Scientific intuition inspired by machine learning-generated hypotheses

View the article online for updates and enhancements.

MACHINE LEARNING
Science and Technology

Scientific intuition inspired by machine learning-generated hypotheses

Pascal Friederich[1,2,3,4] (ID), **Mario Krenn**[1,2,5] (ID), **Isaac Tamblyn**[5,6] (ID) and **Alán Aspuru-Guzik**[1,2,5,7] (ID)

1   Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Canada
2   Department of Computer Science, University of Toronto, Toronto, Canada
3   Institute of Theoretical Informatics, Karlsruhe Institute of Technology, Am Fasanengarten 5, 76131 Karlsruhe, Germany
4   Institute of Nanotechnology, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
5   Vector Institute for Artificial Intelligence, Toronto, Canada
6   National Research Council of Canada, Ottawa, Canada
7   Canadian Institute for Advanced Research (CIFAR) Lebovic Fellow, Toronto, Canada

E-mail: pascal.friederich@kit.edu and alan@aspuru.com

Abstract

Machine learning with application to questions in the physical sciences has become a widely used tool, successfully applied to classification, regression and optimization tasks in many areas. Research focus mostly lies in improving the accuracy of the machine learning models in numerical predictions, while scientific understanding is still almost exclusively generated by human researchers analysing numerical results and drawing conclusions. In this work, we shift the focus on the insights and the knowledge obtained by the machine learning models themselves. In particular, we study how it can be extracted and used to inspire human scientists to increase their intuitions and understanding of natural systems. We apply gradient boosting in decision trees to extract human-interpretable insights from big data sets from chemistry and physics. In chemistry, we not only rediscover widely know rules of thumb but also find new interesting motifs that tell us how to control solubility and energy levels of organic molecules. At the same time, in quantum physics, we gain new understanding on experiments for quantum entanglement. The ability to go beyond numerics and to enter the realm of scientific insight and hypothesis generation opens the door to use machine learning to accelerate the discovery of conceptual understanding in some of the most challenging domains of science.

1. Introduction

Machine learning (ML) recently has become a widely used tool with many applications in the physical sciences [1], ranging from chemistry (for example, prediction of quantum chemistry properties [2], solving Schrödinger's equation [3], predicting reactions [4], materials discovery [5] or inverse materials design [6, 7]) to physics (for example, identification of phases of matter [8], astronomical object recognition [9], or validation of quantum experiments [10]) and biology (for example, prediction of protein structures [11] or drug design [12, 13]). Some open challenges regarding the application of machine learning models in natural sciences include the accessibility, homogeneity, amount and quality of available data, as well as a lack of machine learning models which inherently include physical laws, limiting the interpretability of the models' predictions. While ML models are successfully used and optimized to accelerate numerical predictions or to recognize or generate patterns in existing data, it is rarely inquired how the machine finds solutions, i.e. which patterns and correlations it detected and exploited. Thus, the scientific insight obtained by the model is not directly transferred to human scientists. First attempts to use artificial intelligence in physical sciences aimed to directly answer scientific questions, e.g. determine the location of protein encodings in the genome [14]. Further attempts to employ machine learning models to obtain insight and

Figure 1. Workflow for automated hypothesis generation. (a) General workflow, starting with a database of graphs and respective properties, followed by training of a machine learning model that allows for the extraction of feature importances, e.g. gradient boosting regression. Features with high importance are combined and analysed in a way that facilitates interpretation in order to stimulate scientific insight. (b) Schematic illustration of the gradient boosting regression method [19], where multiple simple decision tree models are trained sequentially. Each new decision tree is trained to correct the residual errors (red lines) of the previous models, so the final prediction $F_0(x)$ can be written as a sum of the mean label c_0 and a weighted series of models $h_i(x)$, where each h_i predicts the deviation of the previous $i - 1$ models from the ground truth. (c) Each decision tree is trained on samples that are represented using predefined input features (coloured squares) and uses their values to split the data set sequentially into smaller subsets which are used for the predictions. The subgraph based input representation used in this work allows a direct interpretation of the feature importances (d) that are computed based on a quantification of how meaningful features are for the accuracy of the machine learning model.

help scientists to develop theories were focused on rediscovering solutions to already solved problems, e.g. to rediscover the coordinate transformation in astrophysical [15] and non-linear dynamical systems [16], or to detect symmetries and conservation laws [17]. The methods used in these cases enforce information bottlenecks or interpretable transformations in the ML model that then can inspire scientific understanding [18]. However, to our knowledge such methods were mostly applied to solved problems and have not been used yet to obtain novel insight and answers to questions that are not well understood yet.

In this work, we propose to use machine learning and systematic data analysis to automate further the process of generation of *interpretable* scientific hypotheses. We demonstrate the applicability of the approach using two questions in the natural sciences — a rediscovery task of chemistry knowledge (hydrophobicity and molecular energy levels in simple as well as application relevant molecules) and the discovery of new intuitions in physics (quantum optics). We show that our approach 'rediscovers' but also extends known chemical rules of thumb for solubility and energy levels of organic molecules with application in organic photovoltaics and organic light-emitting diodes, and helps us to better understand the entanglement created in quantum optical experiments.

Our model represents its findings in a graph representation which is directly related to chemical or physical instances in the specific scientific domain. The results are statements regarding distinct subgraphs that can easily be comprehended and therefore, scientifically interpreted and understood by experts. This is in stark contrast to conventional machine learning models where the internal representations are only indirectly connected with the real physical entities and thus hard to impossible to interpret.

