Active learning machine learns to create new quantum experiments

Alexey A. Melnikov[a,1,2], Hendrik Poulsen Nautrup[a,1], Mario Krenn[b,c], Vedran Dunjko[a,3], Markus Tiersch[a], Anton Zeilinger[b,c,2], and Hans J. Briegel[a,d]

[a]Institute for Theoretical Physics, University of Innsbruck, 6020 Innsbruck, Austria; [b]Vienna Center for Quantum Science and Technology, Faculty of Physics, University of Vienna, 1090 Vienna, Austria; [c]Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, 1090 Vienna, Austria; and [d]Department of Philosophy, University of Konstanz, 78457 Konstanz, Germany

How useful can machine learning be in a quantum laboratory? Here we raise the question of the potential of intelligent machines in the context of scientific research. A major motivation for the present work is the unknown reachability of various entanglement classes in quantum experiments. We investigate this question by using the projective simulation model, a physics-oriented approach to artificial intelligence. In our approach, the projective simulation system is challenged to design complex photonic quantum experiments that produce high-dimensional entangled multiphoton states, which are of high interest in modern quantum experiments. The artificial intelligence system learns to create a variety of entangled states and improves the efficiency of their realization. In the process, the system autonomously (re)discovers experimental techniques which are only now becoming standard in modern quantum optical experiments—a trait which was not explicitly demanded from the system but emerged through the process of learning. Such features highlight the possibility that machines could have a significantly more creative role in future research.

machine learning | quantum experiments | quantum entanglement | artificial intelligence | quantum machine learning

Automated procedures are indispensable in modern science. Computers rapidly perform large calculations and help us visualize results, and specialized robots perform preparatory laboratory work to staggering precision. But what is the true limit of the utility of machines for science? To what extent can a machine help us understand experimental results? Could it—the machine—discover new useful experimental tools? The impressive recent results in the field of machine learning, from reliably analyzing photographs (1) to beating the world champion in the game of Go (2), support optimism in this regard. Researchers from various research fields now use machine learning algorithms (3), and the success of machine learning applied to physics (4–7) in particular is already noteworthy. Machine learning has claimed its place as the new data analysis tool in the physicist's toolbox (8). However, the true limit of machines lies beyond data analysis, in the domain of the broader theme of artificial intelligence (AI), which is much less explored in this context. In an AI picture, we consider devices (generally called intelligent agents) that interact with an environment (the laboratory) and learn from previous experience (9). To broach the broad questions raised above, in this work we address the specific yet central question of whether an intelligent machine can propose novel and elucidating quantum experiments. We answer in the affirmative in the context of photonic quantum experiments, although our techniques are more generally applicable. We design a learning agent, which interacts with (the simulations of) optical tables and learns how to generate novel and interesting experiments. More concretely, we phrase the task of the development of new optical experiments in a reinforcement learning (RL) framework (10), vital in modern AI (2, 11).

The usefulness of automated designs of quantum experiments has been shown in ref. 12. There, the algorithm MELVIN starts from a toolbox of experimentally available optical elements to randomly create simulations of setups. Those setups are reported if they result in high-dimensional multipartite entangled states. The algorithm uses handcrafted rules to simplify setups and is capable of learning by extending its toolbox with previously successful experimental setups. Several of these experimental proposals have been implemented successfully in the laboratory (13–16) and have led to the discovery of new quantum techniques (17, 18).

Inspired by the success of the MELVIN algorithm in the context of finding specific optical setups, here we investigate the broader potential of learning machines and AI in designing quantum experiments (19). Specifically, we are interested in their potential to contribute to novel research, beyond rapidly identifying solutions to fully specified problems. To investigate this question, we use a more general model of a learning agent, formulated within the projective simulation (PS) framework for artificial intelligence (19), which we apply to the concrete testbed of ref. 12. In the process of generating specified optical experiments, the learning agent builds up a memory network of correlations between different optical components—a feature it later

Significance

Quantum experiments push the envelope of our understanding of fundamental concepts in quantum physics. Modern experiments have exhaustively probed the basic notions of quantum theory. Arguably, further breakthroughs require the tackling of complex quantum phenomena and consequently require complex experiments and involved techniques. The designing of such complex experiments is difficult and often clashes with human intuition. We present an autonomous learning model which learns to design such complex experiments, without relying on previous knowledge or often flawed intuition. Our system not only learns how to design desired experiments more efficiently than the best previous approaches, but in the process also discovers nontrivial experimental techniques. Our work demonstrates that learning machines can offer dramatic advances in how experiments are generated.

