

# Noisy intermediate-scale quantum (NISQ) algorithms

Kishor Bharti,<sup>1,\*</sup> Alba Cervera-Lierta,<sup>2,3,\*</sup> Thi Ha Kyaw,<sup>2,3,\*</sup> Tobias Haug,<sup>4</sup> Sumner Alperin-Lea,<sup>3</sup> Abhinav Anand,<sup>3</sup> Matthias Degroote,<sup>2,3,5</sup> Hermanni Heimonen,<sup>1</sup> Jakob S. Kottmann,<sup>2,3</sup> Tim Menke,<sup>6,7,8</sup> Wai-Keong Mok,<sup>1</sup> Sukin Sim,<sup>9</sup> Leong-Chuan Kwek,<sup>1,10,11,†</sup> and Alán Aspuru-Guzik<sup>2,3,12,13,‡</sup>

<sup>1</sup>Centre for Quantum Technologies, National University of Singapore 117543, Singapore  
<sup>2</sup>Department of Computer Science, University of Toronto, Toronto, Ontario M5S 2E4, Canada  
<sup>3</sup>Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario M5G 1Z8, Canada  
<sup>4</sup>QOLS, Blackett Laboratory, Imperial College London SW7 2AZ, UK  
<sup>5</sup>current address: Boehringer Ingelheim, Amsterdam, Netherlands  
<sup>6</sup>Department of Physics, Harvard University, Cambridge, MA 02138, USA  
<sup>7</sup>Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA  
<sup>8</sup>Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA  
<sup>9</sup>Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA  
<sup>10</sup>MajLab, CNRS-UNS-NUS-NTU International Joint Research Unit UMI 3654, Singapore  
<sup>11</sup>National Institute of Education and Institute of Advanced Studies, Nanyang Technological University 637616, Singapore  
<sup>12</sup>Vector Institute for Artificial Intelligence, Toronto, Ontario M5S 1M1, Canada  
<sup>13</sup>Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada

(Dated: January 22, 2021)

A universal fault-tolerant quantum computer that can solve efficiently problems such as integer factorization and unstructured database search requires millions of qubits with low error rates and long coherence times. While the experimental advancement towards realizing such devices will potentially take decades of research, noisy intermediate-scale quantum (NISQ) computers already exist. These computers are composed of hundreds of noisy qubits, i.e. qubits that are not error-corrected, and therefore perform imperfect operations in a limited coherence time. In the search for quantum advantage with these devices, algorithms have been proposed for applications in various disciplines spanning physics, machine learning, quantum chemistry and combinatorial optimization. The goal of such algorithms is to leverage the limited available resources to perform classically challenging tasks. In this review, we provide a thorough summary of NISQ computational paradigms and algorithms. We discuss the key structure of these algorithms, their limitations, and advantages. We additionally provide a comprehensive overview of various benchmarking and software tools useful for programming and testing NISQ devices.

## CONTENTS

I. Introduction	2	3. Resource-aware optimizers	17
A. Computational complexity theory in a nutshell	2	III. Other NISQ approaches	19
B. Experimental progress	3	A. Quantum annealing	19
C. NISQ and near-term	4	B. Gaussian boson sampling	21
D. Scope of the review	5	1. The protocol	21
II. Building blocks of variational quantum algorithms	5	2. Applications	22
A. Objective function	5	C. Analog quantum simulation	22
1. Pauli strings	7	1. Implementations	23
2. Fidelity	7	2. Programmable quantum simulators	23
3. Other objective functions	7	D. Digital-analog quantum simulation and computation	23
B. Parameterized quantum circuits	8	E. Iterative quantum assisted eigensolver	24
1. Problem-inspired ansätze	8	IV. Maximizing NISQ utility	25
2. Hardware-efficient ansätze	11	A. Quantum error mitigation (QEM)	25
C. Measurement	12	1. Zero-noise extrapolation	25
D. Parameter optimization	13	2. Probabilistic error cancellation	27
1. Gradient-based approaches	13	3. Other QEM strategies	28
2. Gradient-free approaches	16	B. Barren plateaus	29
		C. Expressibility of variational ansätze	31
		D. Reachability	32
		E. Theoretical guarantees of the QAOA algorithm	32
		F. Circuit compilation	33
		1. Native and universal gate sets	33
		2. Circuit decompositions	34
		3. The qubit mapping problem	34
		V. Applications	35

