

Review

# Quantum Reinforcement Learning with Quantum Photonics

Lucas Lamata 

Departamento de Física Atómica, Molecular y Nuclear, Universidad de Sevilla, Apartado de Correos 1065, 41080 Sevilla, Spain; llamata@us.es

**Abstract:** Quantum machine learning has emerged as a promising paradigm that could accelerate machine learning calculations. Inside this field, quantum reinforcement learning aims at designing and building quantum agents that may exchange information with their environment and adapt to it, with the aim of achieving some goal. Different quantum platforms have been considered for quantum machine learning and specifically for quantum reinforcement learning. Here, we review the field of quantum reinforcement learning and its implementation with quantum photonics. This quantum technology may enhance quantum computation and communication, as well as machine learning, via the fruitful marriage between these previously unrelated fields.

**Keywords:** quantum machine learning; quantum reinforcement learning; quantum photonics; quantum technologies; quantum communication

## 1. Introduction

The field of quantum machine learning promises to employ quantum systems for accelerating machine learning [1] calculations, as well as employing machine learning techniques to better control quantum systems. In the past few years, several books as well as reviews on this topic have appeared [2–8].

Inside artificial intelligence and machine learning, the area of reinforcement learning designs “intelligent” agents capable of interacting with their outer world, the “environment”, and adapt to it, via reward mechanisms [9], see Figure 1. These agents aim at achieving a final goal that maximizes their long-term rewards. This kind of machine learning protocol is, arguably, the most similar one to the way the human brain learns. The field of quantum machine learning is recently exploring the fruitful combination of reinforcement learning protocols with quantum systems, giving rise to quantum reinforcement learning [10–39].

Different quantum platforms are being considered for the implementation of quantum machine learning. Among them, quantum photonics seems promising because of the good integration with communication networks, information processing at the speed of light, as well as possible realization of quantum computations with integrated photonics [40]. Moreover, in the scenario with a reduced amount of measurements, quantum reinforcement learning with quantum photonics has been shown to perform better than standard quantum tomography [16]. Quantum reinforcement learning with quantum photonics has been proposed [15,17,20] and implemented [16,19] in diverse works. Even before these articles were produced, a pioneering experiment of quantum supervised and unsupervised learning with quantum photonics was carried out [41].

In this review, we first give an overview of the field of quantum reinforcement learning, focusing mainly on quantum devices employed for reinforcement learning algorithms [10–20], in Section 2. Later on, we review the proposal for measurement-based adaptation protocol with quantum reinforcement learning [15] and its experimental implementation with quantum photonics [16], in Section 3. Subsequently, we describe further works in quantum reinforcement learning with quantum photonics [19,20], in Section 4. Finally, we give our conclusions in Section 5.



**Citation:** Lamata, L. Quantum Reinforcement Learning with Quantum Photonics. *Photonics* **2021**, *8*, 33. <https://doi.org/10.3390/photonics8020033>

Received: 26 December 2020

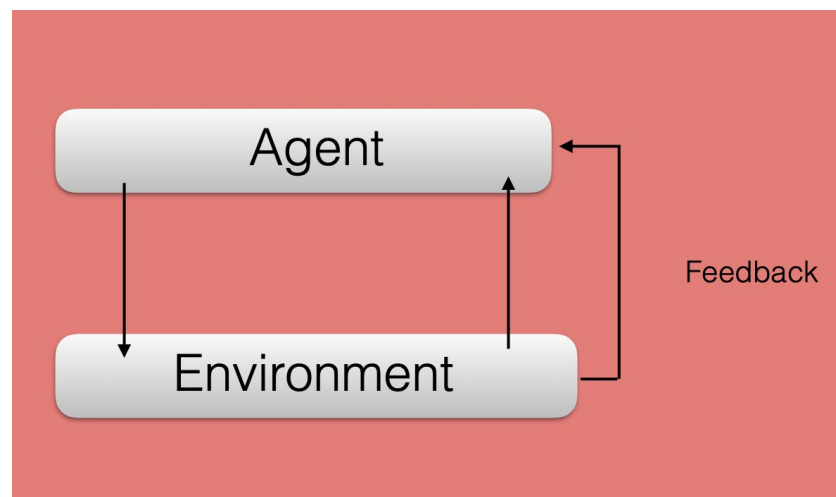
Accepted: 25 January 2021

Published: 28 January 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



**Figure 1.** Reinforcement learning protocol. A system, called *agent*, interacts with its external world, the *environment*, carrying out some action on it, while receiving information from it. Afterwards, the agent acts accordingly in order to achieve some long-term goal, via feedback with rewards, iterating the process several times.