PHYSICS

exploits when asked to generate targeted experiments efficiently. In the process of learning, it also develops notions [technically speaking, composite clips (19)], for components that "work well" in combination. In effect, the learning agent autonomously discovers subsetups (or gadgets) which are useful outside of the given task of searching for particular high-dimensional multiphoton entanglement. (The discovery of such devices is in fact a byproduct stemming from the more involved structure of the learning model.)

We concentrate on the investigation of multipartite entanglement in high dimensions (20–22) and give two examples where learning agents can help. The understanding of multipartite entanglement remains one of the outstanding challenges of quantum information science, not only because of the fundamental conceptual role of entanglement in quantum mechanics, but also because of the many applications of entangled states in quantum communication and computation (23–28). As the first example, the agent is tasked to find the simplest setup that produces a quantum state with a desired set of properties. Such efficient setups are important as they can be robustly implemented in the laboratory. The second task is to generate as many experiments, which create states with such particular properties, as possible. Having many different setups available is important for understanding the structure of the state space reachable in experiments and for exploration of different possibilities that are accessible in experiments. In both tasks the desired property is chosen to be a certain type of entanglement in the generated states. More precisely we target high-dimensional ($d > 2$) many-particle ($n > 2$) entangled states. The orbital angular momentum (OAM) of photons (29–32) can be used for investigations into high-dimensional (33–37) or multiphoton entanglement (38, 39) and, since recently, also both simultaneously (13). OAM setups have been used as testbeds for other new techniques for quantum optics experiments (12, 17), and they are the system of choice in this work as well.

Results

Our learning setup can be put in terms of a scheme visualized in Fig. 1*A*. The agent (our learning algorithm) interacts with a virtual environment—the simulation of an optical table. The agent has access to a set of optical elements (its toolbox), with which it generates experiments: It sequentially places the chosen elements on the (simulated) table; following the placement of an element, the quantum state generated by the corresponding setup, i.e., configuration of optical elements, is analyzed. Depending on the state, and the chosen task, the agent either receives a reward or not. The agent then observes the current table configuration and places another element; thereafter, the procedure is iterated. We are interested in finite experiments, so the maximal number of optical elements in one experiment is limited. This goes along with practical restrictions: Due to accumulation of imperfections, e.g., through misalignment and interferometric instability, for longer experiments we expect a decreasing overall fidelity of the resulting state or signal. The experiment ends (successfully) when the reward is given or (unsuccessfully) when the maximal number of elements on the table is reached without obtaining the reward. Over time, the agent learns which element to place next, given a table.

Initially, we fix the analyzer in Fig. 1*A* to reward the successful generation of a state from a certain class of high-dimensional multipartite entangled states. Conceptually, this choice of the reward function explicitly specifies our criterion for an experimental setup (and the states that are created) to be interesting, but we will come back to a more general scenario momentarily. We characterize interesting states by a Schmidt–Rank vector (SRV) (21, 22)—the numerical vector containing the rank of the reduced-density matrix of each of the subsystems—together with the requirement that the states are maximally entangled in the OAM basis (12, 13).

The initial state of the simulated experiment is generated by a double spontaneous parametric down-conversion (SPDC) process in two nonlinear crystals. Similar to refs. 12 and 13, we ignore higher-order terms with OAM $|m| > 1$ from down-conversion as their amplitudes are significantly smaller than those of the low-order terms. Neglecting these higher-order terms in the down-conversion, the initial state $|\psi(0)\rangle$ can be written as a tensor product of two pairs of OAM-entangled photons (13),

$$|\psi(0)\rangle = \frac{1}{3}\left(\sum_{m=-1}^{1} |m\rangle_a |-m\rangle_b\right) \otimes \left(\sum_{m=-1}^{1} |m\rangle_c |-m\rangle_d\right), \quad [1]$$