\* These authors contributed equally to this work.

kishor.bharti1@gmail.com

a.cervera.lierta@gmail.com

thihakyaw.phy@gmail.com

† cqtclc@gmail.com

‡ alan@aspuru.com

A. Many-body physics and chemistry	35
1. Qubit encodings	35
2. Constructing electronic Hamiltonians	36
3. Variational quantum eigensolver	37
4. Variational quantum eigensolver for excited states	38
5. Hamiltonian simulation	40
6. Quantum information scrambling and thermalization	41
7. Simulating open quantum systems	41
8. Nonequilibrium steady state	42
9. Gibbs state preparation	43
10. Many-body ground state preparation	43
11. Quantum autoencoder	44
12. Quantum computer-aided design	44
B. Machine learning	45
1. Supervised learning	46
2. Unsupervised learning	48
3. Reinforcement learning	49
C. Combinatorial optimization	50
1. Max-Cut	50
2. Other combinatorial optimization problems	52
D. Numerical solvers	52
1. Variational quantum factoring	52
2. Singular value decomposition	53
3. Linear system problem	53
4. Non-linear differential equations	54
E. Finance	54
1. Portfolio optimization	55
2. Fraud detection	56
F. Other applications	56
1. Quantum foundations	56
2. Quantum optimal control	56
3. Quantum metrology	57
4. Fidelity estimation	57
5. Quantum error correction	57
6. Nuclear physics	57
7. Entanglement properties	58
VI. Benchmarking	58
A. Randomized benchmarking	58
B. Quantum volume	59
C. Cross-entropy benchmarking	60
D. Application benchmarks	61
VII. Quantum software tools	61
VIII. Outlook	62
A. NISQ goals	63
B. Long-term goal: fault-tolerant quantum computing	64
Acknowledgements	65
References	65
Tables of applications	78
Table of software packages	81
Table of external libraries	82

## I. INTRODUCTION

Quantum computing originated in the eighties when physicists started to speculate about computational models that integrate the laws of quantum mechanics (Kaiser, 2011). Starting with the pioneering works of Benioff and Deutsch, which involved the study of quantum Turing

machines and the notion of universal quantum computation (Benioff, 1980; Deutsch, 1985), the field continued to develop towards its proposed theoretical application: the simulation of quantum systems (Feynman, 1982; Lloyd, 1996; Manin, 1980). Arguably, the drive for quantum computing took off in 1994 when Peter Shor provided an efficient quantum algorithm for finding prime factors of composite integers, rendering most classical cryptographic protocols unsafe (Shor, 1994). Since then, the study of quantum algorithms has matured as a sub-field of quantum computing with applications in search and optimization, machine learning, simulation of quantum systems and cryptography (Montanaro, 2016).

In the last forty years, many scientific disciplines have converged towards the study and development of quantum algorithms and their experimental realization. Quantum computers are, from the computational complexity perspective, fundamentally different tools available to computationally intensive fields. The implementation of aforementioned quantum algorithms requires that the minimal quantum information units, *qubits*, are as reliable as classical bits. Qubits need to be protected from environmental noise that induces decoherence but, at the same time, must allow their states to be controlled by external agents. This control includes the interaction that generates entanglement between qubits and the measurement operation that extracts the output of the quantum computation, called *read out*. It is technically possible to tame the effect of noise without compromising the quantum information process by developing quantum error correction (QEC) protocols (Lidar and Brun, 2013; Shor, 1995; Terhal, 2015). Unfortunately, the overhead of QEC in terms of the number of qubits is, at present day, still far from current experimental capabilities.