## 2. Quantum Reinforcement Learning

The fields of reinforcement learning and quantum technologies have started to merge recently in a novel area, named quantum reinforcement learning [10–31,35]. A subset inside this field is composed of articles studying quantum systems that carry out reinforcement learning algorithms, ideally with some speedup [10–20].

In Ref. [10], a pioneer proposal for reinforcement learning using quantum systems was put forward. This employed a Grover-like search algorithm, which could provide a quadratic speedup in the learning process as compared to classical computers [10].

Ref. [11] provided a quantum algorithm for reinforcement learning in which a quantum agent, possessing a quantum processor, can couple classically with a classical environment, obtaining classical information from it. The speedup in this case would come from the quantum processing of the classical information, which could be done faster than with classical computers. This is also based on Grover search, with a corresponding quadratic speedup.

In Ref. [12], a quantum algorithm considers a quantum agent coupled to a quantum oracular environment, attaining a proven speedup with this kind of configuration, which can be exponential in some situations. The quantum algorithm could be applied to diverse kinds of learning, namely reinforcement learning, but also supervised and unsupervised learning.

Refs. [10–12] have speedups with respect to classical algorithms. While the first two rely on a polynomial gain due to a Grover-like algorithm, the latter achieves its proven speedup via a quantum oracular environment.

The series of articles in Refs. [13–18] study quantum reinforcement learning protocols with basic quantum systems coupled to small quantum environments. These works focus mainly on proposals for implementations [13–15,17] as well as experimental realizations in quantum photonics [16] and superconducting circuits [18]. In the theoretical proposals, small few-qubit quantum systems are proposed both for quantum agents and quantum environments. In Ref. [13], the aim of the agent is to achieve a final state which cannot be distinguished from the environment state, even if the latter has to be modified, as it is a single-copy protocol. In order to achieve this goal, measurements are allowed, as well as classical feedback inside the coherence time. Ref. [14] extends the previous protocol to the case in which measurements are not considered, but instead further ancillary qubits coupled via entangling gates to agent and environment are employed, and later on disregarded. In Ref. [15], several identical copies of the environment state are considered, such that the agent, via trial and error, or, equivalently, a balance between exploration and exploitation,

iteratively approaches the environment state. This proposal was carried out in a quantum photonics experiment [16] as well as with superconducting circuits [18]. In Ref. [17], a further extension of Ref. [15] to operator estimation, instead of state estimation, was proposed and analyzed.

Ref. [16] obtained a speedup as well with respect to standard quantum tomography, in the scenario with a reduced amount of resources, in the sense of reduced number of measurements.

Finally, Ref. [20] considered different paradigms of learning inside a reinforcement learning framework, which included projective simulation [42] and a possible implementation with quantum photonics devices. The latter, with high-repetition rates, high-bandwidth and low crosstalks, as well as the possibility to propagate to long distances, makes this quantum platform an attractive one for this kind of protocol.

### 3. Measurement-Based Adaptation Protocol with Quantum Reinforcement Learning Implemented with Quantum Photonics

#### 3.1. Theoretical Proposal

In this section we review the proposal in Ref. [15] introducing a quantum agent that can obtain information from several identical copies of an unknown environment state, via measurement, feedback, and application of random unitary gates. The randomness of previous operations is reduced as long as the agent approaches the environment state, ideally converging to a large extent to this.

The detailed protocol was as follows. One assumes a quantum system, the agent ( $A$ ), and many identical copies of an unknown quantum state, the environment ( $E$ ). An auxiliary system, the register ( $R$ ), which interacts with  $E$ , was also considered. Then, information about  $E$  is obtained by measuring  $R$ , and the result is employed as input to a reward function (RF). Finally, one carries out a partially random unitary operation on  $A$ , depending on the output of the RF. The aim is to increase the overlap between  $A$  and  $E$ , without measuring the state of  $A$ .

In the rest of the description of the protocol we use a notation as follows: the subscripts  $A$ ,  $R$ , and  $E$  are referred to each subsystem, while the superscripts denote the iteration. For example,  $\mathcal{O}_\alpha^{(k)}$  denotes the operator  $\mathcal{O}$  acting on subsystem  $\alpha$  during the  $k$ th iteration. In case we do not employ indices we refer to a general entity in the iterations as well as in the subsystems.