where the indexes a, b, c, and d specify four arms in the optical setup. The toolbox contains a basic set of elements (12) including beam splitters (BS), mirrors (Refl), shift-parameterized holograms (Holo), and Dove prisms (DP). Taking into account that each distinct element can be placed in any one (or two in the
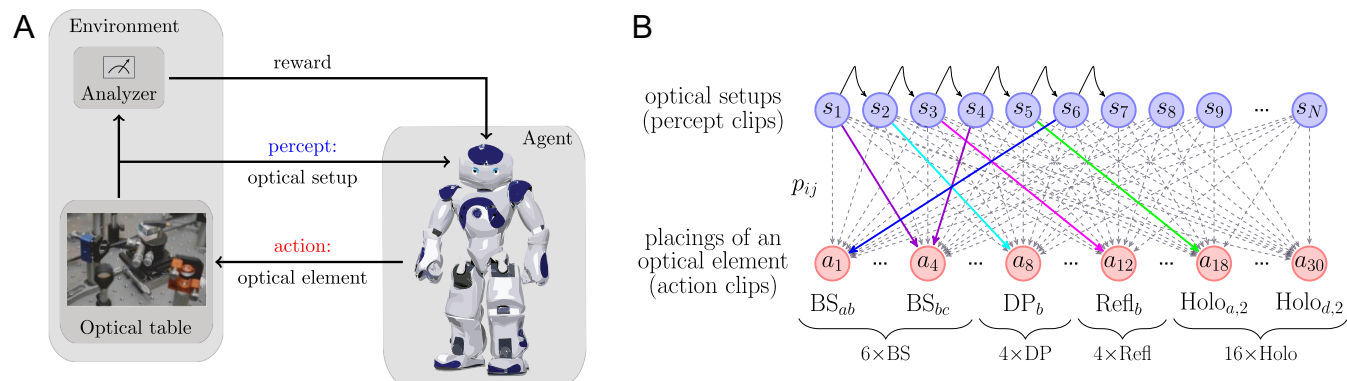
Fig. 1. The learning agent. (*A*) An agent is always situated in an environment (9). Through sensors it perceives optical setups and with actuators it can place optical elements in an experiment. Note that, in this paper, the interaction between the agent and the environment was entirely simulated on a classical computer. (One could imagine that in the future, a real robot builds up the experiment designed by the computer.) On the side, an analyzer evaluates a proposed experiment corresponding to the current optical setup and gives rewards according to a specified task. Image of the optical table courtesy of Manuel Erhard (University of Vienna, Vienna) and is a part of the (3, 3, 3) experiment. Image of the agent reproduced from https://openclipart.org/detail/266420/Request. (*B*) The memory network that represents the internal structure of the PS agent. Dashed arrows indicate possible transitions from percept clips (blue circles) to action clips (red circles). Solid, colored arrows depict a scenario where a sequence of actions leads to the experiment {BS$_{bc}$, DP$_b$, Refl$_b$, BS$_{bc}$, Refl$_b$, Holo$_{a,2}$}. Arrows between percepts correspond to deterministic transitions from one experiment to another after placement of an optical element.

case of BS) of the four arms a, b, c, d, we allow in total 30 different choices of elements. Since none of the optical elements in our toolbox creates entanglement in OAM degrees of freedom, we use a measurement in arm a and postselection to "trigger" a tripartite state in the other three arms.

Our learning agent is based on the PS (19) model for AI. PS is a physics-motivated framework which can be used to construct RL agents. PS was shown to perform well in standard RL problems (40–43) and in advanced robotics applications (44), and it is also amenable for quantum enhancements (45–47). The main component of the PS agent is its memory network (shown in Fig. 1B) comprising units of episodic memory called clips. Here, clips include remembered percepts (in our case, the observed optical tables) and actions (corresponding to the placing of an optical element). Each percept (clip) s_i, $i \in [1, \ldots, N]$ is connected to every action (clip) a_j, $j \in [1, \ldots, 30]$ via a weighted directed edge (i, j), which represents the possibility of taking an action a_j in a situation s_i with probability $p_{i,j}$ (Fig. 1B). The process of learning is manifested in the creation of new clips and in the adjustment of the probabilities. Intuitively, the probabilities of percept–action transitions which eventually lead to a reward will be enhanced, leading to a higher likelihood of rewarding behavior in the future (see *Projective Simulation* for details).

Designing Short Experiments. As mentioned, a maximum number of elements placed in one experiment is introduced to limit the influence of experimental imperfections. For the same reason, it is valuable to identify the shortest experiments that produce high-dimensional tripartite entanglement characterized by a desired SRV. To relate our work to existing experiments (13), we task the agent with designing a setup that creates a state with SRV $(3, 3, 2)$. To further emphasize the benefits of learning, we then investigate whether the agent can use its knowledge, attained through the learning setting described previously, to discover a more complex quantum state. Thus, the task of finding a state with SRV $(3, 3, 2)$ is followed by the task of finding a state with SRV $(3, 3, 3)$, after 5×10^4 simulated experiments. As always, a reward is issued whenever a target state is found, and the table is reset. Fig. 2A shows the success probability throughout the learning process: The PS agent first learns to construct a $(3, 3, 2)$ state and then, with probability 0.5, very quickly [compared to the by itself simpler task of finding a $(3, 3, 2)$ state] learns to design a setup that corresponds to a $(3, 3, 3)$ state. Our results suggest that either the knowledge of constructing a $(3, 3, 2)$ state is highly beneficial in the second phase or

