
VARIATIONAL QUANTUM CIRCUITS AND DEEP REINFORCEMENT LEARNING

Samuel Yen-Chi Chen
Department of Physics
National Taiwan University
Taipei, Taiwan
ycchen1989@gmail.com

Hsi-Sheng Goan
Department of Physics
National Taiwan University
Taipei, Taiwan
goan@phys.ntu.edu.tw

ABSTRACT

Recently, machine learning has prevailed in many academia and industrial applications. At the same time, quantum computing, once seen as not realizable, has been brought to markets by several tech giants. However, these machines are not fault-tolerant and can not execute very deep circuits. Therefore, it is urgent to design suitable algorithms and applications implementable on these machines. In this work, we demonstrate a novel approach which applies variational quantum circuits to deep reinforcement learning. With the proposed method, we can implement famous deep reinforcement learning algorithms such as experience replay and target network with variational quantum circuits. In this framework, with appropriate information encoding scheme, the possible quantum advantage is the number of circuit parameters with $poly(\log N)$ compared to $poly(N)$ in conventional neural network where N is the dimension of input vectors. Such an approach can be deployed on near-term noisy intermediate-scale quantum machines.

Keywords Variational Quantum Circuit · Quantum Circuit Learning · Deep Reinforcement Learning · Machine Learning

1 Introduction

Machine Learning has grown to a vast subject in the recent decade and has already prevailed in many academic fields and has successful commercial applications. It has been successful in computer vision [32] [36] [39], natural language processing [34], financial trading [12] [15] [28] [29] [33] and even the game of Go[30] which is previously considered formidable for machines. It also helped basics physics research such as quantum many-body physics [4] [8] [9] [16] [21] [24] [38] [37], phase-transitions [7], quantum control [1] [14] and quantum error correction [2] [22]. In the meantime, quantum computing, previously thought to be not realizable, has been brought to the market by several startups and large tech companies. However, large circuits still cannot be deployed on these machines due to the lack of quantum error correction. Therefore, it is urgent to develop algorithms and applications which can be implemented on such devices. In view of the continuously increasing demand for machine learning applications in the near future and also the quest for more advanced artificial intelligence, this research is targeted at reinforcement learning, which is referred to an agent interacting with the environment to gain the knowledge of surroundings and then derive the policy of decision-making accordingly [35]. This kind of learning is the form which can be regarded as *intelligence*. The present work focuses on the applications of variational quantum circuits in reinforcement learning, especially in *action-value* function approximation, which is currently based on deep neural networks [19] [20]. With appropriate circuit design and encoding scheme, it is possible to have quantum advantages with less parameters compared to conventional neural networks because of quantum entanglement. Such circuits will be implementable on currently commercially available quantum computers which are Noisy Intermediate-Scale Quantum machines(NISQ) and will show invaluable applications in both academia and industries.

2 Methods

2.1 Reinforcement Learning

Reinforcement learning is a machine learning paradigm which an *agent* interacting with an *environment* \mathcal{E} over a number of discrete time steps[35]. At each time step t , the agent receives a *state* or *observation* s_t and then chooses an *action* a_t from the set of possible actions \mathcal{A} according to its *policy* π , where the policy is a function mapping the state s_t to action a_t . After executing the action a_t , the agent receives the state of next time step s_{t+1} and a scalar *reward* r_t . The process continues until the agent reach the terminal state. An *episode* is defined as an agent starting from a randomly selected initial state and following the above-mentioned process until a terminal state.

The return $R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'}$ is the total discounted return from time step t , where γ is the discount factor lies in $(0, 1]$. The discount factor controls the importance of future rewards. For large γ , the agent weights future reward more and vice versa. The agent is to be trained to maximize the expected return from each state s_t . The *action-value function* or *Q-value function* $Q^\pi(s, a) = \mathbb{E}[R_t | s_t = s, a]$ is the expected return for selecting action a in state s following the policy π . The optimal action value function $Q^*(s, a) = \max_{\pi} Q^\pi(s, a)$ gives the maximal action-value across all possible policy. The value of state s under policy π is $V^\pi(s) = \mathbb{E}[R_t | s_t = s]$, it is the expected return by following policy π from state s .