Most of the originally proposed quantum algorithms require millions of physical qubits to incorporate these QEC techniques successfully. However, existing quantum devices contain on the order of 100 qubits. Realizing the daunting goal of building an error-corrected quantum computer with millions of physical qubits may take decades. Currently realized quantum computers are the so called “*Noisy Intermediate-Scale Quantum* (NISQ)” devices (Preskill, 2018), i.e. those devices whose qubits and quantum operations are substantially imperfect. One of the goals in the NISQ era is to extract the maximum quantum computational power from current devices while developing techniques (Preskill, 2018) that may also be suited for the long-term goal of the fault-tolerant quantum computation (Terhal, 2015).

### A. Computational complexity theory in a nutshell

Defining a new computational paradigm increases ambition for tackling unsolved problems, but also exposes the limits of the new paradigm. New computational com-

plexity classes have been recognized through the study of quantum computing, and proposed algorithms and goals have to be balanced within well-known mathematical boundaries.

In this review, we will often use some computational complexity-theoretic ideas to establish the complexity domain and efficiency of the quantum algorithms covered. For this reason, we provide in this subsection a brief synopsis for a general audience and refer to (Arora and Barak, 2009) for a more comprehensive treatment.

The computational complexity of functions is typically expressed using asymptotic notation. Some of the more commonly used notations are  $f(n) = O(T(n))$  ( $f$  asymptotically bounded above by  $T$  up to a multiplicative constant),  $f(n) = \Theta(T(n))$  (asymptotically bounded above and below by  $T$  up to multiplicative constants  $k_1$  and  $k_2$ ) and  $f(n) = \Omega(T(n))$  ( $f$  asymptotically bounded below by  $T$  up to a multiplicative constant).

Complexity classes are groupings of problems by hardness, namely the scaling of the cost of solving the problem with respect to some resource, as a function of the “size” of an instance of the problem. The most well-known ones being described informally in the following lines:

- (1) *P*: problems that can be solved in time polynomial with respect to input size by a deterministic classical computer.
- (2) *NP*: a problem is said to be in *NP*, if the problem of verifying the correctness of a proposed solution lies in *P*, irrespective of the difficulty of obtaining a correct solution.
- (3) *PH*: stands for Polynomial Hierarchy. This class is a generalization of *NP* in the sense that it contains all the problems which one gets if one starts with a problem in the class *NP* and adds additional layers of complexity using quantifiers, i.e. there exists ( $\exists$ ) and for all ( $\forall$ ). As we add more quantifiers to a problem, it becomes more complex and is placed higher up in the polynomial hierarchy.
- (4) *BPP*: stands for Bounded-error Probabilistic Polynomial-time. A problem is said to be in *BPP*, if it can be solved in time polynomial in the input size by a probabilistic classical computer.
- (5) *BQP*: stands for Bounded-error Quantum Polynomial-time. Such problems can be solved in time polynomial in the input size by a quantum computer.
- (6) *PSPACE*: stands for Polynomial Space. The problems in *PSPACE* can be solved in space polynomial in the input size by a deterministic classical computer.
- (7) *EXPTIME*: stands for Exponential Time. The problems in *EXPTIME* can be solved in time exponential in the input size by a deterministic classical computer.
- (8) *QMA*: stands for Quantum Merlin Arthur and is the quantum analog of the complexity class *NP*. A problem is said to be in *QMA*, if given a “yes” as an answer, the solution can be verified in time polynomial (in the input size) by a quantum computer.

Widely believed containment relations for some of the complexity classes are shown in a schematic way in Fig. 1.

To understand the internal structure of complexity classes, the idea of “reductions” can be quite useful. One says that problem *A* is reducible to problem *B* if a method for solving *B* implies a method for solving *A*; one denotes the same by  $A \leq B$ . It is a common practice to assume the reductions as polynomial-time reductions. Intuitively, it could be thought as solving *B* is at least as difficult as solving *A*. Given a class *C*, a problem *X* is said to be *C*-hard if every problem in class *C* reduces to *X*. We say a problem *X* to be *C*-complete if *X* is *C*-hard and also a member of *C*. The *C*-complete problems could be understood as capturing the difficulty of class *C*, since any algorithm which solves one *C*-complete problem can be used to solve any problem in *C*.

