



# Smooth input preparation for quantum and quantum-inspired machine learning

Zhikuan Zhao<sup>1,2,3</sup> · Jack K. Fitzsimons<sup>4</sup> · Patrick Rebentrost<sup>3</sup> · Vedran Dunjko<sup>5,6</sup> · Joseph F. Fitzsimons<sup>2,3,7</sup>

Received: 4 November 2020 / Accepted: 19 March 2021 / Published online: 26 April 2021  
© The Author(s) 2021

## Abstract

Machine learning has recently emerged as a fruitful area for finding potential quantum computational advantage. Many of the quantum-enhanced machine learning algorithms critically hinge upon the ability to efficiently produce states proportional to high-dimensional data points stored in a quantum accessible memory. Even given query access to exponentially many entries stored in a database, the construction of which is considered a one-off overhead, it has been argued that the cost of preparing such amplitude-encoded states may offset any exponential quantum advantage. Here we prove using smoothed analysis that if the data analysis algorithm is robust against small entry-wise input perturbation, state preparation can always be achieved with constant queries. This criterion is typically satisfied in realistic machine learning applications, where input data is subjective to moderate noise. Our results are equally applicable to the recent seminal progress in quantum-inspired algorithms, where specially constructed databases suffice for polylogarithmic classical algorithm in low-rank cases. The consequence of our finding is that for the purpose of practical machine learning, polylogarithmic processing time is possible under a general and flexible input model with quantum algorithms or quantum-inspired classical algorithms in the low-rank cases.

**Keywords** Data structure for quantum machine learning · Input encoding · Query complexities · Smoothed complexity analysis

## 1 Introduction

In recent years, there has been substantial interest in algorithms based on “quantum linear algebra”, where quantum

states are used to represent vectors with exponentially large dimensions, which are manipulated by large matrices representing quantum operations (Zhao et al. 2019; Gilyén et al. 2018). A particularly fruitful domain of applying such methods is quantum machine learning, where quantum algorithms promise for exponential or high-degree polynomial improvements on computational bottlenecks, or enhancement in model expressiveness due to the ability to construct kernels using access to exponentially large Hilbert spaces (Aïmeur et al. 2006; Lloyd et al. 2013; 2014b; Schuld et al. 2014; Zhao et al. 2019; Dunjko and Briegel 2018; Zhao et al. 2019; Zhao et al. 2019; Schuld and Killoran 2019; Tiwari et al. 2020; Dehdashti et al. 2020; Kübler et al. 2019; Havlíček et al. 2019). Many of these methods leverage quantum algorithms for solving linear systems that trace back to the seminal result of Harrow, Hasidim and Lloyd (2009). This algorithm provides a way to generate quantum states proportional to  $A^{-1}\mathbf{x}$ , in time logarithmic in the dimension of the vector  $\mathbf{x}$  which was subsequently improved to achieve exponentially better precision (Childs et al. 2017). However, as Aaronson pointed out (Aaronson 2015), these quantum-assisted machine learning

---

✉ Zhikuan Zhao  
zhikuan.zhao@inf.ethz.ch

- <sup>1</sup> Department of Computer Science, ETH Zurich, Universitatstrasse 6, 8092 Zurich, Switzerland
- <sup>2</sup> Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, Singapore
- <sup>3</sup> Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore
- <sup>4</sup> Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK
- <sup>5</sup> Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany
- <sup>6</sup> LIACS, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
- <sup>7</sup> Horizon Quantum Computing, 79 Ayer Rajah Crescent, Singapore 139955, Singapore

approach comes with several caveats relating to the sparsity and conditioning of the matrix involved, and the structure of the input vector. Crucially, nearly all the algorithms assume access to a quantum accessible database which can produce quantum states with amplitudes proportional to the classical input entries, which is referred to as amplitude encoding. This requirement persists even under the reasonable assumption that the loading and construction of the required databases constitute a one-off overhead, and does not contribute to the computational complexity analysis of data processing.