2.1.1 Deep Q-Learning

The action-value function $Q(s, a)$ can be explicitly represented by a two dimensional table with total number of entries $s \times a$. However, when the state space or the action space is large or even continuous the tabular method is unfeasible. In this situation, the action-value function is represented with function approximators such as neural networks[19][20].

Using neural networks as function approximators to represent the Q -value function has been studied extensively [19][20] and has already gained great success in a variety of tasks such as playing video games. In this setting, the action-value function $Q(s, a; \theta)$ is parameterized by θ , the θ can be derived by a series of iterations from a variety of optimization methods adapted from other machine learning tasks. The simplest form is the *Q-learning*. In this method, the goal is to directly approximate the optimal action-value function $Q^*(s, a)$ with the loss function

$$L(\theta) = \mathbb{E}[(r_t + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta))^2] \quad (1)$$

The target is $r_t + \gamma \max_{a'} Q(s', a'; \theta^-)$ and the prediction is $Q(s, a; \theta)$, where s' is the state encountered after playing action a_t at state s .

Reinforcement learning is known to be hard to converge or even to diverge when a nonlinear function approximator such as a neural network is used to represent the action-value function [20]. There are several reasons to blame: The first is the correlation among the states or observations. It is obvious that each state or observation on a trajectory will be strongly correlated to each other. With this kind of training input, which is in general not iid sampling, the Q function may change dramatically, and therefore change the policy at relatively large scale. The second reason may lie in the fact that there is also large correlation between the action-value Q and the target values $r_t + \gamma \max_{a'} Q(s', a')$. In supervised learning, the targets are labels for the given inputs, and these labels will not change from time to time. However, under the setting of deep reinforcement learning, the target changes accordingly with the Q , causing the Q to be chasing an unstationary target.

The *deep Q-learning* presented in the work [20] addressed these issues through the following two mechanisms: *Experience replay*: To perform experience replay, we store each transition the agent encounters. The transition is stored as a tuple in the following form: (s_t, a_t, r_t, s_{t+1}) at each time step t . When updating Q -learning parameters, randomly samples a batch of experiences from the replay memory and then perform gradient descent with the following loss function: $L(\theta) = \mathbb{E}_{\text{replaybatch}}[(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta))^2]$. The key importance of experience replay is to lower the correlation of inputs for training the Q -function.

Target Network: θ^- is the paramater of the target network and these parameters are only updated every C steps. This setting helps stabilizing the Q -value function training since the target is relatively stationary compared to the action-value function.

2.2 Variational Quantum Circuits

Variational quantum circuit is a kind of quantum circuits with tunable parameters subject to iterative optimization. These parameters can be seen as the *weights* in artificial neural networks. Due to the lack of quantum error correction in the near-term quantum devices, these variational circuits can relatively be robust against noises since these deviations

Algorithm 1 Variational Quantum Deep Q Learning

```

Initialize replay memory  $\mathcal{D}$  to capacity  $N$ 
Initialize action-value function circuit  $Q$  with random parameters
for episode = 1,  $M$  do
    Initialise state  $s_1$ 
    for  $t = 1, T$  do
        With probability  $\epsilon$  select a random action  $a_t$ 
        otherwise select  $a_t = \max_a Q^*(s_t, a; \theta)$ 
        Execute action  $a_t$  in emulator and observe reward  $r_t$  and next state  $s_{t+1}$ 
        Store transition  $(s_t, a_t, r_t, s_{t+1})$  in  $\mathcal{D}$ 
        Sample random minibatch of transitions  $(s_j, a_j, r_j, s_{j+1})$  from  $\mathcal{D}$ 
        Set  $y_j = \begin{cases} r_j & \text{for terminal } s_{j+1} \\ r_j + \gamma \max_{a'} Q(s_{j+1}, a'; \theta) & \text{for non-terminal } s_{j+1} \end{cases}$ 
        Perform a gradient descent step on  $(y_j - Q(s_j, a_j; \theta))^2$ 
    end for
end for
    
```

are absorbed into the parameters during iterative optimization process. Previous work [13] [18] [26] have already demonstrate the capability of variational circuits representing function approximators, classifiers and even modeling quantum-many-body physics which are intractable on classical computers. For example, in the work [18], it has been shown that variational quantum circuits can be used to approximate analytical functions $f(x)$ and such kind of quantum circuits are in general hard to be simulated by classical computers, which implied more expressive power than classical function approximators like neural network.