2. Method

2.1. Computer generated hypotheses
We suggest an automated workflow for ML-based generation of human-interpretable scientific hypotheses as illustrated in figure 1(a). The workflow is based on a reference database of calculated (potentially also measured) data points with graph-based structure and with corresponding target properties. A binary feature vector describing presence/absence of automatically generated subgraphs [20] is used to train a tree ensemble method, e.g. gradient boosting [19] or random forest regression/classification [21, 22], that allows for the quantification of feature importances. Based on the features with the highest importance, a list of hypotheses is generated. Each hypothesis has the human understandable form

Feature *i* leads to an increase/decrease of the target property with strength *s*

where i is the index of the corresponding feature (subgraph) in the input and strength s quantifies the degree of correlation between feature i and the target property. High feature importance does not necessarily correspond to a high direct correlation with the target feature. In many cases, multiple features have to be combined in order to become predictive, even if the single features individually do not help in the predicting the target property. Therefore, important features are combined using logical operations (*and*, *xor*,…) to automatically generate combined features which, especially in presence of higher-order correlations, can be directly interpreted by researchers.

2.2. Input representation and experiments

In this work, we test this workflow on two experiments in chemistry and physics. The first experiment targets the automated generation of intuitive rules that determine molecular properties, whereas the second aims at hypothesis generation for entanglement properties of quantum optical experiments. In both cases, we can describe the data points as graphs (molecules and quantum optical experiments), where nodes are chemical elements or optical instruments while edges are chemical bonds or photon paths travelling through the setup. This allows us to use fingerprinting techniques to generate input representations (bit-vectors), e.g. using the algorithm for circular extended-connectivity fingerprints [20]. This iterative algorithm generates a unique representation of each node, including its local environment. In each iteration, hashing functions are used to aggregate the information (predefined node and edge features) of the next nearest neighbors of each node, thus implicitly integrating information of one additional neighbor shell in each iteration. In the end, a hashing function is used to map all subgraphs found in the graphs to bit-vectors. Each entry in these bit-vectors encodes the presence or absence of a certain subgraph. A similar approach has been used in Lopez *et al* [23] to determine molecular substructures in molecules for organic solar cells that lead to high power conversion efficiencies. Other models that link the presence of subgraphs (or more generally features) in the input data to properties can potentially be employed in our workflow (see e.g. Duvenaud *et al* [24] where molecular fragments are identified that correlate with toxicity, the Grad-CAM method by Selvaraju *et al* [25] for convolutional neural networks or the GNNExplainer by Ying *et al* [26]). In contrast to this work, some of these approaches depend on the analysis of single samples and thus only indirectly allow to conclude about an entire data set. Furthermore, these approaches assign importance indicators to single nodes or edges of a graph, which are not necessarily binary numbers, which complicates the direct interpretation. Due to their general applicability to all graphs where node and edges can be represented by one or multiple categorical features, we focused on automatically generated circular fingerprints in this work.
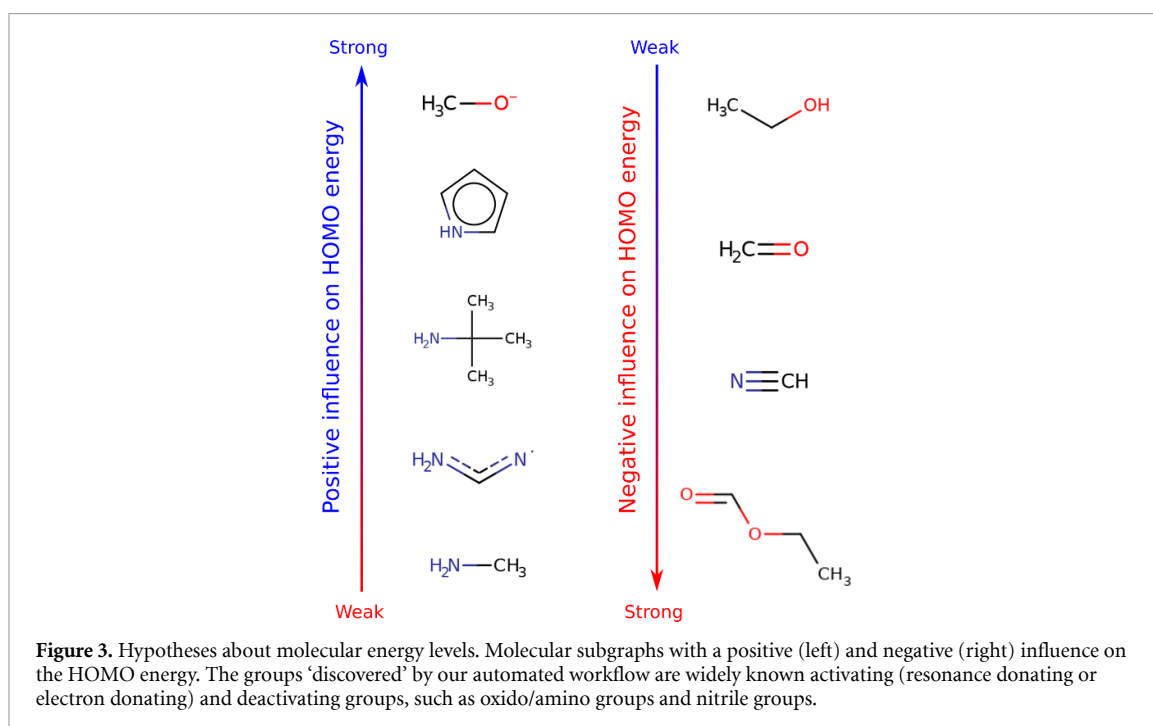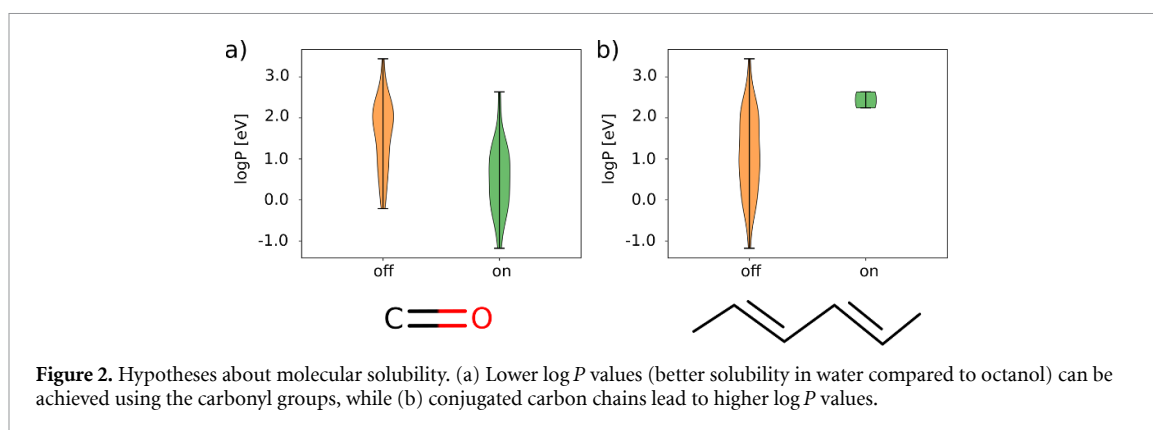
3. Results