Here we will review the case for which the subsystems are respectively described by single-qubit states [15]. For a multilevel state description we refer to Ref. [15]. One considers that  $A(R)$  is encoded in  $|0\rangle_{A(R)}$ , while  $E$  is represented by an arbitrary state described as  $|\mathcal{E}\rangle_E = \cos(\theta^{(1)}/2)|0\rangle_E + e^{-i\phi^{(1)}} \sin(\theta^{(1)}/2)|1\rangle_E$ . The initial state is given by

$$|\psi^{(1)}\rangle = |0\rangle_A |0\rangle_R [\cos(\theta^{(1)}/2)|0\rangle_E + e^{i\phi^{(1)}} \sin(\theta^{(1)}/2)|1\rangle_E]. \tag{1}$$

One subsequently introduces the ingredients of the reinforcement learning protocol, including the policy, the RF, as well as the value function (VF). We point out that we consider a definition for the VF which is different from standard reinforcement learning but which will fulfill our purposes. For carrying out the policy, one performs a controlled-NOT (CNOT) gate ( $U_{E,R}^{NOT}$ ) where  $E$  is the control and  $R$  is the target (namely, the interaction with the environment system), in order that the information of  $E$  is transferred into  $R$ , achieving

$$|\Psi_1\rangle = U_{E,R}^{NOT} |\psi^{(1)}\rangle = |0\rangle_A \left[ \cos(\theta^{(1)}/2)|0\rangle_R |0\rangle_E + e^{i\phi^{(1)}} \sin(\theta^{(1)}/2)|1\rangle_R |1\rangle_E \right]. \tag{2}$$

One would then measure the register qubit in the  $\{|0\rangle, |1\rangle\}$  basis, with probabilities  $p_0^{(1)} = \cos^2(\theta_1/2)$  and  $p_1^{(1)} = \sin^2(\theta_1/2)$ , to achieve state  $|0\rangle$  or  $|1\rangle$ , respectively (namely, extraction of information). If the outcome is  $|0\rangle$ , one will have collapsed  $E$  into  $A$  such

that one does nothing, while if the outcome is  $|1\rangle$ , one has consequently measured the projection of  $E$  orthogonal to  $A$ , such that one accordingly updates the agent. Given that no further information about the environment is available, one carries out a partially-random unitary gate on  $A$  given by  $U_A^{(1)}(\alpha^{(1)}, \beta^{(1)}) = e^{-iS_A^{z(1)}\alpha^{(1)}} e^{-iS_A^{x(1)}\beta^{(1)}}$  (action), being  $\alpha^{(1)}$  and  $\beta^{(1)}$  random angles given by  $\alpha(\beta)^{(1)} = \zeta_{\alpha(\beta)}\Delta^{(1)}$ , while  $\zeta_{\alpha(\beta)} \in [-1/2, 1/2]$  is a random number,  $\Delta^{(1)}$  is the random angle range,  $\alpha(\beta)^{(1)} \in [-\Delta^{(1)}/2, \Delta^{(1)}/2]$ , and  $S_A^{k(1)} = S^k$  is the  $k$ th component of the spin. Subsequently, one initializes the register qubit and considers a novel copy of  $E$ , achieving the following initial state for the second iteration:

$$|\psi^{(2)}\rangle = \mathcal{U}_A^{(1)}|0\rangle_A|0\rangle_R|\mathcal{E}\rangle_E = |\bar{0}\rangle_A^{(2)}|0\rangle_R|\mathcal{E}\rangle_E, \tag{3}$$

with

$$\mathcal{U}_A^{(1)} = [m^{(1)}U_A^{(1)}(\alpha^{(1)}, \beta^{(1)}) + (1 - m^{(1)})\mathbb{I}_A]. \tag{4}$$

Here we denote with  $m^{(1)} = \{0, 1\}$  the result of the measurement, while  $\mathbb{I}$  is the unity operator, and we express the new agent state in the form  $|\bar{0}\rangle_1^{(2)} = \mathcal{U}_1^{(1)}|0\rangle_1$ .