A breakthrough came with the seminal work by Kerenidis and Prakash who provided the first end-to-end quantum machine learning algorithm with explicit description of quantum accessible data structure for efficient input preparation, which was applied to recommendation systems (Kerenidis and Prakash 2017). The same data structure was subsequently applied to improve the performance of the quantum linear system algorithm in dense cases (Wossnig et al. 2018). The Kerenidis-Prakash (KP) model explicitly stores the amplitudes in a binary tree structure. As such in addition to the ability of realising amplitude encoding for the row vectors of a matrix, the KP structure further allows for a stronger ability for preparing classical probability distributions proportional to vectors of entries squared – so-called  $\ell_2$ -sampling. From a classical perspective, intriguingly, recent results of quantum-inspired machine learning algorithms (Tang 2018a; Gilyén et al. 2018; Tang 2018b; Chia et al. 2018) have shown that assuming such a data structure capable of efficient  $\ell_2$ -sampling leads to equally efficient classical algorithms in the low-rank cases. The KP structure allows for input preparation both in the quantum and classical cases with a logarithmic storage overhead. The entries in the KP structure, i.e. the partial sums, are fixed beforehand and stored in the data structure. As a result, specific vectors for which the correct partial sums are stored can be directly accessed as amplitude-encoded quantum state without post-selection. However, practical situations may arise where computation and re-computation of the partial sums is inefficient. For instance, it could be desirable to generate input vectors collecting entries from differing rows of a matrix stored in the data structure in statistical resampling techniques for cross-validation, or specific entry-wise functions are required for different algorithms using the same input data. In such situations, a more flexible and general input query model would have significant relevance. Alternatively, some quantum machine learning algorithms operate on an input model in which data is stored as entries in density matrices and subsequently read out by exponentiating density matrices and simulating them as Hamiltonians (Lloyd et al. 2014a; Rebentrost et al. 2014). This method requires explicit encoding and storage of data as quantum mixed states, and restricted by a similar rigidity as the KP model.

## 2 Input Model and Amplitude Encoding

In this work, we examine the generic input model where only entry-wise access is allowed, and prove using smoothed analysis that both quantum amplitude encoding and classical  $\ell_2$ -sampling can be achieved with constant queries in realistic machine learning application with moderately noisy input data. In the quantum setting, entry-wise access is equivalent to having access to an oracle  $O_{\mathbf{x}}$  that implements the unitary transformation,

$$\sum_{i,j} \alpha_{ij} |i\rangle |j\rangle \xrightarrow{O_{\mathbf{x}}} \sum_{i,j} \alpha_{ij} |i\rangle |j + f(x_i)\rangle. \quad (1)$$

whereas in the classical setting, it is the standard random access memory that provides the mapping,  $i \rightarrow f(x_i)$ . Without committing to an ad hoc data structure either in the form of the KP model or density matrix access, primitive input assumption reserves the flexibility of loading the data as arbitrary entry-wise functions. We assume in the sequel  $f(x_i) = x_i$  for simplicity. Generating the exact amplitude-encoded states (or  $\ell_2$ -samples) from the entry-wise oracle, by the Grover (or unstructured) search lower bound requires  $\Omega(D^{1/2})$  ( $\Omega(D)$ , respectively) calls for  $\mathbf{x} \in \mathbb{R}^D$ . This observation is a common critique on early quantum machine proposals which did not explicitly address the issue of initial state preparation, yet attempted to argue for polylogarithmic runtimes. Here we show that, in fact, for practical machine learning where the algorithm is robust against moderate input noise represented by entry-wise perturbations, quantum state preparation based on amplitude encoding has high success probability with only constant query cost. Furthermore, By an analogous argument, we also show that low-rank quantum-inspired algorithms operating under the same practical assumptions do not require access to special data structures to attain polylogarithmic runtime.

For quantum amplitude encoding, our objective is to prepare the state

$$|\mathbf{x}\rangle = \|\mathbf{x}\|_2^{-1} \sum_{i=1}^D x_i |i\rangle \quad (2)$$

with amplitudes proportional to the entries  $x_i$  of  $\mathbf{x} \in \mathbb{R}^D$ , given access to an oracle that realises the mapping in Eq. 1. Quantum random access memory (QRAM) provides one way to implementing such an operation. In this case,  $\mathbf{x}$  is stored classically and one is able to access the corresponding memory cells in quantum superposition (Giovannetti et al. 2008). In other cases, the oracle can also be constructed if  $x_i$  is efficiently computable given the index  $i$ . To produce  $|\mathbf{x}\rangle$  probabilistically for any  $\mathbf{x}$ , one can start with the state  $D^{-\frac{1}{2}} \sum_i |i\rangle |0\rangle$  as the query state for the