the more complicated $(3, 3, 3)$ state is easier to generate. To resolve this dichotomy, we simulated a learning process where the PS agent is required to construct a $(3, 3, 3)$ state within 6×10^4 experiments, without having learned to build a $(3, 3, 2)$ state during the first 5×10^4 experiments. The results of the simulations are shown in Fig. 2A as dashed lines (lower and upper edges of the frame). It is apparent that the agent without previous training on $(3, 3, 2)$ states does not show any significant progress in constructing a $(3, 3, 3)$ experiment. Furthermore, the PS agent constantly and autonomously improves by constructing shorter and shorter experiments (Fig. S1A). By the end of the first phase, PS almost always constructs experiments of length 4—the shortest length of an experiment producing a state from the $(3, 3, 2)$ SRV class. During the $(3, 3, 3)$ learning phase, the PS agent produces a $(3, 3, 3)$ state of the shortest length in half of the cases. Experimental setups useful for the second phase are almost exclusively so-called parity sorters (Fig. 3). Other setups that are not equivalent to parity sorters seem not to be beneficial in finding a $(3, 3, 3)$ state. As we show later, the PS agent tends to use parity sorters more frequently while exploring the space of different SRV classes. This is particularly surprising since the parity sorter itself was originally designed for a different task.

Designing New Experiments. The connection between high-dimensional entangled states and the structure of optical tables which generate them is not well understood (12). Having a database of such experiments would allow us to deepen our understanding of the structure of the set of entangled states that can be accessed by optical tables. In particular, such a database could then be further analyzed to identify useful subsetups or gadgets (12)—certain few-element combinations that appear frequently in larger setups—which are useful for generating complex states. With our second task, we have challenged the agent to generate such a database by finding high-dimensional entangled states. As a bonus, we found that the agent does the postprocessing above for us implicitly and in runtime. The outcome of this postprocessing is encoded in the structure of the memory network of the PS agent. Specifically, the subsetups which were particularly useful in solving this task are clearly visible, or embodied, in the agent's memory.

To find as many different high-dimensional three-photon entangled states as possible, we reward the agent for every new implementation of an interesting experiment. To avoid trivial extensions of such implementations, a reward is given only if the obtained SRV was not reached before within the same experiment. Fig. 2B displays the total number of new, interesting
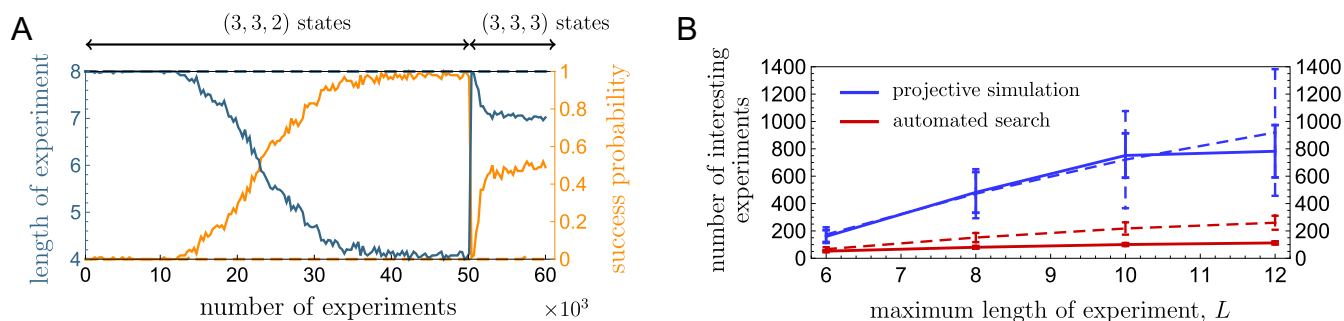
Fig. 2. Results of learning new experiments. (A) Average length of experiment and success probability in each of the 6×10^4 experiments. The maximal length of an experiment is $L = 8$. During the first 5×10^4 experiments an agent is rewarded for obtaining a $(3, 3, 2)$ state, and during the last 10^4 experiments the same agent is rewarded when finding a $(3, 3, 3)$ state. The average success probability shows how probable it is for the PS agent to find a rewarded state in a given experiment. Solid/dashed lines show simulations of the PS agent that learns how to generate a $(3, 3, 3)$ state from the beginning with/without prior learning of setups that produce a $(3, 3, 2)$ state. (B) Average number of interesting experiments obtained after placing 1.2×10^5 optical elements. Data points are shown for $L = 6, 8, 10,$ and 12. Dashed/solid blue and red lines correspond to PS with/without action composition and automated random search (12) with/without action composition, respectively (main text). Vertical bars indicate the mean-squared deviation. (A and B) All data points are obtained by averaging over 100 agents. Parameters of the PS agents are specified in *Parameters of the PS Agents*.