To test the automated hypothesis generation workflow, we performed experiments in two scientific domains, molecular chemistry (section 3.1) and quantum optical experiments (section 3.2). We computed physical properties of these graphs and used the generated data sets and the workflow described in figure 1 to automatically generate hypotheses that can be either compared to a collection of widely known chemical rules of thumb or that can help to better understand entanglement in quantum optical experiments for designing future experiments.

3.1. Chemical intuition for solubility, energy levels

In case of the chemistry experiment, we used two prototypical target properties — the water–octanol partition coefficient which describes the solubility of molecules in water (polar) vs. octanol (non-polar) as well as the energy of the highest occupied molecular orbital. Both properties are of high relevance for the application of molecules as pharmaceuticals or in electronic devices, e.g. for organic solar cells, organic light-emitting diode (OLED) displays or organic flow batteries. We furthermore analysed existing application-specific data sets, namely a data set of thermally activated delayed fluorescent (TADF) molecules as emitter molecules for OLEDs [27], the Harvard Clean Energy project data set [28, 29] and a data set of non-fullerene acceptor molecules for organic solar cells [23]. Solubility and energy levels are relatively well understood and for both properties there exist several widely known rules of thumb, often described as chemical intuition, which describes how certain functional groups influence them. Our experiment aims to test whether the automated hypothesis generation method can 'rediscover' those rules and potentially add new or refined rules. For frontier orbital gaps reported in the Harvard Clean Energy data set and the non-fullerene acceptor data set as well as for singlet-triplet energy splittings reported in the TADF data set, there exists less chemical intuition on how to influence and tune them.
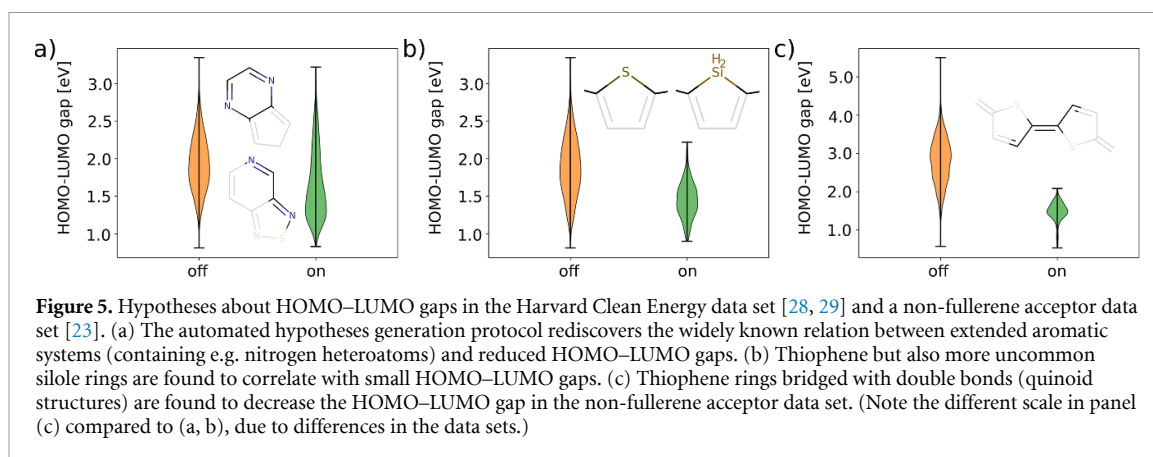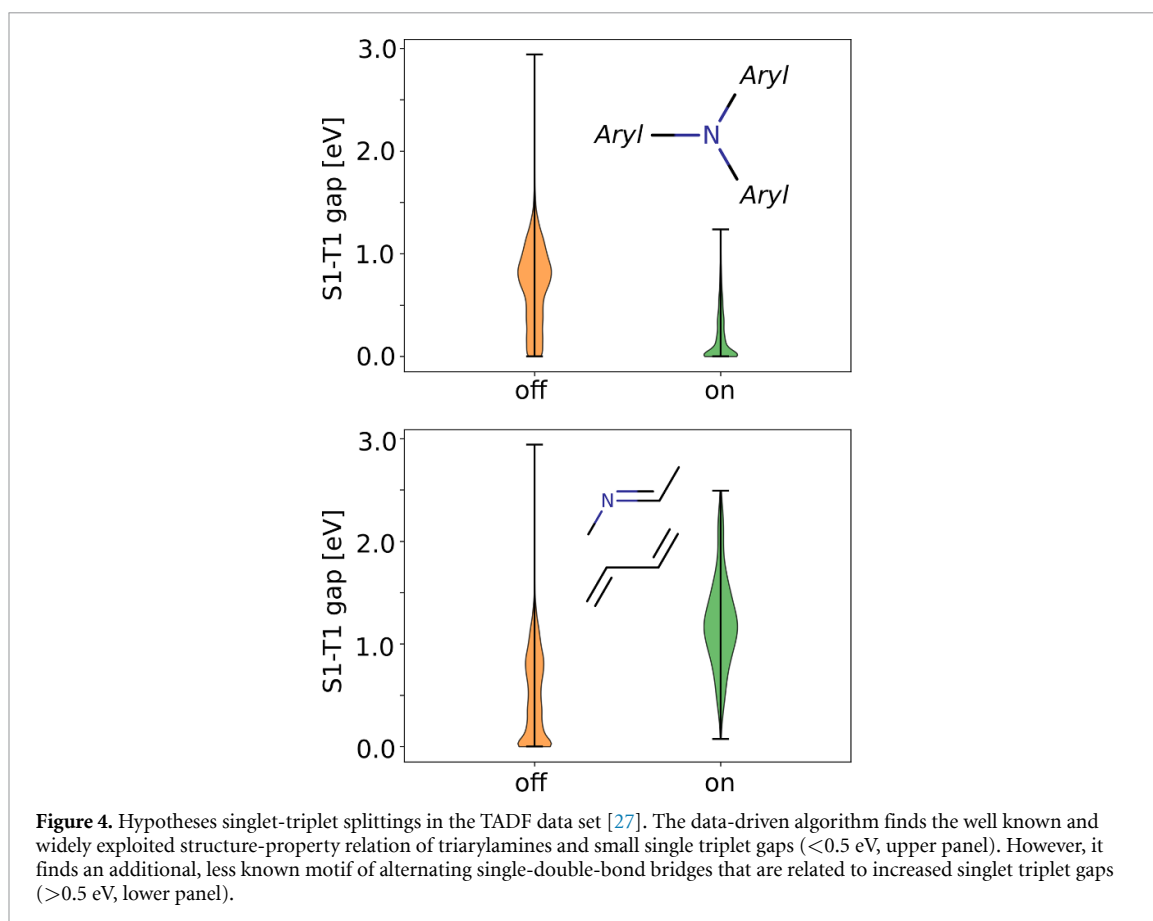
Figure 2 shows two solubility related hypotheses that were generated using our workflow. Without prior knowledge, the algorithm predicts two widely known chemical groups/motifs for increasing solubility in polar solvents (carbonyl group in figure 2(a) and to increase solubility in non-polar solvents (conjugated