2.3 Variational Quantum Deep Q Learning

In this work, we explore the possibility of expanding the capability of variational circuits to represent the action-value function. With the possible quantum advantages of having less parameters needed than conventional neural network, which is promising for modeling more complex environments. For *target network*, we construct two sets of circuit parameters while with the same circuit architecture. The target circuit is updated every 20 steps. For *experience replay*, the replay memory is set with the length of 80 and the training batch size is 5. The optimization in this work needs to calculate gradients of expectation values of quantum measurements, which can be done with the same circuit architecture and slightly different parameters [25]. We encode the state with *computational basis encoding*. In the FrozenLake environment, there are total 16 states, thus it needs 4 qubits to represent all states. For a general discussion about different encoding scheme, please refer to [27]

2.3.1 Computational Basis Encoding

For a general n -qubit state, it can be written as

$$|\psi\rangle = \sum_{(q_1, q_2, \dots, q_n) \in \{0,1\}^n} c_{q_1, \dots, q_n} |q_1\rangle \otimes |q_2\rangle \otimes |q_3\rangle \otimes \dots \otimes |q_n\rangle$$

where $c_{q_1, \dots, q_n} \in \mathbb{C}$ is the *amplitude* of each quantum state and each $q_n \in \{0, 1\}$.

The square of the amplitude is the *probability* of measurement with the output $|q_1\rangle \otimes |q_2\rangle \otimes |q_3\rangle \otimes \dots \otimes |q_n\rangle$ and the total probability should sum to 1.

$$\sum_{(q_1, q_2, \dots, q_n) \in \{0,1\}^n} \|c_{q_1, \dots, q_n}\|^2 = 1$$

For the *states* in the environment considered in this work, each is labeled with a integer number in the range from 0 to 15. Each *decimal* number is first converted into a *binary* number and then encoded in the quantum states. For example, the *state* observed by the agent, say, 12, is first converted to binary number 1100. It is denoted the binary number as $b_1 b_2 b_3 b_4$, therefore, the state indexed by the number 12 will be denoted as $b_1 = 1, b_2 = 1, b_3 = 0, b_4 = 0$. The quantum state for this can be denoted as $|1\rangle \otimes |1\rangle \otimes |0\rangle \otimes |0\rangle$. Next step is to construct the appropriate rotation to

Location	Reward
HOLE	-0.2
GOAL	+1.0
OTHER	-0.01

Table 1: The list of reward in our testing environment FrozenLake. Also note that the environment is non-slippery. Setting of this not only encourage the agent to achieve the goal but also encourage it to select the shortest path.

encode this state into the quantum circuit Fig1. The rotation angles are:

$$\begin{aligned}\theta_i &= \pi \times b_i \\ \phi_i &= \pi \times b_i\end{aligned}$$

where i represents the index of each qubit, in this work, the total number of qubits is 4 therefore the index is the set $\{1, 2, 3, 4\}$.

2.4 Numerical Simulation

We set up the experiment as follows Fig1. The quantum circuit as shown in the figure is numerically simulated with the software package PennyLane [3], in order to accelerate the simulation, the linear algebra manipulation is helped with the well-known package PyTorch [23].

The testing environment for all reinforcement learning agent is provided by OpenAI Gym[6], in this work, we choose a simple environment FrozenLake to implement the proof-of-concept experiments. The optimization is chosen to be *RMSprop* with *learning rate* = 0.01, *alpha* = 0.99 and *eps* = 10^{-8} , which is used widely in deep reinforcement learning. The batch-size for the experience replay is 5. The ϵ -greedy strategy used in this work is the following:

$$\epsilon \leftarrow \frac{\epsilon}{\frac{episode}{100} + 1}$$

with initial $\epsilon = 1.0$ for encouraging more exploration in early episodes and shifting to more exploitation in later episodes.