A canonical example of a problem in the class *BQP* is integer factorization, which can be solved in polynomial time by a quantum computer using Shor’s factoring algorithm (Shor, 1994). However, no classical polynomial-time algorithm is known for the aforementioned problem. While analyzing the performance of algorithms, it is prudent to perform complexity-theoretic sanity checks. For example, though quantum computers are believed to be powerful, they are not widely expected to be able to solve *NP*-Complete problems, such as the travelling-salesman problem, in polynomial time.

## B. Experimental progress

Experimental progress in quantum computation can be measured by different figures of merit. On one side, the number of physical qubits available must exceed a certain threshold to unravel applications beyond the capabilities of a classical computer. Conversely, there exist several classical techniques capable of efficiently simulating certain quantum many-body systems. The success of some of these techniques, such as Tensor Networks (Orús, 2014; Verstraete et al., 2008), rely on the efficient representation of states that are not highly entangled (Vidal, 2003, 2004). With the advent of universal quantum computers, one would expect to be able to generate highly entangled quantum states which are not efficiently expressible via tensor network techniques. Hence, one imminent and practical direction towards demonstrating quantum advantage over classical machines is to go beyond the state space that the best supercomputers can reach. In particular, one may focus on a region of Hilbert space whose states the current best classical computing algorithms have difficulty representing efficiently. Alternatively, one

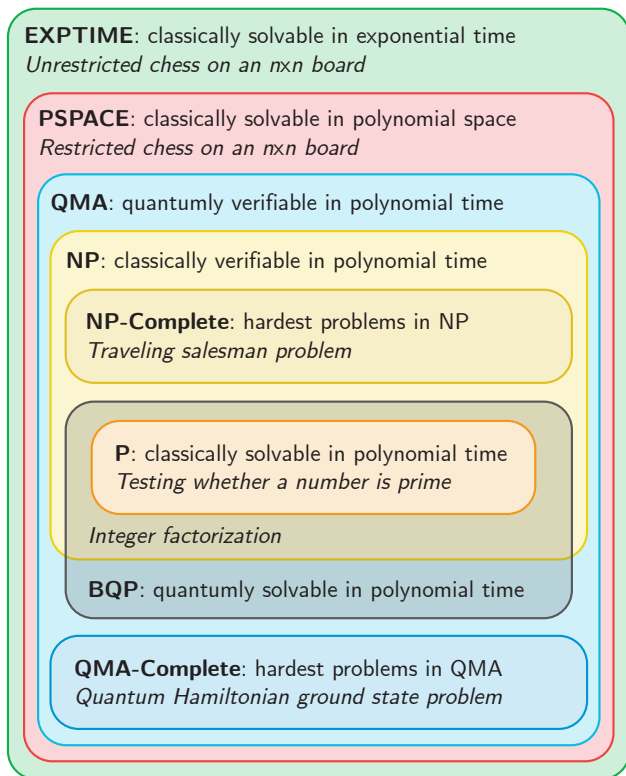


Figure 1 An illustrative picture of some relevant complexity classes and their widely believed containment relations. Example problems for a few of the complexity classes have been mentioned. The word “restricted” for the chess example refers to a polynomial upper bound on the number of moves. As a word of caution, these containment relations are suggestive. They have not been mathematically proven for many of the complexity classes in the figure—a typical open problem being whether  $P$  is equal to  $NP$ .

might tackle certain computational tasks which are believed to be intractable with any classical computer.

Two recent experimental ventures exhibit this focus. In 2019, the Google AI Quantum team implemented an experiment with the 53-qubit Sycamore chip (Arute *et al.*, 2019), in which single-qubit gate fidelities of 99.85% and two-qubit gate fidelities of 99.64% were attained on average. Quantum advantage was demonstrated against the best current classical computers in the task of sampling the output of a pseudo-random quantum circuit. A later work by the same Google AI Quantum team implemented a quantum-chemistry experiment (Arute *et al.*, 2020a) to demonstrate that the *Sycamore* chip is a fully programmable quantum processor with high fidelity quantum gates. However, the latter experiment was done on the 12-qubit subset of the former 53-qubit processor, showing the challenges to run certain purpose-specific quantum algorithms. An additional quantum advantage experiment was carried out by Jian-Wei Pan’s group using a *Jiuzhang* photonic quantum computer performing

Gaussian boson sampling (GBS) with 50 indistinguishable single-mode squeezed states (Zhong *et al.*, 2020) (see Sec. III.B for brief explanation of GBS). Here, quantum advantage was seen in sampling time complexity of a Torontonian (Quesada *et al.*, 2018) matrix, which scales exponentially with output photon clicks.