Subsequently, one considers the RF to change the exploration interval of the  $k$ th iteration  $\Delta^{(k)}$  as

$$\Delta^{(k)} = [(1 - m^{(k-1)})\mathcal{R} + m^{(k-1)}\mathcal{P}]\Delta^{(k-1)}, \tag{5}$$

where we denote with  $m^{(k-1)}$  the result of the  $(k - 1)$ th iteration and with  $\mathcal{R}$  and  $\mathcal{P}$  the reward and punishment ratios, respectively.

Equation (5) represents the fact that  $\Delta$  is changed by  $\mathcal{R}\Delta$  for the subsequent iteration whenever  $m = 0$  and by  $\mathcal{P}\Delta$  whenever the result is  $m = 1$ . In the described protocol, one takes for the sake of simplicity  $\mathcal{R} = \epsilon < 1$  and  $\mathcal{P} = 1/\epsilon > 1$ , in such a way that the value of  $\Delta$  is reduced every time  $|0\rangle$  is measured, and it grows in the other situation. Moreover, given the fact that  $\mathcal{R} \cdot \mathcal{P} = 1$ , reward and punishment are of similar value, or, equivalently, if the algorithm provides equal number of results 0 and 1, the exploration interval is not modified. Additionally, one defines the VF as the  $\Delta^{(n)}$  after all the iterations have taken place. Thus,  $\Delta^{(n)} \rightarrow 0$  whenever the algorithm converges to a maximal overlap between  $A$  and  $E$ .

In order to show in further detail how the algorithm works, we consider the  $k$ th step. The initial state in the protocol reads

$$|\psi\rangle^{(k)} = |\bar{0}\rangle_A^{(k)}|0\rangle_R|\mathcal{E}\rangle_E, \tag{6}$$

with  $|\bar{0}\rangle_A^{(k)} = \mathbb{U}_A^{(k)}|0\rangle_A$  and  $\mathbb{U}_A^{(k)} = \mathcal{U}_A^{(k-1)}\mathbb{U}_A^{(k-1)}$ , where we have  $\mathbb{U}_A^{(1)} = \mathbb{I}_A$  while  $\mathcal{U}_A^{(j)}$  is provided by Equation (4). Moreover  $U_A^{(j)} = e^{-iS_A^{z(j)}\alpha^{(j)}} e^{-iS_A^{x(j)}\beta^{(j)}}$ , where one has

$$\begin{aligned} S_A^{z(j)} &= \frac{1}{2}(|\bar{0}\rangle_A^{(j)}\langle\bar{0}| - |\bar{1}\rangle_A^{(j)}\langle\bar{1}|) = \mathcal{U}_A^{(j-1)\dagger} S_A^{z(j-1)} \mathcal{U}_A^{(j-1)}, \\ S_A^{x(j)} &= \frac{1}{2}(|\bar{0}\rangle_A^{(j)}\langle\bar{1}| + |\bar{1}\rangle_A^{(j)}\langle\bar{0}|) = \mathcal{U}_A^{(j-1)\dagger} S_A^{x(j-1)} \mathcal{U}_A^{(j-1)}, \end{aligned} \tag{7}$$

where  ${}^j_A \langle \bar{0} | \bar{1} \rangle_A^j = 0$ . One can then express the state of system  $E$  in the Bloch basis, employing  $|\bar{0}_j\rangle$  as reference and subsequently act with the quantum gate  $\mathbb{U}_E^{(k)\dagger}$ , achieving for  $E$ ,

$$\begin{aligned} & \mathbb{U}_E^{(k)\dagger} |\mathcal{E}\rangle_E \\ &= \mathbb{U}_E^{(k)\dagger} \left[ \cos(\theta^{(k)}/2) |\bar{0}\rangle_E^{(k)} + e^{i\phi^{(k)}} \sin(\theta^{(k)}/2) |\bar{1}\rangle_E^{(k)} \right] \\ &= \cos(\theta^{(k)}/2) |0\rangle_E + e^{i\phi^{(k)}} \sin(\theta^{(k)}/2) |1\rangle_E = |\mathcal{E}\rangle_E^{(k)}. \end{aligned} \tag{8}$$