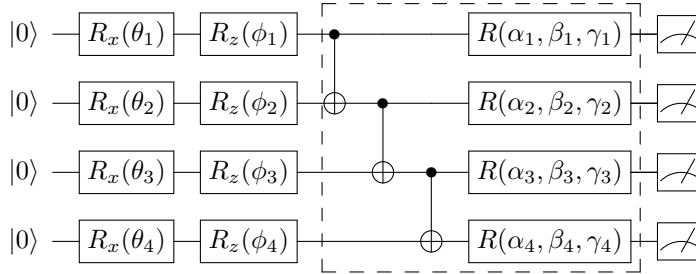


Figure 1: **Circuit architecture for the deep reinforcement learning.** This is the circuit architecture for deep Q learning. The parameters labeled θ and ϕ are for state preparation and are not for iterative optimization. The CNOT gates are used to entangle states. Parameters labeled α , β and γ are the ones for optimization. Note that the grouped box may repeat several times thus increase the number of parameters. In this work, the grouped circuit repeat two times and therefore the total number of parameters subject to optimization is $4 \times 3 \times 2 = 24$

3 Results

3.1 Numerical Simulations

In the numerical simulation, we ran 500 episodes and the agent converge to reward 0.9 after about 200 episodes. It is noted that, however, several sub-optimal results occurred. This phenomenon is probably due to the ϵ -greedy policy selection.

For the numerical simulation code and the parameters, please refer to the GitHub repository:

<https://github.com/ycchen1989/VariationalQuantumCircuitsAndDeepReinforcementLearning>

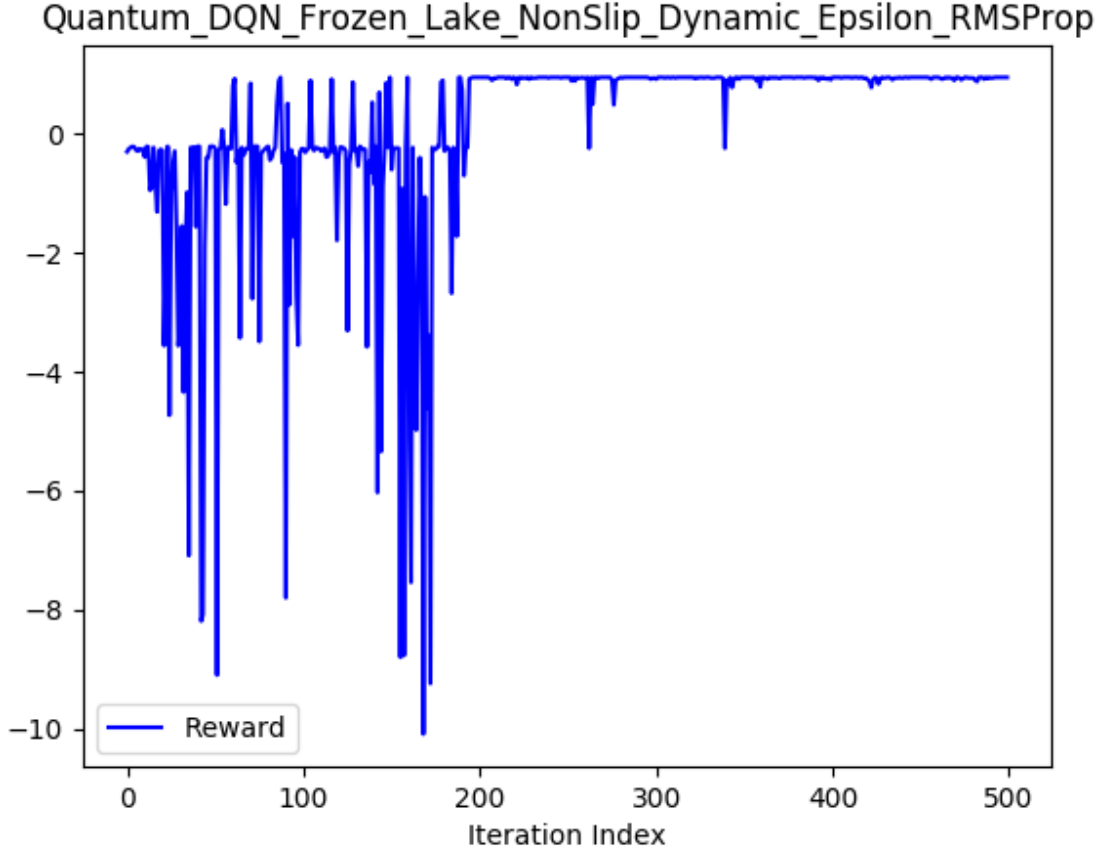


Figure 2: **Variational quantum circuit for Deep Q Learning** In 500 iterations, we can see that the agent reach the optimal policy with total reward about 0.9 after about 200 iterations.