There are several quantum computing platforms that researchers are actively developing at present in order to achieve scalable practical universal quantum computers. By “universal”, is meant that such a quantum computer can perform native gate operations that allow it to easily and accurately approximate any unitary gate. Two of the most promising platforms, superconducting circuits and quantum optics, have already been mentioned; In addition to these, trapped-ion devices are also leading candidates (see Sec. VI.B), where scaling up to 2D architecture is being pushed forward by pioneers in the field (see (Wan *et al.*, 2020) and references therein).

Scientists and engineers developing alternate quantum computing platforms are trying to achieve similar feats described above. These alternate quantum devices might not necessarily possess universal quantum gate sets, as many are built to solve specific problems. Notably, coherent Ising machines (Inagaki *et al.*, 2016; Marandi *et al.*, 2014; McMahon *et al.*, 2016; Utsunomiya *et al.*, 2011; Wang *et al.*, 2013) based on mutually coupled optical parametric oscillators are promising and have shown success in solving instances of hard combinatorial optimization problems. Recently, it has been shown that the efficiency of these machines can be improved with error detection and correction feedback mechanisms (Kako *et al.*, 2020). The reader is advised to refer to the recent review article (Yamamoto *et al.*, 2020) for an in-depth discussion about coherent Ising machines.

### C. NISQ and near-term

The experimental state-of-the-art and the demand for QEC have encouraged the development of innovative algorithms capable of reaching the long-expected *quantum advantage*, i.e. a purpose-specific computation that involves a quantum device and that can not be performed classically with a reasonable amount of time and energy resources. The fact that physical qubits possess limited coherence times as well as gate implementations and qubits readouts are imperfect has lead to the coinage of the term *noisy* to refer to non-QEC qubits, to distinguish them from the QEC ones, called *logical*.

To cluster all these quantum algorithms specially developed to be run on current quantum computing hardware or those which could be developed in the next few years, the term *near-term* quantum computation has been coined. It is important to note that NISQ is a hardware-focused definition, and does not necessarily imply a temporal connotation. *near-term algorithms*, how-



ever, refers to those algorithms designed for quantum devices available in the next few years and carries no explicit reference to the absence of QEC. In other words, the NISQ era corresponds to the period when only a few hundred noisy qubits are available. In contrast, the near-term era involves any quantum computation performed in the next few years.

#### D. Scope of the review

This review aims to accomplish three main objectives. The first is to provide a proper compilation of the available algorithms suited for the NISQ era and which can deliver results in the near-term. We present a summary of the crucial tools and techniques that have been proposed and harnessed to design such algorithms. The second objective is to discuss the implications of these algorithms in various applications such as quantum machine learning, quantum chemistry, and combinatorial optimization, etc. Finally, the third objective is to give some perspectives on potential future developments given recent quantum hardware progress.

Most of the current NISQ algorithms rely on harnessing the power of quantum computers in a hybrid quantum-classical arrangement. Such algorithms delegate the classically difficult part of some computation to the quantum computer and perform the classically tractable part on some sufficiently powerful classical device. These algorithms variationally update the parameters of a parametrized quantum circuit and hence are referred to as *Variational Quantum Algorithms* (VQA) (Cao *et al.*, 2019; Cerezo *et al.*, 2020b; Endo *et al.*, 2020a; McArdle *et al.*, 2020) (sometimes also called Hybrid Quantum-Classical Algorithms). The first proposals of VQA were the Variational Quantum Eigensolver (VQE) (McClean *et al.*, 2016; Peruzzo *et al.*, 2014; Wecker *et al.*, 2015), originally proposed to solve quantum chemistry problems, and the Quantum Approximate Optimization Algorithm (QAOA) (Farhi *et al.*, 2014), proposed to solve combinatorial optimization problems. These two algorithms may be thought of as the parents of the whole VQA family. We cover their main blocks in Sec. II.