Figure 2. Hypotheses about molecular solubility. (a) Lower log P values (better solubility in water compared to octanol) can be achieved using the carbonyl groups, while (b) conjugated carbon chains lead to higher log P values.

Figure 3. Hypotheses about molecular energy levels. Molecular subgraphs with a positive (left) and negative (right) influence on the HOMO energy. The groups 'discovered' by our automated workflow are widely known activating (resonance donating or electron donating) and deactivating groups, such as oxido/amino groups and nitrile groups.

carbon chain in figure 2(b)). Figure 3 shows an overview of molecular subgraphs that positively and negatively influence the HOMO energy of a molecule. To our surprise, five of the nine groups shown in the figure can directly be found in chemistry textbooks or Wikipedia when searching for electrophilic aromatic directing groups which can change the energy levels of molecules through the inductive effect and the mesomeric effect. Specifically, the oxido (O^-) group that shows the strongest positive influence on the HOMO is well known for a strong resonance donating and a strong inductive effect which both leads to an increase in HOMO energy. Furthermore, heterocycles that contain nitrogen, as well as amine (NH_2) groups are also known for lifting the HOMO level to higher energies. On the other hand, the nitrile group ($C{\equiv}N$) is one of the most widely known electron-withdrawing groups that lowers the HOMO energy of molecules due to its resonance withdrawing and inductively withdrawing nature.

The patterns found to be relevant for small HOMO–LUMO gaps in the Harvard Clean Energy data set as well as in the non-fullerene acceptor data set are mostly related to extended aromatic systems and fused aromatic rings (see figures 5(a) and S1(a) (available online at stacks.iop.org/MLST/2/025027/mmedia)). This finding is well-understood by chemists due to the widely known relation between the size of an aromatic system (i.e. the degree of delocalization of π-electrons) and the frontier orbital gap [30]. In the limit of infinite delocalization (e.g. in graphene), the HOMO–LUMO gap closes completely. This relation was also exploited in the development of conductive polymers, which was awarded with the Nobel Price in Chemistry in 2000 and which created the field of organic electronics [31].
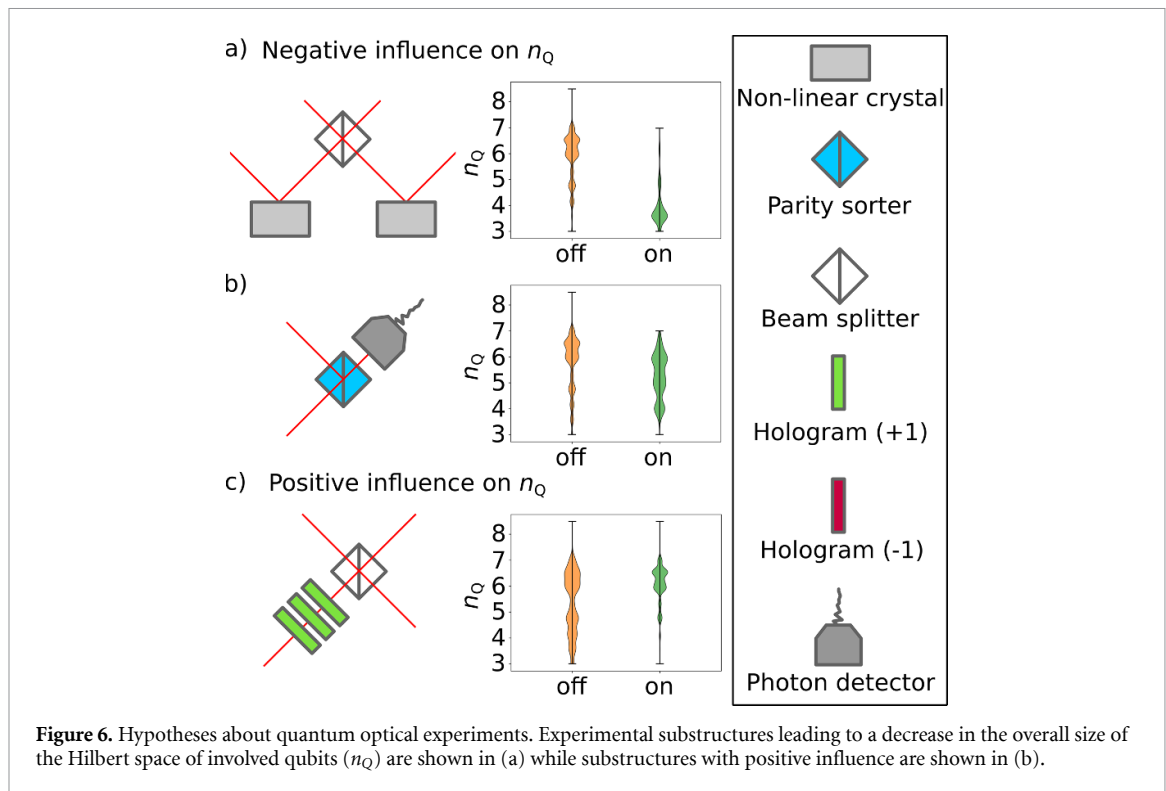
However, we additionally found several interesting and surprising patterns both in the photovoltaic data sets (figures 5(b) and (c)) and in the TADF dataset (figure 4). In case of the Harvard Clean Energy data set, we find that aromatic heterocycles with sulfur (e.g. thiophene rings) as well as silicon heteroatoms (e.g. silole