Fig. 3. Experimental setups frequently used by the PS agent. (*A*) Local parity sorter. (*B*) Nonlocal parity sorter (as discovered by the program). (*C*) Nonlocal parity sorter in the Klyshko wave front picture (53), in which the paths *a* and *d* are identical to the paths *b* and *c*, respectively. (*D*) Setup to increase dimensionality of photons. (*A*–*D*) In a simulation of 100 agents, the highest-weighted subsetups were 11 times experiment *A*, 22 times experiment *B*, and 43 times experiment *D* was part of the highest-weighted subsetup. Only in 24 cases were other subsetups the highest weighted.

experiments designed by the basic PS agent (solid blue curve) and the PS agent with action composition (19) (dashed blue curve). Action composition allows the agent to construct new composite actions from useful optical setups (i.e., placing multiple elements in a fixed configuration), thereby autonomously enhancing the toolbox (see *Projective Simulation* for details). It is a central ingredient for an AI to exhibit even a primitive notion of creativity (50) and was also used in ref. 12 to augment automated random search. For comparison, we provide the total number of interesting experiments obtained by automated random search with and without action composition (Fig. 2*B*, solid and dashed red curves). As we will see later, action composition will allow for additional insight into the agent's behavior and helps provide useful information about quantum optical setups in general. We found that the PS model discovers significantly more interesting experiments than both automated random search and automated random search with action composition (Fig. 2*B*).

Ingredients for Successful Learning. In general, successful learning relies on a structure hidden in the task environment (or dataset). The results presented thus far show that PS is highly successful

in the task of designing new interesting experiments, and here we elucidate why this should be the case. The following analysis also sheds light on other settings where we can be confident that RL techniques can be applied as well.

First, the space of optical setups can be illustrated using a graph as given in Fig. 4*C*, where the building of an optical experiment corresponds to a walk on the directed graph. Note that optical setups that create a certain state are not unique: Two or more different setups can generate the same quantum state. Due to this fact, this graph does not have a tree structure but rather resembles a maze. Navigating in a maze, in turn, constitutes one of the classic textbook RL problems (10, 41, 48, 49). Second, our empirical analysis suggests that experiments generating high-dimensional multipartite entanglement tend to have some structural similarities (12) (Fig. 4 *A* and *B* partially displays the exploration space). Fig. 4 shows regions where the density of interesting experiments (large colored nodes) is high and others where it is low—interesting experiments seem to be clustered (Fig. S2). In turn, RL is particularly useful when one needs to handle situations which are similar to those previously encountered—once one maze (optical experiment) is learned, similar mazes (experiments) are tackled more easily, as we have seen before. In other words, whenever the experimental task has a maze-type underlying structure, which is often the case, PS can likely help—and critically, without having any a priori information about the structure itself (41, 51). In fact, PS gathers information about the underlying structure throughout the learning process. This information can then be extracted by an external user or potentially be used further by the agent itself.

The Potential of Learning from Experiments. Thus far, we have established that a machine can indeed design new quantum experiments in the setting where the task is precisely specified (via the rewarding rule). Intuitively, this could be considered the limit of what a machine can do for us, as machines are specified by our programs. However, this falls short from what, for instance, a human researcher can achieve. How could we, even in principle, design a machine to do something (interesting) we have not specified it to do? To develop an intuition for the type of behavior we could hope for, consider, for the moment, what we may expect a human, say a good PhD student, would do in situations similar to those studied thus far.