operation in Eq. 1 to obtain  $D^{-\frac{1}{2}} \sum_i |i\rangle |x_i\rangle$ . An ancillary qubits is prepared in state  $|0\rangle$  and then conditionally rotated based on the second register to obtain

$$D^{-\frac{1}{2}} \sum_i |i\rangle |x_i\rangle \left( \sqrt{1 - |x_i|^2} |0\rangle + x_i |1\rangle \right), \quad (3)$$

where we have assumed for simplicity that  $\mathbf{x}$  is normalised such that  $|x_i| \leq 1$ . Performing a second oracle call to uncompute the registers  $|x_i\rangle$ , and post-selecting onto  $|1\rangle$  results in the desired state  $|\mathbf{x}\rangle$ .

A known barrier to state preparation in the oracle setting is the probability of projecting onto the correct subspace in the final step, which is given by  $D^{-1} \sum_i |x_i|^2$ . When the entries of  $\mathbf{x}$  are of similar magnitude,  $|\mathbf{x}\rangle$  can be prepared using a constant number queries. However, in the case where a few entries are much larger than the rest, the lower bounds on unordered search (Boyer et al. 1996) imply that the corresponding state requires  $\Omega(\sqrt{D})$  oracle queries (Soklakov and Schack 2006). This argument can be extended to the case where it is only necessary to prepare any  $|\mathbf{x}'\rangle$  such that  $\|\mathbf{x}' - \mathbf{x}\|_2$  is sufficiently small.

However, here we make the observation that if real-valued data comes from a realistic source subjective to noise, such as datasets of ocean temperatures (International Comprehensive Ocean-Atmosphere Data Set) and geostatistics (Practical Geostatistics 2000 Data Sets), and the algorithm is robust against such realistic input noise, then quantum state preparation can be done efficiently with constant oracle queries. Formally, our argument is based on analysing the smoothed complexity for the amplitude encoding procedure when the input vectors are subjective to small perturbations. The smoothed complexity was introduced by Spielman and Teng (2009), originally to explain the efficient performance of the simplex algorithm for linear programming in typical real-world scenarios. The same line of reasoning was subsequently applied to analyse the practical efficiency of various important algorithms in mathematical programming, machine learning, numerical analysis discrete mathematics and combinatorics optimisation (Spielman and Teng 2001).

### 3 Smoothed Analysis

The key intuition here is to analyse the performance of the algorithms when the worst-case data input is subject to noise, which is represented by a small Gaussian element-wise perturbation. Following the convention of Spielman and Teng (2009), we state the definition of smoothed complexity and then prove that preparing amplitude-encoded states has a constant smoothed complexity:

**Definition 1** (Smoothed Complexity (Spielman and Teng 2009)) Given an algorithm  $\mathcal{A}$  with an input domain  $\Omega_D = \mathbb{R}^D$ , the smoothed complexity of  $\mathcal{A}$  with  $\sigma$ -Gaussian perturbation is defined as

$$\text{Smoothed}_{\mathcal{A}}^{\sigma}(D) = \max_{\mathbf{x} \in [-1, 1]^D} \mathbb{E}_{\mathbf{g}}[T_{\mathcal{A}}(\mathbf{x} + \mathbf{g})], \quad (4)$$

where  $\mathbf{g}$  is a Gaussian random vector with variance  $\sigma^2$ , and  $T_{\mathcal{A}}$  denotes the runtime of  $\mathcal{A}$ .

Furthermore,  $\mathcal{A}$  is said to have polynomial smoothed complexity if there exist positive constants  $k_1, k_2, D_0, \sigma_0$  and  $c$  such that for all  $D \geq D_0$  and  $0 \leq \sigma \leq \sigma_0$ , it holds that  $\text{Smoothed}_{\mathcal{A}}^{\sigma}(D) \leq c\sigma^{-k_2} D^{k_1}$ . (5)

**Theorem 1** Given oracle access,  $O_{\mathbf{x}}$  to the entries of  $\mathbf{x} \in \mathbb{R}^D$ , the amplitude encoding of  $\mathbf{x}$  into  $|\mathbf{x}\rangle$  has smoothed complexity  $\mathcal{O}(1/\sigma)$ .

