[3] This neural network will have to be identify first, which we anticipate should not be hard.

[4] Bohr technology – https://www.bohr.technology/

devices is strongly related to quantum simulations. In this scenario one is able to perform a finite number of allowed operations with a given device. This phase is called training. Here, the outcomes from device can be recorded in a quantum memory for further usage. Later, in the retrieving phase, using only information stored in memory one needs to simulate the action of the device on an unknown input state.

In the next stage we will consider quantum learning of generalized measurements. In this case the device in the black box performs an unknown quantum measurement. We store the result of the training phase in a quantum state and keep it in the quantum memory. In the the retrieving phase, it enters together with the unknown state into the retrieving channel that mimics the action of the transformation an a given state.

In the final stage we will consider the problem of quality of learning other classes of CPTP maps. It is important to note, that even for finite dimensional quantum systems there does not exist a finite number of learning phases for which the quantum learning works perfectly. Indeed, even if the training part of the strategy would encode full information about CPTP map into the finite dimensional state, the no programming theorem of Nielsen prevents us to retrieve the transformation perfectly.

3. Task 3: Certification of quantum devices

In order to create a certification procedure, one needs to design a scheme which produces an *Accept / Reject* outcome with a given quality threshold (e.g. fidelity) F and maximal failure probability p. The scheme must encapsulate the continuum of statistical alternatives – all possible inaccurate implementations of the device. For this reason, as a preliminary result we must have a good understanding of hypothesis testing with a simple – point alternative. Results concerning this subjects in the case of simple quantum devices – quantum measurements was recently obtained by Puchała et.al. [25, 28]. For a given quality function Q, we can define

Definition: (Quantum measurement certification). For a given quality threshold $F_T < 1$ and a maximal failure probability $p > 0$. Test which takes as input a classical description of measurement $\mathscr{M}$ and a black box $\mathscr{X}$ which is an implementation of $\mathscr{M}$ and outputs *Accept / Reject* labels is a certification test for $\mathscr{X}$ if: with probability at least $1-p$ it rejects every $\mathscr{X}$ for which $Q(\mathscr{M},\mathscr{X}) < F_T$ and accepts $\mathscr{M} = \mathscr{X}$.

We say that any $\mathscr{X}$ accepted by such a test is a certified implementation of $\mathscr{M}$. The main task is to design certification procedures, which satisfy the above definition. The procedures will be custom tailored to various classes of quantum devices • quantum states, • von Neumann measurements, • generalized measurements.

References

[1] Aaronson, S. *Quantum Computing Since Democritus* (Cambridge University Press, New York, NY, USA, 2013).

[2] Feynman, R. P. There's plenty of room at the bottom. *Caltech Eng. Sci.* **23**, 22 (1960).

[3] T. Lanting, *at al.*, Entanglement in a quantum annealing processor. *Phys. Rev. X* **4**, 021041 (2014).

[4] Harris, R. at al., Phase transitions in a programmable quantum spin glass simulator. *Science* **361**, 162 (2018).

[5] King, A. at al., Observation of topological phenomena in a programmable lattice of 1,800 qubits. *Nature* **560**, 456 (2018).

[6] Vapnik, V. *The nature of statistical learning theory* (Springer science & business media, 2013).

[7] Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N. & Lloyd, S. Quantum machine learning. *Nature* **549**, 195 (2017).

[8] Schuld, M. & Petruccione, F. Supervised learning with quantum computers. (2018).

[9] Wendin, G. Quantum information processing with superconducting circuits: a review. *Reports on Progress in Physics* **80**, 106001 (2017).

[10] Preskill, J. Quantum Computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).

[11] Michielsen, K., Nocon, M., Willsch, D., Jin, F., Lippert, T. & Raedt, H. D. Benchmarking gate-based quantum computers. *Computer Physics Communications* **220**, 44 (2017).

[12] Aolita, L., Gogolin, C., Kliesch, M. & Eisert, J. Reliable quantum certification of photonic state preparations. *Nature communications* **6**, 8498 (2015).

[13] Sucar, L. E. *Probabilistic Graphical Models: Principles and Applications* (Springer Publishing Company, Incorporated, 2015).

[14] Cooper, G. F. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence* **42**, 393 (1990).

[15] Lucas, A. Ising formulations of many np problems. *Frontiers in Physics* **2**, 5 (2014).

[16] Vidal, G. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.* **91**, 147902 (2003).

[17] Schollwöck, U. The density-matrix renormalization group. *Rev. Mod. Phys.* **77**, 259 (2005).

[18] Lucas, A. Ising formulations of many NP problems. *Frontiers in Physics* **2**, 5 (2014).

[19] Barahona, F. On the computational complexity of Ising spin glass models. *J. Phys. A: Math. Gen.* **15**, 3241 (1982).

[20] Orus, R. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Ann. Phys.* **349**, 117 (2014).

[21] Baccari, F., Gogolin, C. & Peter Wittek, A. A. Verification of quantum optimizers. Preprint at arXiv:1808.01275 (2018).

[22] Gardas, B., Rams, M. M. & Dziarmaga, J. Quantum artificial intelligence to simulate many body quantum systems. Preprint at arXiv:1805.05462v1 (2018).

[23] Haferkamp, J., Hangleiter, D., Eisert, J. & Gluza, M. Contracting projected entangled pair states is average-case hard. Preprint at arXiv:11810.00738 (2018).

[24] Gardas, B., Dziarmaga, J., Zurek, W. H. & Zwolak, M. Defects in Quantum Computers. *Sci. Rep.* **8**, 4539 (2018).

[25] Puchała, Z., Pawela, L., Krawiec, A. & Kukulski, R. Strategies for optimal single-shot discrimination of quantum measurements. *Phys. Rev. A* **98**, 042103 (2018).

[26] Duan, R., Guo, C., Li, C.-K. & Li, Y. Parallel distinguishability of quantum operations. in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2016) pp. 2259–2263 arXiv:1605.02294.

[27] Harrow, A. W., Hassidim, A., Leung, D. W. & Watrous, J. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Physical Review A* **81**, 032339 (2010).

[28] Puchała, Z., Pawela, Ł., Krawiec, A., Kukulski, R. & Oszmaniec, M. Multiple-shot and unambiguous discrimination of von neumann measurements. *arXiv preprint arXiv:1810.05122* (2018).