Fig. 4. Exploration space of optical setups. Different setups are represented by vertices with colors specifying an associated SRV [biseparable states are depicted in blue]. Arrows represent the placing of optical elements. (*A*) A randomly generated space of optical setups. Here we allow up to 6 elements on the optical table and a standard toolbox of 30 elements. Large, colored vertices represent interesting experiments. If two nodes share a color, they can generate a state with the same SRV. Running for 1.6×10^4 experiments, the graph that is shown here has 45,605 nodes, of which 67 represent interesting setups. (*B*) A part of graph *A*, which demonstrates the nontrivial structure of the network of optical setups. (*C*) A detailed view of one part of the bigger network. The depicted colored maze represents an analogy between the task of finding the shortest implementation of an experiment and the task of navigating in a maze (10, 41, 48, 49). Arrows of different colors represent distinct optical elements that are placed in the experiment. The initial state is represented by an empty table $\varnothing$. The shortest path to a setup that produces a state with SRV (3, 3, 2) and (3, 3, 3) is highlighted. Labels along this path coincide with the labels of the percept clips in Fig. 1*B*.

To begin with, a human cannot go through all conceivable optical setups to find those that are interesting. Arguably, she would try to identify prominent subsetups and techniques that are helpful in the solving of the problem. Furthermore, she would learn that such techniques are probably useful beyond the specified tasks and may provide new insights in other contexts. Could a machine, even in principle, have such insight? Arguably, traces of this can be found already in our, comparatively simple, learning agent. By analyzing the memory network of the agent (ranking clips according to the sum of the weights of their incident edges), specifically the composite actions it learned to generate, we can extract subsetups that have been particularly useful in the endeavor of finding many different interesting experiments in Fig. 2B for $L = 6$.

For example, PS composes and then extensively uses a combination corresponding to an optical interferometer as displayed in Fig. 3A which is usually used to sort OAM modes with different parities (52)—in essence, the agent (re)discovered a parity sorter. This interferometer has already been identified as an essential part of many quantum experiments that create high-dimensional multipartite entanglement (12), especially those involving more than two photons (Fig. 2A and ref. 13).

One of the PS agent's assignments was to discover as many different interesting experiments as possible. In the process of learning, even if a parity sorter was often (implicitly) rewarded, over time it will no longer be, as in this scenario only novel experiments are rewarded. This, again implicitly, drives the agent to "invent" new, different-looking configurations, which are similarly useful.

Indeed, in many instances the most rewarded action is no longer the original parity sorter in the form of a Mach–Zehnder interferometer (Fig. 3A) but a nonlocal version thereof (Fig. 3B). As it turns out, the two setups are equivalent in the Klyshko wave front picture (53, 54), where the time of two photons in arms b and c in Fig. 3B is inverted, and these photons are considered to be reflected at their origin (represented by Refl_x)—the nonlinear crystals. This transformation results in a reduction of the four-photon experiment in Fig. 3B to the two-photon experiment shown in Fig. 3C.

Such a nonlocal interferometer has only recently been analyzed and motivated the work in ref. 17. Furthermore, by slightly changing the analyzer to reward only interesting experiments that produce states with SRV other than the most common SRV $(3, 3, 2)$, subsetups aside from the OAM parity sorter become apparent. For example, the PS discovered and exploited a technique to increase the dimensionality of the OAM of the photons by shifting the OAM mode number in one path and subsequently mixing it with an additional path as displayed in Fig. 3D. Moreover, one can observe that this technique is frequently combined with a local OAM parity sorter. This setup allows the creation of high-dimensional entangled states beyond the initial state dimension of 3.

All of the observed setups are, in fact, modern quantum optical gadgets/devices (designed by humans) that either already found applications in state-of-the-art quantum experiments (13) or could be used (as individual tools) in future experiments which create high-dimensional entanglement starting from lower-dimensional entanglement.

Discussion

One of the fastest-growing trends in recent times is the development of "smart" technologies. Such technologies are not only permeating our everyday lives in the forms of smart phones, smart watches, and in some places even smart self-driving cars, but are expected to induce the next industrial revolution (55). Hence, it should not come as a surprise when "smart laboratories" emerge. Already now, modern laboratories are to a large extent automated (56) and are removing the need for human involvement in tedious (or hazardous) tasks.

In this work we broach the question of the potential of automated laboratories, trying to understand to what extent machines could not only help in research, but perhaps also even genuinely perform it. Our approach highlights two aspects of learning machines, both of which will be assets in quantum experiments of the future. First, we have improved upon the original observation that search algorithms can aid in the context of finding special optical setups (12) by using more sophisticated learning agents. This yields confidence that even more involved techniques from AI research [e.g., generalization (43), meta-learning (51), etc., in the context of the PS framework] may yield ever-improving methods for the autonomous design of experiments. In a complementary direction, we have shown that the structure of learning models commonly applied in the context of AI research [even the modest basic PS reinforcement learning machinery augmented with action-clip composition (19, 40)] possibly allows machines to tackle problems they were not directly instructed or trained to solve. This supports the expectation that AI methodologies will genuinely contribute to research and, very optimistically, the expectation that they will contribute to the discovery of new physics.

1. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Advances of Neural Information Processing Systems*, eds Pereira F, Burges CJC, Bottou L, Weinberger KQ (Curran Associates, Red Hook, NY), Vol 25, pp 1097–1105.
2. Silver D, et al. (2016) Mastering the game of go with deep neural networks and tree search. *Nature* 529:484–489.
3. Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. *Science* 349:255–260.
4. Carrasquilla J, Melko RG (2017) Machine learning phases of matter. *Nat Phys* 13:431–434.
5. van Nieuwenburg EPL, Liu YH, Huber SD (2017) Learning phase transitions by confusion. *Nat Phys* 13:435–439.
6. Schmidt M, Lipson H (2009) Distilling free-form natural laws from experimental data. *Science* 324:81–85.
7. Carleo G, Troyer M (2017) Solving the quantum many-body problem with artificial neural networks. *Science* 355:602–606.
8. Zdeborová L (2017) Machine learning: New tool in the box. *Nat Phys* 13:420–421.
9. Russel S, Norvig P (2010) *Artificial Intelligence–A Modern Approach* (Prentice Hall, Upper Saddle River, NJ), 3rd Ed.
10. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT press, Cambridge, MA).
11. Mnih V, et al. (2015) Human-level control through deep reinforcement learning. *Nature* 518:529–533.
12. Krenn M, Malik M, Fickler R, Lapkiewicz R, Zeilinger A (2016) Automated search for new quantum experiments. *Phys Rev Lett* 116:090405.
13. Malik M, et al. (2016) Multi-photon entanglement in high dimensions. *Nat Photon* 10:248–252.
14. Schlederer F, Krenn M, Fickler R, Malik M, Zeilinger A (2016) Cyclic transformation of orbital angular momentum modes. *New J Phys* 18:043019.
15. Babazadeh A, et al. (2017) High-dimensional single-photon quantum gates: Concepts and experiments. *Phys Rev Lett* 119:180510.
16. Erhard M, Malik M, Krenn M, Zeilinger A (2017) Experimental GHZ entanglement beyond qubits. arXiv:1708.03881.
17. Krenn M, Hochrainer A, Lahiri M, Zeilinger A (2017) Entanglement by path identity. *Phys Rev Lett* 118:080401.
18. Krenn M, Gu X, Zeilinger A (2017) Quantum experiments and graphs: Multiparty states as coherent superpositions of perfect matchings. *Phys Rev Lett* 119:240403.
19. Briegel HJ, De las Cuevas G (2012) Projective simulation for artificial intelligence. *Sci Rep* 2:400.