Figure 4. Hypotheses singlet-triplet splittings in the TADF data set [27]. The data-driven algorithm finds the well known and widely exploited structure-property relation of triarylamines and small single triplet gaps (<0.5 eV, upper panel). However, it finds an additional, less known motif of alternating single-double-bond bridges that are related to increased singlet triplet gaps (>0.5 eV, lower panel).

Figure 5. Hypotheses about HOMO–LUMO gaps in the Harvard Clean Energy data set [28, 29] and a non-fullerene acceptor data set [23]. (a) The automated hypotheses generation protocol rediscovers the widely known relation between extended aromatic systems (containing e.g. nitrogen heteroatoms) and reduced HOMO–LUMO gaps. (b) Thiophene but also more uncommon silole rings are found to correlate with small HOMO–LUMO gaps. (c) Thiophene rings bridged with double bonds (quinoid structures) are found to decrease the HOMO–LUMO gap in the non-fullerene acceptor data set. (Note the different scale in panel (c) compared to (a, b), due to differences in the data sets.)

rings) significantly reduce the HOMO–LUMO gap. While the former are widely used in organic electronics to control energy levels and reduce HOMO–LUMO gaps, silole rings are more unusual.

In the non-fullerene acceptor data set (see figure 5(c)) we found that thiophene rings connected by double bonds (i.e. forming a quinoid structure instead of aromatic systems) also significantly reduce the HOMO–LUMO gap, which is a know relation first described by Brédas [32]. However, such systems require a specific functionalization in the periphery of the molecule to enforce the quinoid structure of the two thiophene rings, which intrinsically is less stable and thus higher in energy than the aromatic structure.

In case of the TADF data set (see figure 4), we found expected patterns such as triarylamines that correlate with decreased singlet triplet gaps (S1–T1 gaps) as well as rather unexpected patterns (e.g. conjugated bridges) that are identified by our workflow as chemical groups that highly correlate with large singlet triplet gaps. Low singlet–triplet splittings in TADF molecules are typically achieved by decoupling electron donating and electron accepting parts of a molecule to reduce the exchange interaction between the frontier orbitals which would otherwise lower the triplet state compared to the singlet state and open an undesired singlet–triplet splitting. The decoupling of the fragments can be achieved by introducing

Figure 6. Hypotheses about quantum optical experiments. Experimental substructures leading to a decrease in the overall size of the Hilbert space of involved qubits (n_Q) are shown in (a) while substructures with positive influence are shown in (b).

twist angles close to 90° between the fragments. One way to accomplish this are triarylamines bridges between the fragments. We expect that the conjugated bridges between fragments have precisely the opposite effect: They lead to a planar alignment of the adjacent fragments and thus an enhanced exchange interaction, reduced triplet energies and finally increased singlet–triplet splittings.