4 Discussions

4.1 Relevant Works

For a general review in the field *quantum machine learning*, please refer to [10] and [11]. It is noted that some relevant works concerning *quantum reinforcement learning* is around. For example, [5] [17], in these works, a framework called *projective simulation* is proposed. The key concept in this setting is that the *agent* has a *memory* to keep the transition history and before executing each action, the agent will *simulate* several possible outcomes according to historical data in the memory. It is conceptually related to the well-known Monte-Carlo Tree Search in [30] [31] and it is interesting to investigate the quantum counterparts and possible quantum advantages.

4.2 Other RL Environments

In this work, we use a simple testing environment for the proof-of-concept experiment. The reasons are the following. First, for more complex reinforcement learning benchmark environments, for example, Atari games, the inputs are raw pixels. In these situations, we need much more qubits to conduct the simulations. Although it can be encoded with computational basis encoding or amplitude encoding, the number of qubits needed is intractable for numerical simulations on classical computers and also exceed commercially available quantum devices. It is possible to investigate these complex RL environments with the same setting proposed in this work when large quantum machines are released.

4.3 Other Encoding Scheme

Computational basis encoding in this work still not harvest the full quantum advantages compared to amplitude encoding. Although currently limited by the simulator, we still demonstrate the feasibility of applying quantum circuits to deep reinforcement learning problems. To gain the promising logarithmic less parameters with quantum circuits, *amplitude encoding* should be used. In that case, the complexity of parameters is $poly(\log N)$ while at the same input dimension N , conventional neural network is $poly(N)$, therefore, it is promising to investigate the possibility of applying these circuits to represent the high-dimensional data or the more complex environments.

5 Conclusions

This is the first demonstration of variational quantum circuits to approximate the deep Q -value function with experience replay and target network. Under appropriate encoding method, the possible quantum advantage is $poly(\log N)$ compared to conventional neural network with complexity $poly(N)$.

6 Acknowledgements

Thanks to Prof. Yun-Cheng Tsai for constructive discussion and great support. Thanks to Cheng-Lin Hong for helping on setting computational clusters.