Other quantum computing paradigms propose different kinds of algorithms. Inspired by and hybridized with analog approaches, we present their fundamental properties in Sec. III. These include quantum annealing, digital-analog quantum computation, Gaussian Boson Sampling and analog quantum computation. Since most of these other techniques have been covered in other reviews, we center the present one’s efforts upon detailed coverage of VQA.

In Sec. IV, we examine the theoretical and experimental challenges facing these algorithms and the methods developed to best exploit them. We include the theoretical guarantees that some of these algorithms may lay

claim to and techniques to mitigate the errors coming from the use of noisy quantum devices.

Section V presents the large variety of applications that VQA introduces. Techniques to benchmark, compare and quantify current quantum devices’ performance is presented in Sec. VI. Like any other computational paradigm, quantum computing requires a language to establish human-machine communication. We explain different levels of quantum programming and provide a list of open-source quantum software tools in Sec. VII. Finally, we conclude this review in Sec. VIII by highlighting the increasing community involvement in this field and by presenting the NISQ, near-term and long-term goals of quantum computational research.

## II. BUILDING BLOCKS OF VARIATIONAL QUANTUM ALGORITHMS

A variational algorithm comprises several modular components that can be readily combined, extended and improved with developments in quantum hardware and algorithms. Chief among these are the cost function, the equation to be variationally minimized; one or several *Parameterized Quantum Circuit* (PQC), the quantum circuit unitaries whose parameters are manipulated in the minimization of the cost function; the measurement scheme, which extracts the expectation values needed to evaluate the cost function; and the classical optimization, the method used to obtain the optimal circuit parameters that minimize the cost function. In the following subsections, we will define each of these pieces, presented diagrammatically in Fig. 2.

### A. Objective function

The Hamiltonian encodes information about a given system, and it naturally arises in the description of many physical systems, such as in quantum chemistry problems. Its expectation value yields the energy of a quantum state, which is often used as the minimization target of a VQA. Other problems not related to real physical systems can also be encoded into a Hamiltonian form, thereby opening a path to solve them on a quantum computer. Hamiltonian operators are not all that can be measured on quantum devices; in general, any expectation value of a function written in an operational form can be extracted from a quantum computer. After the Hamiltonian of a problem has been determined, or the target operator has been defined, it must be decomposed into a set of particular operators that can be measured with a quantum processor. Such a decomposition, which is further discussed in Sec. II.A.1, is an important step of many quantum algorithms in general and of VQA in particular.