*Proof* Let  $\mathcal{A}$  be the algorithm that maps  $D^{-\frac{1}{2}} \sum_i |i\rangle |x_i\rangle$  into  $|\mathbf{x}\rangle = \|\mathbf{x}\|_2^{-1} \sum_{i=1}^D x_i |i\rangle$ . After applying the controlled rotation and uncomputing the second register, the optimal success probability of projecting the state,

$$D^{-\frac{1}{2}} \sum_i |i\rangle \left( \sqrt{1 - |x_i|^2} |0\rangle + x_i |1\rangle \right) \quad (6)$$

onto the desired state  $|\mathbf{x}\rangle$  is given by  $P_{\mathcal{A}} = \mathcal{O}(\frac{\|\mathbf{x}\|_2}{\sqrt{D}})$  with fixed-point amplitude amplification (Gilyén et al. 2018). From the definition of smoothed complexity, Eq. 4, in the worst case, we have

$$\begin{aligned} \text{Smoothed}_{\mathcal{A}}^{\sigma}(D) &= \max_{\mathbf{x} \in [-1, 1]^D} \mathbb{E}_{\mathbf{g}}[T_{\mathcal{A}}(\mathbf{x} + \mathbf{g})] \\ &= (\min_{\mathbf{x} \in [-1, 1]^D} \mathbb{E}_{\mathbf{g}}[P_{\mathcal{A}}(\mathbf{x} + \mathbf{g})])^{-1}, \end{aligned} \quad (7)$$

where the second line follows from the fact that the expected runtime is inversely proportional to the expected success probability. Note that since  $P_{\mathcal{A}}(\mathbf{x} + \mathbf{g}) = \mathcal{O}(\frac{\|\mathbf{x} + \mathbf{g}\|_2}{\sqrt{D}})$  and  $\mathbf{g}$  is a zero-mean Gaussian random vector, the minimum of expected success probability is obtained at  $\mathbf{x} = 0$ . We thus have

$$\begin{aligned} \text{Smoothed}_{\mathcal{A}}^{\sigma}(D) &= (\mathbb{E}_{\mathbf{g}}[P_{\mathcal{A}}(\mathbf{g})])^{-1} \\ &= D^{\frac{1}{2}} (\mathbb{E}[\|\mathbf{g}\|_2])^{-1}. \\ &= \mathcal{O}\left(\frac{1}{\sigma}\right), \end{aligned} \quad (8)$$

where the last equality followed by noting that the random variable,  $\|\mathbf{g}\|_2$ , by definition follows a chi distribution with mean,

$$\mathbb{E}[\|\mathbf{g}\|_2] = \sqrt{2}\sigma \frac{\Gamma((D+1)/2)}{\Gamma(D/2)} = \mathcal{O}(\sigma\sqrt{D}). \quad (9)$$

□

The result of Eq. 8 implies that when the input vector is subjective to a certain level of noise represented by an element-wise Gaussian perturbation, the query complexity of preparing amplitude encoding is independent of the dimensionality.

We have seen that the quantum state preparation based on amplitude encoding has a constant runtime given that the input is subjective to a finite variance Gaussian noise. An analogous reasoning applies in the fully classical setting. Classically, particularly in the context of quantum-inspired ML algorithm,  $\ell_2$ -sampling can be achieved by simple rejection sampling: an entry is chosen uniformly at random, and a value  $x_j$  is read (we assume  $x_j \leq c$ , and  $c$  is a known constant upper bound on the entries). Then a random real is sampled from the interval  $(0, c)$ , and if this value is below  $|x_j|^2$ , the value  $j$  is output. Otherwise the process is repeated. As desired, the acceptance probability of the element  $j$  is given by  $P_j^{\text{accept}} = |x_j|^2$ , which leads to the following average runtime for producing an  $\ell_2$ -sample from the correct distribution:

$$\mathbb{E}[T_{\mathcal{R}}] = \left( \frac{1}{D} \sum_j P_j^{\text{accept}} \right)^{-1} = D \left( \sum_j |x_j|^2 \right)^{-1}. \quad (10)$$

We can make an analogous smoothed analysis for this classical rejection sampling process. Denoting  $\mathcal{R}$  as the algorithm that performs rejection sampling, we have

$$\begin{aligned} \text{Smoothed}_{\mathcal{R}}^{\sigma}(D) &= \max_{\mathbf{x} \in [-1, 1]^D} \mathbb{E}_{\mathbf{g}}[T_{\mathcal{R}}(\mathbf{x} + \mathbf{g})] \\ &= D (\mathbb{E}[\|\mathbf{g}\|_2])^{-2}. \\ &= \mathcal{O}\left(\frac{1}{\sigma^2}\right). \end{aligned} \quad (11)$$

Thus the smoothed complexity of preparing quantum amplitude encoding from QRAM queries and classical  $\ell_2$ -samplings from classical RAM are  $\mathcal{O}(\frac{1}{\sigma})$  and  $\mathcal{O}(\frac{1}{\sigma^2})$  respectively given a  $\sigma^2$ -variance Gaussian perturbation on the input. The quadratic quantum improvement in the dependency on  $\sigma$  comes directly from amplitude amplification.