3.2. Physical intuitions for quantum experiments

As a second example, we use quantum optical experiments for producing high-dimensional, multipartite quantum entanglement [33, 34]. These experiments grow in interest as they allow the investigation of fundamental physical properties — such as local realism [35] — in laboratories. Furthermore, such quantum states are the key resources for large and complex quantum communication networks [36, 37], which are on the edge of commercial availability. The experimental setups that we consider consist of standard optical components that are used in labs, such as non-linear crystals for the creation of photon pairs, single-photon detectors, beam splitters, holograms or Dove prisms. Under approximations that are closely resembled in experiments, the final emergent quantum state can be reliably calculated [38].
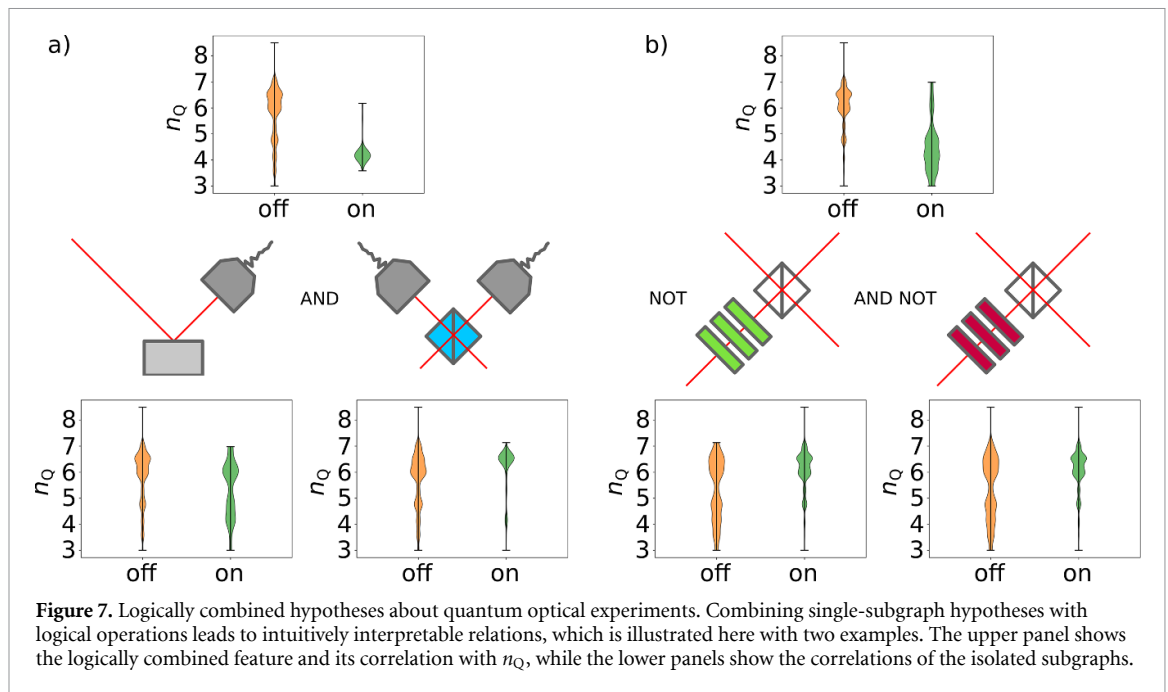
A key challenge lies in the design of experiments which creates certain desired quantum systems. The difficulty arises from counter-intuitive quantum phenomena, which raises the question of whether human intuition is the best way to design new experiments. Several studies have therefore developed automated and machine-learning augmented approaches for the design of experiments [39–44]. The goal in our approach is to tackle this challenge in a completely different way, namely by improving the scientist's intuition about these systems.

Specifically, we are investigating optical setups with three-photon entanglement in high dimensions, using a fourth photon as a trigger. The experimental setups can be represented as graphs where vertices represent optical elements, and edges correspond to the photon paths connecting these elements. Analogously to chemical elements, the optical elements can have one to four connections. For example, a beam splitter has four input-output modes, while a detector has only one input. As a measure of entanglement, we use the overall size of the involved Hilbert space in terms of involved qubits, $n_Q = \log_2(d_1 d_2 d_3)$, where d_i stands for the rank of density matrix after tracing out photon i [45, 46].

We used the same fingerprint-based graph representation as in section 3.1 and trained a gradient boosting regression model to predict n_Q. Using the algorithm outlined in figure 1, we form a list of hypotheses of subgraphs features that influence n_Q most. This computer-generated list was analysed and interpreted by a domain expert.

The two features which influence n_Q most negatively contradict the intuition in the field, see figures 6(a), (b) and S2. Surprisingly, both of them represent subgraphs that are core elements of two experimental setups

Figure 7. Logically combined hypotheses about quantum optical experiments. Combining single-subgraph hypotheses with logical operations leads to intuitively interpretable relations, which is illustrated here with two examples. The upper panel shows the logically combined feature and its correlation with n_Q, while the lower panels show the correlations of the isolated subgraphs.

which have produced high-dimensional multipartite entanglement in the laboratory [47, 48]. Specifically, if the outputs of two non-linear crystals (both crystals produce entangled photon pairs in the same 3D mode space) are connected directly via a beam splitter or interferometer, the entanglement of the resulting state is predicted to be comparably low. This can be interpreted in the following way: the photons from the two different crystals need to combine at some point, otherwise, they remain bi-separable. However, if they combine directly after their generation, the equal mode spaces mix in such a way that it is difficult to increase their dimensionality subsequently. It is therefore explicitly enlightening that several of the features that positively influence n_Q correspond to elements which shift the entire mode space by plus or minus three before or after the beam splitters or non-linear crystals. The insight for a human researcher now is to shift the mode space by three (as the local dimension is three), before combining photons from different non-linear crystals to achieve a high n_Q. This leads to mode spaces of twice the original size and thereby increasing the probability for large overall entanglement dimensionalities.

A different feature which was used in the two experimentally demonstrations, but significantly negatively influences n_Q is the following: one output of a non-linear crystal is directly connected to the detector. For human designers, this leads to the convenient fact that it simplifies the initial state (as double-emissions from one crystal can be ignored in this case). However, the entanglement of this photon with the other two photons can never be larger than three (as the local mode space is three). A similar, negatively influencing feature is a certain interferometer, which sorts the parity of the involved modes, directly connected to a detector. This acts as a filter, thus reducing the mode space of the incoming photon by half, thereby reducing the overall possible entanglement significantly.

3.3. Logically combined features

We can logically combine graph features, as described in 2, and find the most significant macro-features for quantum experiments. In figure 7(a), two small sub-experiments are combined with a logical *and*, i.e. the feature is the combination of both structures. Individually, the presence of the first feature has a negative influence on n_Q. The second feature, a parity sorter followed by two detectors, influences n_Q positively. Surprisingly, their combination has a significant negative influence on n_Q and can be seen as an almost sufficient condition for $n_Q \approx 4$. This behaviour can be interpreted using the Klyshko advanced wave-picture for quantum correlations in quantum optics [49]. The detector after the photon pair creation heralds a specific quantum state in the other photonic path. If those photons deterministically split at the parity sorter, the ability to mix with the photons from the other input ports (thus from the other crystal) vanish. From this insight, the human designer can learn that a heralded single-photon should be combined in a probabilistic way with the photons of the other crystal, using beam splitters instead of parity sorters.