PHYSICS

20. Lawrence J (2014) Rotational covariance and Greenberger-Horne-Zeilinger theorems for three or more particles of any dimension. *Phys Rev A* 89:012105.
21. Huber M, de Vicente JI (2013) Structure of multidimensional entanglement in multipartite systems. *Phys Rev Lett* 110:030501.
22. Huber M, Perarnau-Llobet M, de Vicente JI (2013) Entropy vector formalism and the structure of multidimensional entanglement in multipartite systems. *Phys Rev A* 88:042328.
23. Horodecki R, Horodecki P, Horodecki M, Horodecki K (2009) Quantum entanglement. *Rev Mod Phys* 81:865–942.
24. Amico L, Fazio R, Osterloh A, Vedral V (2008) Entanglement in many-body systems. *Rev Mod Phys* 80:517–576.
25. Pan JW, et al. (2012) Multiphoton entanglement and interferometry. *Rev Mod Phys* 84:777–838.
26. Hein M, Eisert J, Briegel HJ (2004) Multiparty entanglement in graph states. *Phys Rev A* 69:062311.
27. Scarani V, Gisin N (2001) Quantum communication between *N* partners and Bell's inequalities. *Phys Rev Lett* 87:117901.
28. Scarani V, et al. (2009) The security of practical quantum key distribution. *Rev Mod Phys* 81:1301–1350.
29. Allen L, Beijersbergen MW, Spreeuw R, Woerdman J (1992) Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes. *Phys Rev A* 45:8185–8189.
30. Mair A, Vaziri A, Weihs G, Zeilinger A (2001) Entanglement of the orbital angular momentum states of photons. *Nature* 412:313–316.
31. Molina-Terriza G, Torres JP, Torner L (2007) Twisted photons. *Nat Phys* 3:305–310.
32. Krenn M, Malik M, Erhard M, Zeilinger A (2017) Orbital angular momentum of photons and the entanglement of Laguerre–Gaussian modes. *Philos Trans R Soc A* 375:20150442.
33. Vaziri A, Weihs G, Zeilinger A (2002) Experimental two-photon, three-dimensional entanglement for quantum communication. *Phys Rev Lett* 89:240401.
34. Dada AC, Leach J, Buller GS, Padgett MJ, Andersson E (2011) Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities. *Nat Phys* 7:677–680.
35. Agnew M, Leach J, McLaren M, Roux FS, Boyd RW (2011) Tomography of the quantum state of photons entangled in high dimensions. *Phys Rev A* 84:062101.
36. Krenn M, et al. (2014) Generation and confirmation of a (100× 100)-dimensional entangled quantum system. *Proc Natl Acad Sci USA* 111:6243–6247.
37. Zhang Y, et al. (2016) Engineering two-photon high-dimensional states through quantum interference. *Sci Adv* 2:e1501165.
38. Hiesmayr B, de Dood M, Löffler W (2016) Observation of four-photon orbital angular momentum entanglement. *Phys Rev Lett* 116:073601.
39. Wang XL, et al. (2015) Quantum teleportation of multiple degrees of freedom of a single photon. *Nature* 518:516–519.
40. Mautner J, Makmal A, Manzano D, Tiersch M, Briegel HJ (2015) Projective simulation for classical learning agents: A comprehensive investigation. *New Generat Comput* 33:69–114.
41. Melnikov AA, Makmal A, Briegel HJ (2014) Projective simulation applied to the grid-world and the mountain-car problem. arXiv:1405.5459.
42. Bjerland ØF (2015) Projective Simulation compared to reinforcement learning. Master's thesis (Department of Computer Science, University of Bergen, Bergen, Norway).
43. Melnikov AA, Makmal A, Dunjko V, Briegel HJ (2017) Projective simulation with generalization. *Sci Rep* 7:14430.
44. Hangl S, Ugur E, Szedmak S, Piater J (2016) Robotic playing for hierarchical complex skill learning. *Proc IEEE/RSJ Int Conf Intell Robots Syst* (IEEE, New York), pp 2799–2804.
45. Paparo GD, Dunjko V, Makmal A, Martin-Delgado MA, Briegel HJ (2014) Quantum speed-up for active learning agents. *Phys Rev X* 4:031002.
46. Dunjko V, Friis N, Briegel HJ (2015) Quantum-enhanced deliberation of learning agents using trapped ions. *New J Phys* 17:023006.
47. Friis N, Melnikov AA, Kirchmair G, Briegel HJ (2015) Coherent controlization using superconducting qubits. *Sci Rep* 5:18036.
48. Mirowski P, et al. (2016) Learning to navigate in complex environments. arXiv: 1611.03673.
49. Mannucci T, van Kampen EJ (2016) A hierarchical maze navigation algorithm with reinforcement learning and mapping. *Proc IEEE Symp Ser Comput Intelligence* (IEEE, New York), pp 1–8.
50. Briegel HJ (2012) On creative machines and the physical origins of freedom. *Sci Rep* 2:522.
51. Makmal A, Melnikov AA, Dunjko V, Briegel HJ (2016) Meta-learning within projective simulation. *IEEE Access* 4:2110–2122.
52. Leach J, Padgett MJ, Barnett SM, Franke-Arnold S, Courtial J (2002) Measuring the orbital angular momentum of a single photon. *Phys Rev Lett* 88: 257901.
53. Klyshko DN (1988) A simple method of preparing pure states of an optical field, of implementing the Einstein–Podolsky–Rosen experiment, and of demonstrating the complementarity principle. *Sov Phys Usp* 31:74–85.
54. Aspden RS, Tasca DS, Forbes A, Boyd RW, Padgett MJ (2014) Experimental demonstration of Klyshko's advanced-wave picture using a coincidence-count based, camera-enabled imaging system. *J Mod Opt* 61:547–551.
55. Schwab K (2017) *The Fourth Industrial Revolution* (Penguin, London).
56. King RD, et al. (2009) The automation of science. *Science* 324:85–89.