## 4 Element-Wise Perturbation: A Closer Look

In practical settings, the effect of a noisy element-wise perturbation on input data is well-studied in the machine learning literature (Nettleton et al. 2010; Kalapanidas et al. 2003; Cesa-Bianchi et al. 2011). As a specific example, in the domain of computer vision, Dodge and Karam (2016) examined the effect of various synthetic noise effects on the performance of popular deep learning computer vision models such as Caffe, VGG16, VGG-CNN-S, and GoogLeNet. Two relevant results from the work, the pixel-wise Gaussian noise and a change in contrast to the image,

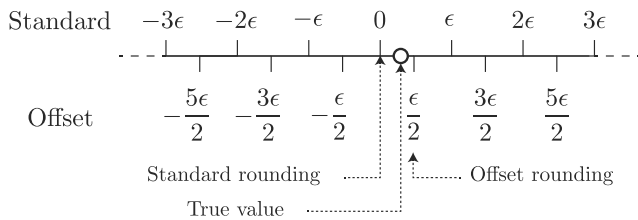
make a little effect on the performance. While adversarial attacks aim to find the worst-case corruption which leads to misclassification, Dodge and Karam (2016) empirically show that random perturbations of a half pixel and small shifts in the pixel values will likely have a negligible effect. Furthermore, if the entry-wise perturbation represents a systematic shift in the data points, it has no effect on a large class of useful machine learning models. For examples, kernel methods such as Gaussian processes, support vector machines, determinantal point processes and Gaussian Markov random fields, to name but a few, most commonly use stationary kernels which are shift invariant (Genton 2001). More generally, any digital data processing with floating-point arithmetic only makes sense if the overall results of such a computation maintain their validity when the features of the input vector had been perturbed below the machine precision. This is also practically reasonable, since real-world data come from measurements of finite precision.

An element-wise perturbation can be represented by an off-set in the  $\infty$ -norm induced distance between the original vector  $\mathbf{x}$  and the perturbed vector  $\mathbf{x}'$ . Assuming that the data analysis process, along with the particular instance of the input data, is robust against such small  $\infty$ -norm perturbations, we can in fact choose to work with the vector entries  $x'_i$  which are half-integer multiples of the base precision  $\epsilon$ . Such that  $\mathbf{x}'$  is chosen to be the closest representable vector to  $\mathbf{x}$  (as shown in Fig. 1), which satisfies  $|\mathbf{x}' - \mathbf{x}|_{\infty} \leq \frac{\epsilon}{2}$  and the distance from the true value of the data is less than  $\epsilon$ . This offset rounding can be implemented in the oracle access stage, or effectively realised at the controlled rotation stage, as discussed earlier. The aforementioned robustness assumption in data processing implies the analysis of results are insensitive to such offset rounding. In the quantum setting, the key benefit in using the offset rounding is that, since the exact representation of 0 is not included in this offset rounding convention, we have  $|x_i| \geq \frac{\epsilon}{2}$  and the probability of the final projection step succeeding is at least

$$D^{-1} \sum_i |x_i|^2 \geq \frac{\epsilon^2}{4}, \quad (12)$$

and hence independent of  $D$ . Inverting the lower bound of success probability leads to an upper bound of number of expected queries of  $\mathcal{O}(1/\epsilon^2)$ . The query complexity can be improved to  $\mathcal{O}(1/\epsilon)$  using fixed-point quantum amplitude amplification based on the methods presented in Gilyén et al. (2018), which achieves the same optimal query complexity as in Yoder et al. (2014), but without introducing an additional phase that could be undesirable if multiple vectors are required to be prepared in superposition. Note that in some cases, a systematic perturbation of data-vectors, by utilising, say, a positive sign offset  $(+\epsilon/2)$  to data points