References

- [1] AN, Zheng ; ZHOU, D. L.: Deep Reinforcement Learning for Quantum Gate Control. (2019), 2. – URL <http://arxiv.org/abs/1902.08418>
- [2] ANDREASSON, Philip ; JOHANSSON, Joel ; LILJESTRAND, Simon ; GRANATH, Mats: Quantum error correction for the toric code using deep reinforcement learning. (2018), 11. – URL <http://arxiv.org/abs/1811.12338>
- [3] BERGHOLM, Ville ; IZAAC, Josh ; SCHULD, Maria ; GOGOLIN, Christian ; BLANK, Carsten ; MCKIERNAN, Keri ; KILLORAN, Nathan: PennyLane: Automatic differentiation of hybrid quantum-classical computations. (2018), 11. – URL <http://arxiv.org/abs/1811.04968>
- [4] BORIN, Artem ; ABANIN, Dmitry A.: Approximating power of machine-learning ansatz for quantum many-body states. (2019), 1. – URL <http://arxiv.org/abs/1901.08615>
- [5] BRIEGEL, Hans J. ; CUEVAS, Gemma De las: Projective simulation for artificial intelligence. In: *Scientific Reports* 2 (2012), 12, Nr. 1, S. 400. – URL <http://www.nature.com/articles/srep00400>. – ISSN 2045-2322
- [6] BROCKMAN, Greg ; CHEUNG, Vicki ; PETTERSSON, Ludwig ; SCHNEIDER, Jonas ; SCHULMAN, John ; TANG, Jie ; ZAREMBA, Wojciech: OpenAI Gym. (2016), 6. – URL <http://arxiv.org/abs/1606.01540>
- [7] CANABARRO, Askery ; FANCHINI, Felipe F. ; MALVEZZI, André L. ; PEREIRA, Rodrigo ; CHAVES, Rafael: Unveiling phase transitions with machine learning. (2019), 4. – URL <http://arxiv.org/abs/1904.01486>
- [8] CARLEO, Giuseppe ; CHOO, Kenny ; HOFMANN, Damian ; SMITH, James E. T. ; WESTERHOUT, Tom ; ALET, Fabien ; DAVIS, Emily J. ; EFTHYMIU, Stavros ; GLASSER, Ivan ; LIN, Sheng-Hsuan ; MAURI, Marta ; MAZZOLA, Guglielmo ; MENDL, Christian B. ; NIEUWENBURG, Evert van ; O'REILLY, Ossian ; THÉVENIAUT, Hugo ; TORLAI, Giacomo ; WIETEK, Alexander: NetKet: A Machine Learning Toolkit for Many-Body Quantum Systems. (2019), 3. – URL <http://arxiv.org/abs/1904.00031>
- [9] CARLEO, Giuseppe ; CIRAC, Ignacio ; CRANMER, Kyle ; DAUDET, Laurent ; SCHULD, Maria ; TISHBY, Naftali ; VOGT-MARANTO, Leslie ; ZDEBOROVÁ, Lenka: Machine learning and the physical sciences. (2019), 3. – URL <http://arxiv.org/abs/1903.10563>
- [10] DUNJKO, Vedran ; BRIEGEL, Hans J.: Machine learning & artificial intelligence in the quantum domain: a review of recent progress. In: *Reports on Progress in Physics* 81 (2018), 7, Nr. 7, S. 074001. – URL <http://stacks.iop.org/0034-4885/81/i=7/a=074001?key=crossref.484b39e1cdde454de1cbc5aba8d6de34>. – ISSN 0034-4885
- [11] DUNJKO, Vedran ; TAYLOR, Jacob M. ; BRIEGEL, Hans J.: Advances in quantum reinforcement learning. In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 10 2017, S. 282–287. – URL <http://ieeexplore.ieee.org/document/8122616/>. – ISBN 978-1-5386-1645-1
- [12] EDET, Samuel: Recurrent Neural Networks in Forecasting S&P 500 Index. In: *SSRN Electronic Journal* (2017), 7. – URL <https://www.ssrn.com/abstract=3001046>. – ISSN 1556-5068