One can then express the states  $|\bar{0}^{(k)}\rangle$  and  $|\bar{1}^{(k)}\rangle$  by means of the initial logical vectors  $|0\rangle$  and  $|1\rangle$  as well as  $\theta^{(k)}$ ,  $\theta^{(1)}$ ,  $\phi^{(k)}$ , and  $\phi^{(1)}$  in the following way,

$$\begin{aligned} |\bar{0}\rangle^{(k)} &= \cos\left(\frac{\theta^{(1)} - \theta^{(k)}}{2}\right) |0\rangle + e^{i\phi^{(1)}} \sin\left(\frac{\theta^{(1)} - \theta^{(k)}}{2}\right) |1\rangle, \\ |\bar{1}\rangle^{(k)} &= -e^{-i\phi^{(k)}} \sin\left(\frac{\theta^{(1)} - \theta^{(k)}}{2}\right) |0\rangle \\ &\quad + e^{i(\phi^{(1)} - \phi^{(k)})} \cos\left(\frac{\theta^{(1)} - \theta^{(k)}}{2}\right) |1\rangle. \end{aligned} \tag{9}$$

Accordingly, the unitary gate  $\mathbb{U}^{(k)\dagger}$  carries out the required rotation to modify  $|\bar{0}^{(k)}\rangle \rightarrow |0\rangle$  and  $|\bar{1}^{(k)}\rangle \rightarrow |1\rangle$ . Subsequently, one applies the operation  $U_{E,R}^{NOT}$ ,

$$\begin{aligned} |\Phi^{(k)}\rangle &= U_{E,R}^{NOT} |\bar{0}\rangle_A^{(k)} |0\rangle_R |\bar{\mathcal{E}}\rangle_E \\ &= |\bar{0}\rangle_A^{(k)} \left[ \cos(\theta^{(k)}/2) |0\rangle_R |0\rangle_E \right. \\ &\quad \left. + e^{i\phi^{(k)}} \sin(\theta^{(k)}/2) |1\rangle_R |1\rangle_E \right], \end{aligned} \tag{10}$$

and later one measures  $R$ , with respective probabilities  $p_0^{(k)} = \cos^2(\theta^{(k)}/2)$  and  $p_1^{(k)} = \sin^2(\theta^{(k)}/2)$  for the results  $m^{(k)} = 0$  and  $m^{(k)} = 1$ . Then, one acts with the RF provided by Equation (5). We remark that, statistically, when  $p_0^{(k)} \rightarrow 1$ ,  $\Delta \rightarrow 0$ , and when  $p_1^{(k)} \rightarrow 1$ ,  $\Delta \rightarrow 4\pi$ . With respect to the exploration–exploitation tradeoff, this implies that whenever the exploitation is reduced (one obtains  $|1\rangle$  many times), one increases the exploration (the value of  $\Delta$  grows) in order that the probability of having a beneficial change increases, while whenever the exploitation grows (one obtains  $|0\rangle$  often), one reduces the exploration range to permit only small subsequent modifications.

### 3.2. Implementation with Quantum Photonics

In Ref. [16], an experiment of the previous proposal described in Section 3.1 was carried out. The experiment aimed to estimate an unknown quantum state in a photonic system, in a scenario with a reduced amount of copies. A partially quantum reinforcement learning protocol was used in order to adapt a qubit state, namely, the “agent”, to the given unknown quantum state, i.e., the “environment”, via iterated single-shot projective measurements followed by feedback, with the aim of achieving maximum fidelity. The experimental setup, consisting of a quantum photonics device, could modify the available parameters to change the agent system according to the measurement results “0” and “1” in the environment state, namely, reward/punishment feedback. The experimental results showed that the proposed protocol provides a technique for estimating an unknown photonic quantum state whenever only a limited amount of copies is given, while it can be extended as well to higher dimensionality, multipartite, and density matrix quantum state

cases. The achieved fidelities in the protocol for the single-photon case were over 88% with an appropriate reward/punishment ratio with 50 iterations.

#### 4. Further Developments of Quantum Reinforcement Learning with Quantum Photonics

Without the aim of being exhaustive, here we briefly describe some other works that have appeared in the literature on the field of quantum reinforcement learning with quantum photonics.

In Ref. [19], it was argued that future breakthroughs in experimental quantum science would require dealing with complex quantum systems and therefore require complex and expensive experiments. Designing such complicated experiments is hard and could be enhanced with the aid of artificial intelligence. In this reference, the authors showed an automated learning system which learns to create complex quantum experiments, without considering previous knowledge or, sometimes, incorrect intuition. Their device did not only learn how to create the design of quantum experiments better than in previous works, but it also discovered in the process nontrivial experimental tools. The conclusion they obtained is that learning devices can provide crucial advances in the design and generation of novel experiments.