**Fig. 1** Numerical rounding conventions. In the standard rounding convention scalar values are rounded to the nearest integer multiple of precision  $\epsilon$ . Alternatively, we can consider an offset rounding convention, where the rounding is to the nearest half-integer multiple of  $\epsilon$ . As a result, this numbering convention does not contain an exact representation of 0. In either scheme, the rounded value is always within  $\frac{\epsilon}{2}$  of the true value

is undesirable. In these cases, one can opt for a near-white noise offset using either a suitable pseudo-random number generator seeded by the memory location being queried or by adding random data  $w_i$ , performing  $|i\rangle|x_i\rangle \rightarrow |i\rangle|x_i + w_i\rangle$ . Critically, however, the number of oracle queries necessary to successfully prepare the state always has an upper bound that is independent from the database size.

It is worth cautioning that there exist scenarios, especially in computational learning theory (Arunachalam and de Wolf 2017), for which the robustness assumption against entry-wise data perturbation does not generally hold. For instance, if the input vector contains zeros with special meanings such as indicating a canonical vector or representing unknown values, the model may be sensitive toward shifting the zeros to even a small value. For instance, when loading a high-dimensional data point with constant or polylogarithmic sparsity, a systematic shift on the zero entries will produce a state vector converging to a uniform superposition, hence losing necessary information about the original input for meaningful analysis. Nevertheless, the constraint that  $\|\mathbf{x}' - \mathbf{x}\|_\infty \leq \epsilon$  rather than requiring closeness in the 2-norm is still meaningful to a large class of practical machine learning tasks. We should also note that the quantum state preparation discussed here is useful to quantum algorithms which aim to improve the efficiency of conventional classical algorithms but otherwise realise the same input-output functionality, such that the quantum algorithms inherit the desired robustness property of their conventional counter-parts.

## 5 Conclusion

In summary, we have shown that any application which is robust under small  $\infty$ -norm perturbations, as it is the case in most practical machine learning, allows for efficient input preparation, both in the sense of classical  $\ell_2$ -sampling and the coherent amplitude-encoded state preparation. In

the context of quantum machine learning, this suggests that the caveat related to state preparation raised by Aaronson (2015) can generally be overcome for a wide range of practical use-cases, due to the natural robustness assumption. In the context of quantum-inspired machine learning, this finding removes the necessity of special data structures that involve the storage of partial sums. Hence we have provided a concrete argument for the feasibility of input preparation for both quantum and quantum-inspired algorithms for machine learning under the most general and flexible entry-wise query access model, which we believe will present both conceptual merit and practical utility to the promising exploration between quantum information and machine learning.

**Acknowledgements** The authors thank Ashley Montanaro and Ronald de Wolf for helpful comments on the manuscript. JFF and ZZ acknowledge support from Singapore's Ministry of Education and National Research Foundation. JFF acknowledges support from the Air Force Office of Scientific Research under AOARD grant FA2386-15-1-4082. This material is based on research funded in part by the Singapore National Research Foundation under NRF Award NRF-NRFF2013-01. This work was also in part supported by the Dutch Research Council (NWO/OCW), as part of the Quantum Software Consortium program (project number 024.003.037), and by the project NEASQC funded from the European Union's Horizon 2020 research and innovation programme (grant agreement No 951821).

**Funding** Open Access funding provided by ETH Zurich.

## Declarations

**Competing interests** Dr. Fitzsimons has financial holdings in Horizon Quantum Computing Pte Ltd. Otherwise, the authors have no potential financial or non-financial conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aaronson S (2015) Nat Phys 11(4):291–293
- Aïmeur E, Brassard G, Gambs S (2006) Advances in Artificial Intelligence, Springer, pp. 431–442
- Arunachalam S, de Wolf R (2017) ACM SIGACT News 48(2):41–67
- Boyer M, Brassard G, Høyer P, Tapp A (1996) arXiv preprint [9605034](https://arxiv.org/abs/9605034)
- Cesa-Bianchi N, Shalev-Shwartz S, Shamir O (2011) IEEE Trans Inf Theory 57(12):7907–7931