- [13] FARHI, Edward ; NEVEN, Hartmut: Classification with Quantum Neural Networks on Near Term Processors. (2018), 2. – URL <http://arxiv.org/abs/1802.06002>
- [14] FLURIN, Emmanuel ; MARTIN, Leigh S. ; HACOHEEN-GOURGY, Shay ; SIDDIQI, Irfan: Using a Recurrent Neural Network to Reconstruct Quantum Dynamics of a Superconducting Qubit from Physical Observations. (2018), 11. – URL <http://arxiv.org/abs/1811.12420>
- [15] HEDAYATI, Amin ; HEDAYATI, Moein ; ESFANDYARI, Morteza: *Stock Market Index Prediction Using Artificial Neural Network*. 7 2017. – URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3004032
- [16] LIAN, Wenqian ; WANG, Sheng-Tao ; LU, Sirui ; HUANG, Yuanyuan ; WANG, Fei ; YUAN, Xinxing ; ZHANG, Wengang ; OUYANG, Xiaolong ; WANG, Xin ; HUANG, Xianzhi ; HE, Li ; CHANG, Xiuying ; DENG, Dong-Ling ; DUAN, Lu-Ming: Machine Learning Topological Phases with a Solid-state Quantum Simulator. (2019), 5. – URL <http://arxiv.org/abs/1905.03255><http://dx.doi.org/10.1103/PhysRevLett.122.210503>
- [17] MAUTNER, Julian ; MAKMAL, Adi ; MANZANO, Daniel ; TIERSCH, Markus ; BRIEGEL, Hans J.: Projective Simulation for Classical Learning Agents: A Comprehensive Investigation. In: *New Generation Computing* 33 (2015), 1, Nr. 1, S. 69–114. – URL <http://link.springer.com/10.1007/s00354-015-0102-0>. – ISSN 0288-3635
- [18] MITARAI, Kosuke ; NEGORO, Makoto ; KITAGAWA, Masahiro ; FUJII, Keisuke: Quantum Circuit Learning. (2018), 3. – URL <https://arxiv.org/abs/1803.00745>
- [19] MNIH, Volodymyr ; BADIA, Adrià P. ; MIRZA, Mehdi ; GRAVES, Alex ; LILLICRAP, Timothy P. ; HARLEY, Tim ; SILVER, David ; KAVUKCUOGLU, Koray: Asynchronous Methods for Deep Reinforcement Learning. (2016), 2. – URL <http://arxiv.org/abs/1602.01783>
- [20] MNIH, Volodymyr ; KAVUKCUOGLU, Koray ; SILVER, David ; RUSU, Andrei A. ; VENESS, Joel ; BELLEMARE, Marc G. ; GRAVES, Alex ; RIEDMILLER, Martin ; FIDJELAND, Andreas K. ; OSTROVSKI, Georg ; PETERSEN, Stig ; BEATTIE, Charles ; SADIK, Amir ; ANTONOGLU, Ioannis ; KING, Helen ; KUMARAN, Dharshan ; WIERSTRA, Daan ; LEGG, Shane ; HASSABIS, Demis: Human-level control through deep reinforcement learning. In: *Nature* 518 (2015), 2, Nr. 7540, S. 529–533. – URL <http://www.nature.com/articles/nature14236>. – ISSN 0028-0836
- [21] MUNOZ-BAUZA, Humberto ; HAMZE, Firas ; KATZGRABER, Helmut G.: Learning to find order in disorder. (2019), 3. – URL <http://arxiv.org/abs/1903.06993>
- [22] NAUTRUP, Hendrik P. ; DELFOSSE, Nicolas ; DUNJKO, Vedran ; BRIEGEL, Hans J. ; FRIIS, Nicolai: Optimizing Quantum Error Correction Codes with Reinforcement Learning. (2018), 12. – URL <http://arxiv.org/abs/1812.08451>
- [23] PASZKE, Adam ; GROSS, Sam ; CHINTALA, Soumith ; CHANAN, Gregory ; YANG, Edward ; DEVITO, Zachary ; LIN, Zeming ; DESMAISON, Alban ; ANTIGA, Luca ; LERER, Adam: *Automatic differentiation in PyTorch*. 2017. – URL <https://www.semanticscholar.org/paper/Automatic-differentiation-in-PyTorch-Paszke-Gross/b36a5bb1707bb9c70025294b3a310138aae8327a>
- [24] SARMA, Sankar D. ; DENG, Dong-Ling ; DUAN, Lu-Ming: Machine learning meets quantum physics. (2019), 3. – URL <http://arxiv.org/abs/1903.03516><http://dx.doi.org/10.1063/PT.3.4164>
- [25] SCHULD, Maria ; BERGHOLM, Ville ; GOGOLIN, Christian ; IZAAC, Josh ; KILLORAN, Nathan: Evaluating analytic gradients on quantum hardware. In: *Physical Review A* 99 (2019), 3, Nr. 3, S. 032331. – URL <https://link.aps.org/doi/10.1103/PhysRevA.99.032331>. – ISSN 2469-9926
- [26] SCHULD, Maria ; BOCHAROV, Alex ; SVORE, Krysta ; WIEBE, Nathan: Circuit-centric quantum classifiers. (2018), 4. – URL <http://arxiv.org/abs/1804.00633>
- [27] SCHULD, Maria ; PETRUCCIONE, Francesco: Information Encoding. URL http://link.springer.com/10.1007/978-3-319-96424-9_5, 2018, S. 139–171
- [28] SEN, Jaydip ; CHAUDHURI, Tamal: *A Robust Predictive Model for Stock Price Forecasting*. 12 2017. – URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3068204
- [29] SEZER, Omer B. ; OZBAYOGLU, A. M. ; DOGDU, Erdogan: An Artificial Neural Network-based Stock Trading System Using Technical Analysis and Big Data Framework. In: *Proceedings of the SouthEast Conference on - ACM SE '17*. New York, New York, USA : ACM Press, 2017, S. 223–226. – URL <http://dl.acm.org/citation.cfm?doid=3077286.3077294>. – ISBN 9781450350242
- [30] SILVER, David ; HUANG, Aja ; MADDISON, Chris J. ; GUEZ, Arthur ; SIFRE, Laurent ; DRIESSCHE, George van den ; SCHRITTWIESER, Julian ; ANTONOGLU, Ioannis ; PANNEERSHELVAM, Veda ; LANCTOT, Marc ;