In Ref. [20], the authors introduced a blueprint for a quantum photonics realization of active learning devices employing machine learning algorithms such as, e.g., SARSA, Q-learning, and projective simulation. They carried out numerical calculations to evaluate the performance of their algorithm with customary reinforcement learning situations, obtaining that reasonable amounts of experimental errors could be allowed or, sometimes, benefit the learning protocol. Among other features, they showed that their designed device would enable features like abstraction and generalization, two aspects considered to be crucial for artificial intelligence. They consider for their studied model a quantum photonics platform, which, they argue, is scalable as well as simple, and proof-of-principle integration in quantum photonics devices seems feasible with near-term platforms.

More specifically, in Ref. [20], the main novelties are two. (i) Firstly, they describe a quantum photonics platform that allows reinforcement learning algorithms for working directly in combination with optical applications. For this aim, they focus on linear optics for its simplicity and well-established fabrication technology when compared to solid-state processors. They mention the example that nanosecond-scale reconfigurability and routing have already been achieved, and furthermore, photonic platforms allow one for decision-making at the speed of light, the fastest possible, which is only constrained by generation and detection rates. Energy efficiency in memories is also a bonus that photonic technologies may provide [20]. (ii) The second achievement of the article is the analysis of a variant of projective simulation based on binary decision trees, which is connected closely to standard projective simulation and suitable for photonic circuit implementations. Finally, they discuss how this development would enable key aspects of artificial intelligence, which are generalization and abstraction [20].

#### 5. Conclusions

In this article, we reviewed the field of quantum reinforcement learning with quantum photonics. Without the goal of being exhaustive, we have firstly reviewed the area of quantum reinforcement learning in general, showing that automated quantum agents can provide sometimes enhancements with respect to classical computers. Later on, we have described a theoretical proposal and its quantum photonics experimental realization of a quantum reinforcement learning protocol for state estimation. This protocol has been shown to have a speedup with respect to standard quantum tomography in the reduced resource scenario. Finally, we have briefly reviewed some other works in the field of quantum reinforcement learning with quantum photonics that appeared in the literature.

The field of quantum reinforcement learning may provide quantum systems with a larger amount of autonomy and independence. Quantum photonics is among the quantum platforms where this kind of technology could be highly fruitful. Even though

the amount of qubits is often not as large as in other platforms such as trapped ions and superconducting circuits, quantum photonics processes information at the speed of light, and it can be suitably interfaced with long-distance quantum communication protocols. A long-term scope in quantum reinforcement learning could be to combine this paradigm with quantum artificial life [8]. This will allow one to achieve fully autonomous quantum individuals that can reproduce, evolve, as well as interact and adapt to their environment. Further benefits in areas such as neuroscience could emerge as a consequence of this promising avenue.

**Funding:** This research was funded by PGC2018-095113-B-I00, PID2019-104002GB-C21, and PID2019-104002GB-C22 (MCIU/AEI/FEDER, UE).