- Chia N-H, Lin H-H, Wang C (2018) arXiv preprint arXiv:[1811.04852](https://arxiv.org/abs/1811.04852)
- Childs AM, Kothari R, Somma RD (2017) SIAM J Comput 46(6):1920–1950
- Dehdashti S, Moreira C, Obeid AK, Bruza P (2020) arXiv preprint arXiv:[2006.02904](https://arxiv.org/abs/2006.02904)
- Dodge S, Karam L (2016) In: Quality of multimedia experience (QoMEX), 2016 Eighth international conference on (IEEE), pp. 1–6
- Dunjko V, Briegel HJ (2018) Rep Prog Phys 81 074001
- Genton MG (2001) Journal of machine learning research 2:299
- Gilyén A, Lloyd S, Tang E (2018) arXiv preprint arXiv:[1811.04909](https://arxiv.org/abs/1811.04909)
- Gilyén A, Su Y, Low G, Wiebe N (2018) arXiv preprint arXiv:[1806.01838](https://arxiv.org/abs/1806.01838)
- Giovannetti V, Lloyd S, Maccone L (2008) Phys Rev Lett 100(16):160501
- Harrow AW, Hassidim A, Lloyd S (2009) Phys Rev Lett 103(15):150502
- Havlíček V, Córcoles AD, Temme K, Harrow AW, Kandala A, Chow JM, Gambetta JM (2019) Nature 567(7747):209–212
- International Comprehensive Ocean-Atmosphere Data Set (<http://icoads.noaa.gov/>)
- Kalapanidas E, Avouris N, Craciun M, Neagu D (2003) Machine learning algorithms: a study on noise sensitivity. In Proc. 1st Balcan Conference in Informatics. pp. 356–365
- Kerenidis I, Prakash A (2017) Quantum Recommendation Systems. In: Papadimitriou C. H (ed) 8th Innovations in Theoretical Computer Science Conference (ITCS 2017), vol 67. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, pp 49:1–49:21, <https://doi.org/10.4230/LIPICs.ITCS.2017.49>
- Kübler JM, Muandet K, Schölkopf B (2019) Physical Review Research 1(3):033159
- Lloyd S, Mohseni M, Rebentrost P (2013) arXiv preprint arXiv:[1307.0411](https://arxiv.org/abs/1307.0411)
- Lloyd S, Mohseni M, Rebentrost P (2014) Nat Phys 10(9):631–633
- Lloyd S, Mohseni M, Rebentrost P (2014) Phys Rev Lett 113(13):130503
- Nettleton DF, Orriols-Puig A, Fornells A (2010) Artif Intell Rev 33(4):275–306
- Practical Geostatistics 2000 Data Sets (<http://www.kriging.com/datasets/>)
- Rebentrost P, Mohseni M, Lloyd S (2014) Phys Rev Lett 113:130503
- Schuld M, Killoran N (2019) Phys Rev Lett 122(4):040504
- Schuld M, Sinayskiy I, Petruccione F (2014) In: PRICAI 2014: Trends in artificial intelligence, Springer, pp. 208–220
- Sokolov AN, Schack R (2006) Phys Rev A 73(1):012307
- Spielman D, Teng SH (2001) In: ACM symposium on theory of computing, pp. 296–305
- Spielman DA, Teng S-H (2009) Commun ACM 52(10):76–84
- Tang E (2018) arXiv preprint arXiv:[1807.04271](https://arxiv.org/abs/1807.04271)
- Tang E (2018) arXiv preprint arXiv:[1811.00414](https://arxiv.org/abs/1811.00414)
- Tiwari P, Dehdashti S, Obeid AK, Melucci M, Bruza P (2020) arXiv preprint arXiv:[2007.07887](https://arxiv.org/abs/2007.07887)
- Wossnig L, Zhao Z, Prakash A (2018) Phys Rev Lett 120(5):050502
- Yoder TJ, Low GH, Chuang IL (2014) Phys Rev Lett 113(21):210501
- Zhao L, Zhao Z, Rebentrost P, Fitzsimons J (2019) arXiv preprint arXiv:[1902.10394](https://arxiv.org/abs/1902.10394)
- Zhao Z, Fitzsimons JK, Fitzsimons JF (2019) Phys. Rev. A 99:052331. <https://doi.org/10.1103/PhysRevA.99.052331>
- Zhao Z, Fitzsimons JK, Osborne MA, Roberts SJ, Fitzsimons JF (2019) Phys Rev A 100(1):012304
- Zhao Z, Pozas-Kerstjens A, Rebentrost P, Wittek P (2019) Quantum Machine Intelligence. <https://doi.org/10.1007/s42484-019-00004-7>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.