- DIELEMAN, Sander ; GREWE, Dominik ; NHAM, John ; KALCHBRENNER, Nal ; SUTSKEVER, Ilya ; LILLICRAP, Timothy ; LEACH, Madeleine ; KAVUKCUOGLU, Koray ; GRAEPEL, Thore ; HASSABIS, Demis: Mastering the game of Go with deep neural networks and tree search. In: *Nature* 529 (2016), 1, Nr. 7587, S. 484–489. – URL <http://www.nature.com/articles/nature16961>. – ISSN 0028-0836
- [31] SILVER, David ; SCHRITTWIESER, Julian ; SIMONYAN, Karen ; ANTONOGLU, Ioannis ; HUANG, Aja ; GUEZ, Arthur ; HUBERT, Thomas ; BAKER, Lucas ; LAI, Matthew ; BOLTON, Adrian ; CHEN, Yutian ; LILLICRAP, Timothy ; HUI, Fan ; SIFRE, Laurent ; DRIESSCHE, George van den ; GRAEPEL, Thore ; HASSABIS, Demis: Mastering the game of Go without human knowledge. In: *Nature* 550 (2017), 10, Nr. 7676, S. 354–359. – URL <http://www.nature.com/articles/nature24270>. – ISSN 0028-0836
- [32] SIMONYAN, Karen ; ZISSERMAN, Andrew: Very Deep Convolutional Networks for Large-Scale Image Recognition. (2014), 9. – URL <http://arxiv.org/abs/1409.1556>
- [33] SINGH, Ritika ; SRIVASTAVA, Shashi: Stock prediction using deep learning. In: *Multimedia Tools and Applications* 76 (2017), 9, Nr. 18, S. 18569–18584. – URL <http://link.springer.com/10.1007/s11042-016-4159-7>. – ISSN 1380-7501
- [34] SUTSKEVER, Ilya ; VINYALS, Oriol ; LE, Quoc V.: Sequence to Sequence Learning with Neural Networks. (2014), 9. – URL <http://arxiv.org/abs/1409.3215>
- [35] SUTTON, Richard S. ; BARTO, Andrew G.: *Reinforcement learning : an introduction*. – 526 S. – ISBN 0262039249
- [36] SZEGEDY, Christian ; LIU, Wei ; JIA, Yangqing ; SERMANET, Pierre ; REED, Scott ; ANGUELOV, Dragomir ; ERHAN, Dumitru ; VANHOUCHE, Vincent ; RABINOVICH, Andrew: Going Deeper with Convolutions. (2014), 9. – URL <http://arxiv.org/abs/1409.4842>
- [37] THÉVENIAUT, Hugo ; ALET, Fabien: Neural network setups for a precise detection of the many-body localization transition: finite-size scaling and limitations. (2019), 4. – URL <http://arxiv.org/abs/1904.13165>
- [38] TORLAI, Giacomo ; MELKO, Roger G.: Machine learning quantum states in the NISQ era. (2019), 5. – URL <http://arxiv.org/abs/1905.04312>
- [39] VOULODIMOS, Athanasios ; DOULAMIS, Nikolaos ; DOULAMIS, Anastasios ; PROTOPAPADAKIS, Eftychios: Deep Learning for Computer Vision: A Brief Review. In: *Computational Intelligence and Neuroscience* 2018 (2018), S. 1–13. – URL <http://www.ncbi.nlm.nih.gov/pubmed/29487619><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5816885><https://www.hindawi.com/journals/cin/2018/7068349/>. – ISSN 1687-5265