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson: London, UK, 2009.
2. Wittek, P. *Quantum Machine Learning*; Academic Press: Cambridge, MA, USA, 2014.
3. Schuld, M.; Petruccione, F. *Supervised Learning with Quantum Computers*; Springer: Berlin/Heidelberg, Germany, 2018.
4. Schuld, M.; Sinayskiy, I.; Petruccione, F. An introduction to quantum machine learning. *Contemp. Phys.* **2015**, *56*, 172. [[CrossRef](#)]
5. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 074001. [[CrossRef](#)] [[PubMed](#)]
6. Dunjko, V.; Briegel, H.J. Machine learning & artificial intelligence in the quantum domain: A review of recent progress. *Rep. Prog. Phys.* **2018**, *81*, 074001. [[PubMed](#)]
7. Schuld, M.; Sinayskiy, I.; Petruccione, F. The quest for a Quantum Neural Network. *Quantum Inf. Process.* **2014**, *13*, 2567. [[CrossRef](#)]
8. Lamata, L. Quantum machine learning and quantum biomimetics: A perspective. *Mach. Learn. Sci. Technol.* **2020**, *1*, 033002. [[CrossRef](#)]
9. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
10. Dong, D.; Chen, C.; Li, H.; Tarn, T.-J. Quantum Reinforcement Learning. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2008**, *38*, 1207. [[CrossRef](#)]
11. Paparo, G.D.; Dunjko, V.; Makmal, A.; Martin-Delgado, M.A.; Briegel, H.J. Quantum Speedup for Active Learning Agents. *Phys. Rev. X* **2014**, *4*, 031002. [[CrossRef](#)]
12. Dunjko, V.; Taylor, J.M.; Briegel, H.J. Quantum-Enhanced Machine Learning. *Phys. Rev. Lett.* **2016**, *117*, 130501. [[CrossRef](#)]
13. Lamata, L. Basic protocols in quantum reinforcement learning with superconducting circuits. *Sci. Rep.* **2017**, *7*, 1609. [[CrossRef](#)]
14. Cárdenas-López, F.A.; Lamata, L.; Retamal, J.C.; Solano, E. Multiqubit and multilevel quantum reinforcement learning with quantum technologies. *PLoS ONE* **2018**, *13*, e0200455. [[CrossRef](#)]
15. Albarrán-Arriagada, F.; Retamal, J.C.; Solano, E.; Lamata, L. Measurement-based adaptation protocol with quantum reinforcement learning. *Phys. Rev. A* **2018**, *98*, 042315. [[CrossRef](#)]
16. Yu, S.; Albarrán-Arriagada, F.; Retamal, J.C.; Wang, Y.-T.; Liu, W.; Ke, Z.-J.; Meng, Y.; Li, Z.-P.; Tang, J.-S.; Solano, E.; et al. Reconstruction of a Photonic Qubit State with Reinforcement Learning. *Adv. Quantum Technol.* **2019**, *2*, 1800074. [[CrossRef](#)]
17. Albarrán-Arriagada, F.; Retamal, J.C.; Solano, E.; Lamata, L. Reinforcement learning for semi-autonomous approximate quantum eigensolver. *Mach. Learn. Sci. Technol.* **2020**, *1*, 015002. [[CrossRef](#)]
18. Olivares-Sánchez, J.; Casanova, J.; Solano, E.; Lamata, L. Measurement-Based Adaptation Protocol with Quantum Reinforcement Learning in a Rigetti Quantum Computer. *Quantum Rep.* **2020**, *2*, 293. [[CrossRef](#)]
19. Melnikov, A.A.; Nautrup, H.P.; Krenn, M.; Dunjko, V.; Tiersch, M.; Zeilinger, A.; Briegel, H.J. Active learning machine learns to create new quantum experiments. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 1221. [[CrossRef](#)]
20. Flamini, F.; Hamann, A.; Jerbi, S.; Trenkwalder, L.M.; Poulsen, Nautrup, H.; Briegel, H.J. Photonic architecture for reinforcement learning. *New J. Phys.* **2020**, *22*, 045002. [[CrossRef](#)]
21. Fösel, T.; Tighineanu, P.; Weiss, T.; Marquardt, F. Reinforcement Learning with Neural Networks for Quantum Feedback. *Phys. Rev. X* **2018**, *8*, 031084. [[CrossRef](#)]
22. Bukov, M. Reinforcement learning for autonomous preparation of Floquet-engineered states: Inverting the quantum Kapitza oscillator. *Phys. Rev. B* **2018**, *98*, 224305. [[CrossRef](#)]
23. Bukov, M.; Day, A.G.R.; Sels, D.; Weinberg, P.; Polkovnikov, A.; Mehta, P. Reinforcement Learning in Different Phases of Quantum Control. *Phys. Rev. X* **2018**, *8*, 031086. [[CrossRef](#)]
24. Melnikov, A.A.; Sekatski, P.; Sangouard, N. Setting up experimental Bell test with reinforcement learning. *arXiv* **2020**, arXiv:2005.01697.
25. Mackeprang, J.; Dasari, D.B.R.; Wrachtrup, J. A Reinforcement Learning approach for Quantum State Engineering. *arXiv* **2019**, arXiv:1908.05981.