A second macro-feature, figure 7(b), combines two insights that we gained in figure 6. The macro-feature in figure 7(b) shows that the absence of either three positive or three negative mode shifters in front of a beam splitter has a very negative impact on the n_Q. Thereby, the algorithm has discovered that both

increasing or decreasing helps to have very positive influence on the final entanglement, and thereby suggests that one can be agnostic about the shift direction, and the importance lies in the actual increase of the local Hilbert space before the mixing. This features clearly shows how logical combinations can simplify the interpretation of scientific data.

4. Conclusion and outlook

We presented a data-driven machine learning workflow for automated generation and verification of hypotheses about observations in natural sciences. We presented examples from chemistry and physics, but our method is directly applicable to most applications, where structures can be represented as graphs, e.g. to DNA/RNA data in biology [50, 51], chemical reaction networks [52, 53] or graphs in social sciences. In chemistry, the workflow 'rediscovers' widely known relations regarding solubility and electronic properties of molecules (often referred to as chemical intuition). In physics, the algorithm discovers rules to generate highly entangled three-photon states in quantum optical experiments. These rules are interpretable by human experts in retrospect, yet not known or postulated before, and even contradicting some of the field's current understanding. Finding such rules will not only help researchers to understand complex scientific relationships and thus design better experiments, but also reduce unavoidable and often undetectable bias generated by prior knowledge and anticipations.

4.1. Hypothesis testing
In addition to automated hypothesis generation, protocols for testing of the postulated hypotheses would be beneficial. In case of the chemistry experiment, a possible hypothesis testing protocol would generate mutations of each molecule in the training set to test the hypotheses on molecules with similar representations, where (ideally) only the relevant feature is changed. In case of the quantum optical experiments, not all random mutations will lead to maximally entangled states between all photons, which is a requirement to compute the entanglement of the quantum state. We currently see two options for automated hypothesis verification both of which we are currently implementing. The first follows the same procedure of mutation and computation as in the chemistry experiment, with the caveat that only a small fraction of the mutations will lead to useful results, potentially making the procedure computationally costly. The second option is based on finding other experimental setups within the whole database that are as similar to the reference experiment as possible, with the exception of the feature that is currently analysed. This procedure is computationally costly as well but does not require new computations.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

ORCID iDs

Pascal Friederich ⬤ https://orcid.org/0000-0003-4465-1465
Mario Krenn ⬤ https://orcid.org/0000-0003-1620-9207
Isaac Tamblyn ⬤ https://orcid.org/0000-0002-8146-6667
Alán Aspuru-Guzik ⬤ https://orcid.org/0000-0002-8277-4434

References

[1] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L and Zdeborová L 2019 Machine learning and the physical sciences *Rev. Mod. Phys.* **91** 045002
[2] Ramakrishnan R, Dral P O, Rupp M and Anatole von Lilienfeld O 2015 Big data meets quantum chemistry approximations: the δ-machine learning approach *J. Chem. Theory Comput.* **11** 2087–96
[3] Hermann J, Schätzle Z and Noé F 2020 Deep-neural-network solution of the electronic Schrödinger equation *Nat. Chem.* **12** 891–7

[4] Schwaller P, Laino T, Gaudin T, Bolgar P, Hunter C A, Bekas C and Lee A A 2019 Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction *ACS Cent. Sci.* **5** 1572–83

[5] Li Z *et al* 2020 Robot-accelerated perovskite investigation and discovery *Chem. Mater.* **32** 5650–663 (https://pubs.acs.org/doi/abs/10.1021/acs.chemmater.0c01153)

[6] Gromski P S, Henson A B, Granda J M and Cronin L 2019 How to explore chemical space using algorithms and automation *Nat. Rev. Chem.* **3** 119–28

[7] Sanchez-Lengeling B and Aspuru-Guzik A 2018 Inverse molecular design using machine learning: generative models for matter engineering *Science* **361** 360–5

[8] Carrasquilla J and Melko R G 2017 Machine learning phases of matter *Nat. Phys.* **13** 431–4

[9] Hezaveh Y D, Levasseur L P and Marshall P J 2017 Fast automated analysis of strong gravitational lenses with convolutional neural networks *Nature* **548** 555–7

[10] Agresti I, Viggianiello N, Flamini F, Spagnolo N, Crespi A, Osellame R, Wiebe N and Sciarrino F 2019 Pattern recognition techniques for boson sampling validation *Phys. Rev.* X **9** 011013

[11] Senior A W *et al* 2020 Improved protein structure prediction using potentials from deep learning *Nature* **577** 706–10 (https://www.nature.com/articles/s41586-019-1923-7)

[12] Zhavoronkov A *et al* 2019 Deep learning enables rapid identification of potent DDR1 kinase inhibitors *Nat. Biotechnol.* **37** 1038–40

[13] Stokes J M *et al* 2020 A deep learning approach to antibiotic discovery *Cell* **180** 688–702

[14] King R D *et al* 2009 The automation of science *Science* **324** 85–9

[15] Iten R, Metger T, Wilming H, Del Rio L and Renner R 2020 Discovering physical concepts with neural networks *Phys. Rev. Lett.* **124** 010508

[16] Lusch B, Nathan Kutz J and Brunton S L 2018 Deep learning for universal linear embeddings of nonlinear dynamics *Nat. Commun.* **9** 1–10

[17] Wetzel S J, Melko R G, Scott J, Panju M and Ganesh V 2020 Discovering symmetry invariants and conserved quantities by interpreting Siamese neural networks (arXiv:2003.04299)

[18] Roscher R, Bohn B, Duarte M F and Garcke J 2020 Explainable machine learning for scientific insights and discoveries *IEEE Access* **8** 42200–16

[19] Friedman J H 2001 Greedy function approximation: a gradient boosting machine *Ann. Stat.* **29** 1189–232 (https://www.jstor.org/stable/2699986)

[20] Rogers D and Hahn M 2010 Extended-connectivity fingerprints *J. Chem. Inf. Mod.* **50** 742–54

[21] Ho T K 1995 Random decision forests *Proc. 3rd Int. Conf. Document Analysis and Recognition* vol 1 (IEEE) pp 278–82

[22] Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

[23] Lopez S A, Sanchez-Lengeling B, de Goes Soares J and Aspuru-Guzik A 2017 Design principles and top non-fullerene acceptor candidates for organic photovoltaics *Joule* **1** 857–70

[24] Duvenaud D K, Maclaurin D, Iparraguirre J, Bombarell R, Hirzel T, Aspuru-Guzik A and Adams R P 2015 Convolutional networks on graphs for learning molecular fingerprints *Advances in Neural Information Processing Systems* pp 2224–32

[25] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D and Batra D 2017 Grad-cam: visual explanations from deep networks via gradient-based localization *Proc. IEEE Int. Conf. Computer Vision* pp 618–26

[26] Ying Z, Bourgeois D, You J, Zitnik M and Leskovec J 2019 Gnnexplainer: generating explanations for graph neural networks *Advances in Neural Information Processing Systems* pp 9244–55

[27] Gómez-Bombarelli R *et al* 2016 Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach *Nat. Mater.* **15** 1120–27

[28] Hachmann J, Olivares-Amaya R, Atahan-Evrenk S, Amador-Bedolla C, Sánchez-Carrera R S, Gold-Parker A, Vogt L, Brockway A M and Aspuru-Guzik A 2011 The Harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid *J. Phys. Chem. Lett.* **2** 2241–51

[29] Lopez S A, Pyzer-Knapp E O, Simm G N, Lutzow T, Li K, Seress L R, Hachmann J and Aspuru-Guzik A 2016 The Harvard organic photovoltaic dataset *Sci. Data* **3** 1–7

[30] Gershoni-Poranne R, Rahalkar A P and Stanger A 2018 The predictive power of aromaticity: quantitative correlation between aromaticity and ionization potentials and HOMO–LUMO gaps in oligomers of benzene, pyrrole, furan and thiophene *Phys. Chem. Chem. Phys.* **20** 14808–17

[31] Rasmussen S C 2018 2000 Nobel Prize in chemistry *Acetylene and its Polymers* (Berlin: Springer) pp 125–32

[32] Bredas J-L 1985 Relationship between band gap and bond length alternation in organic conjugated polymers *J. Chem. Phys.* **82** 3808–11

[33] Friis N, Vitagliano G, Malik M and Huber M 2019 Entanglement certification from theory to experiment *Nat. Rev. Phys.* **1** 72–87

[34] Erhard M, Krenn M and Zeilinger A 2020 Advances in high-dimensional quantum entanglement *Nat. Rev. Phys.* **2** 365

[35] Lawrence J 2017 Mermin inequalities for perfect correlations in many-qutrit systems *Phys. Rev.* A **95** 042123

[36] Pivoluska M, Huber M and Malik M 2018 Layered quantum key distribution *Phys. Rev.* A **97** 032312

[37] Hu X-M, Xing W-B, Zhang C, Liu B-H, Pivoluska M, Huber M, Huang Y-F, Li C-F and Guo G-C 2020 Experimental creation of multi-photon high-dimensional layered quantum states (arXiv:2001.06253)

[38] Pan J-W, Chen Z-B, Lu C-Y, Weinfurter H, Zeilinger A and Żukowski M 2012 Multiphoton entanglement and interferometry *Rev. Mod. Phys.* **84** 777

[39] Krenn M, Malik M, Fickler R, Lapkiewicz R and Zeilinger A 2016 Automated search for new quantum experiments *Phys. Rev. Lett.* **116** 090405

[40] Knott P A 2016 A search algorithm for quantum state engineering and metrology *New J. Phys.* **18** 073033

[41] Wallnöfer J, Melnikov A A, Dür W and Briegel H J 2020 Machine learning for long-distance quantum communication *PRX Quantum* **1** 010301

[42] Zhan X, Wang K, Xiao L, Bian Z, Zhang Y, Sanders B C, Zhang C and Xue P 2020 Experimental quantum cloning in a pseudo-unitary system *Phys. Rev.* A **101** 010302

[43] Krenn M, Erhard M and Zeilinger A 2020 Computer-inspired quantum experiments *Nature Rev. Phys.* **2** 649–61

[44] Krenn M, Kottmann J, Tischler N, and Aspuru-Guzik A 2020 Conceptual understanding through efficient inverse-design of quantum optical experiments (arXiv:2005.06443)

[45] Huber M and de Vicente J I 2013 Structure of multidimensional entanglement in multipartite systems *Phys. Rev. Lett.* **110** 030501

[46] Huber M, Perarnau-Llobet M and de Vicente J I 2013 Entropy vector formalism and the structure of multidimensional entanglement in multipartite systems *Phys. Rev.* A **88** 042328

[47] Malik M, Erhard M, Huber M, Krenn M, Fickler R and Zeilinger A 2016 Multi-photon entanglement in high dimensions *Nat. Photon.* **10** 248

[48] Erhard M, Malik M, Krenn M and Zeilinger A 2018 Experimental Greenberger–Horne–Zeilinger entanglement beyond qubits *Nat. Photon.* **12** 759–64

[49] Klyshko D N 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

[50] Gerling T, Wagenbauer K F, Neuner A M and Dietz H 2015 Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components *Science* **347** 1446–52

[51] Praetorius F, Kick B, Behler K L, Honemann M N, Weuster-Botz D and Dietz H 2017 Biotechnological mass production of DNA origami *Nature* **552** 84–7

[52] Temkin O N, Zeigarnik A V and Bonchev D G 1996 *Chemical Reaction Networks: A Graph-Theoretical Approach* (Boca Raton, FL: CRC Press)

[53] Rappoport D, Galvin C J, Zubarev D Y and Aspuru-Guzik A 2014 Complex chemical reaction networks from heuristics-aided quantum chemistry *J. Chem. Theory Comput.* **10** 897–907