26. Schäfer, F.; Kloc, M.; Bruder, C.; Lörch, N. A differentiable programming method for quantum control. *arXiv* **2002**, arXiv:2002.08376.
27. Sgroi, P.; Palma, G.M.; Paternostro, M. Reinforcement learning approach to non-equilibrium quantum thermodynamics. *arXiv* **2020**, arXiv:2004.07770.
28. Wallnöfer, Melnikov, A.A.; Dür, W.; Briegel, H.J. Machine learning for long-distance quantum communication. *arXiv* **2019**, arXiv:1904.10797.
29. Zhang, X.-M.; Wei, Z.; Asad, R.; Yang, X.-C.; Wang, X. When does reinforcement learning stand out in quantum control? A comparative study on state preparation. *npj Quantum Inf.* **2019**, *5*, 85. [[CrossRef](#)]
30. Xu, H.; Li, J.; Liu, L.; Wang, Y.; Yuan, H.; Wang, X. Generalizable control for quantum parameter estimation through reinforcement learning. *npj Quantum Inf.* **2019**, *5*, 82. [[CrossRef](#)]
31. Sweke, R.; Kesselring, M.S.; van Nieuwenburg, E.P.L.; Eisert, J. Reinforcement Learning Decoders for Fault-Tolerant Quantum Computation. *arXiv* **2018**, arXiv:1810.07207.
32. Andreasson, P.; Johansson, J.; Liljestr, S.; Granath, M. Quantum error correction for the toric code using deep reinforcement learning. *Quantum* **2019**, *3*, 183. [[CrossRef](#)]
33. Nautrup, H.P.; Delfosse, N.; Dunjko, V.; Briegel, H.J.; Friis, N. Optimizing Quantum Error Correction Codes with Reinforcement Learning. *Quantum* **2019**, *3*, 215. [[CrossRef](#)]
34. Fitzek, D.; Eliasson, M.; Kockum, A.F.; Granath, M. Deep Q-learning decoder for depolarizing noise on the toric code. *Phys. Rev. Res.* **2020**, *2*, 023230. [[CrossRef](#)]
35. Fösel, T.; Krastanov, S.; Marquardt, F.; Jiang, L. Efficient cavity control with SNAP gates. *arXiv* **2020**, arXiv:2004.14256.
36. McKiernan, K.A.; Davis, E.; Sohaib, Alam, M.; Rigetti, C. Automated quantum programming via reinforcement learning for combinatorial optimization. *arXiv* **2019**, arXiv:1908.08054.
37. Garcia-Saez, A.; Riu, J. Quantum Observables for continuous control of the Quantum Approximate Optimization Algorithm via Reinforcement Learning. *arXiv* **2019**, arXiv:1911.09682.
38. Khairy, K.; Shaydulin, R.; Cincio, L.; Alexeev, Y.; Balaprakash, P. Learning to Optimize Variational Quantum Circuits to Solve Combinatorial Problems. *arXiv* **2019**, arXiv:1911.11071.
39. Yao, J.; Bukov, M.; Lin, L. Policy Gradient based Quantum Approximate Optimization Algorithm. *arXiv* **2020**, arXiv:2002.01068.
40. Flamini, F.; Spagnolo, N.; Sciarrino, F. Photonic quantum information processing: A review. *Rep. Prog. Phys.* **2019**, *82*, 016001. [[CrossRef](#)]
41. Cai, X.-D.; Wu, D.; Su, Z.-E.; Chen, M.-C.; Wang, X.-L.; Li, L.; Liu, N.-L.; Lu, C.-Y.; Pan, J.-W. Entanglement-Based Machine Learning on a Quantum Computer. *Phys. Rev. Lett.* **2015**, *114*, 110504. [[CrossRef](#)]
42. Briegel, H.J.; De las Cuevas, G. Projective simulation for artificial intelligence. *Sci. Rep.* **2012**, *2*, 1. [[CrossRef](#)]

## Biography

Prof. Lucas Lamata is Associate Professor of Theoretical Physics at the Departamento de Física Atómica, Molecular y Nuclear, Facultad de Física, Universidad de Sevilla, Spain. He carried out his PhD at CSIC, Madrid, and Universidad Autónoma de Madrid, in 2007, with an Extraordinary Award for a PhD in Physics. Later on, he was Humboldt Fellow and Max-Planck Postdoctoral Fellow for more than three years at the Max-Planck Institute of Quantum Optics, Garching, Germany. Subsequently, he was a Marie Curie IEF Postdoctoral Fellow, Ramón y Cajal Fellow, and tenured scientist, at Universidad del País Vasco, Bilbao, Spain, for more than eight years. Since 2019, he achieved his tenured Associate Professor position at Universidad de Sevilla, where he leads his group on quantum optics, quantum technologies, and quantum artificial intelligence.