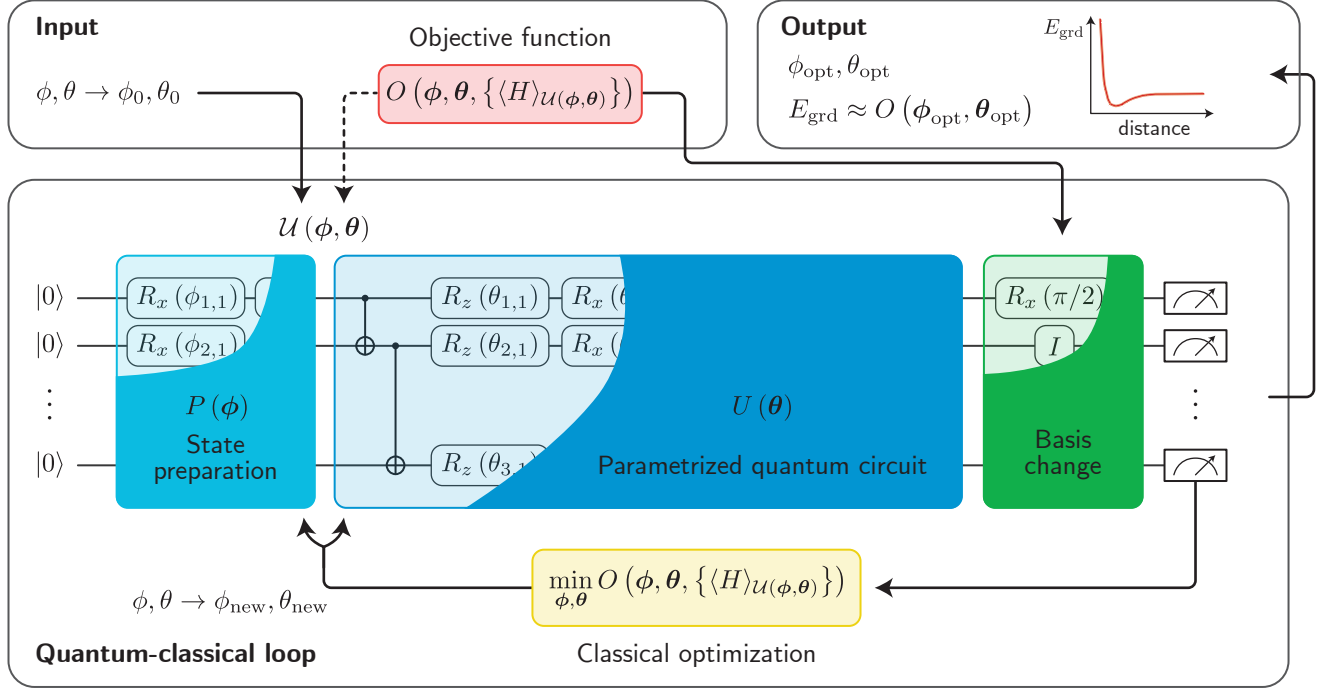


Figure 2 Diagrammatic representation of a Variational Quantum Algorithm (VQA). A VQA workflow can be divided into four main components: *i*) the objective function  $O$  that encodes the problem to be solved; *ii*) the parameterized quantum circuit (PQC), in which its parameters  $\theta$  are tuned to minimize the objective; *iii*) the measurement scheme, which performs the basis changes and measurements needed to compute expectation values  $\langle H \rangle$  that are used to evaluate the objective; and *iv*) the classical optimizer that minimizes the objective and proposes a new set of variational parameters. The PQC can be defined heuristically, following hardware-inspired ansätze, or designed from the knowledge about the problem Hamiltonian  $H$ . It can also include a state preparation unitary  $P(\phi)$  which situates the algorithm to start in a particular region of parameter space. Inputs of a VQA are the circuit ansatz  $\mathcal{U}(\theta, \phi)$  and the initial parameter values  $\theta_0, \phi_0$ . Outputs include optimized parameter values  $\theta_{\text{opt}}, \phi_{\text{opt}}$  and the minimum of the objective.

Within a VQA, one has access to measurements on qubits whose outcome probabilities are determined by the prepared quantum state. To begin, consider only measurements on individual qubits in the standard computational basis and denote the probability to measure qubit  $q$  in state  $|0\rangle$  by  $p_0^q$ , where the qubit label  $q$  will be omitted whenever possible. The central element of a variational quantum algorithm is a parametrized cost or *objective function*  $O$  subject to a classical optimization algorithm

$$\min_{\theta} O(\theta, \{p_0(\theta)\}). \quad (1)$$

The objective function  $O$  and the measurement outcomes  $p_0$ , of one or many quantum circuit evaluations depend on the set of parameters  $\theta$ .

In practice it is often inconvenient to work with the probabilities of the measurement outcomes directly when evaluating the objective function. Higher level formulations employ expectation values

$$\langle H \rangle_{U(\theta)} \equiv \langle 0 | U^\dagger(\theta) H U(\theta) | 0 \rangle \quad (2)$$

of qubit Hamiltonians  $H$ , describing measurements on

the quantum state generated by the unitary  $U(\theta)$ , instead of using the probabilities for the individual qubit measurements directly. See Eq. (5) for the decomposition of arbitrary observables into basic measurements of Pauli strings that can be transformed into basic measurements in the standard computational basis (see Sec. II.C). Restricting ourselves to expectation values instead of pure measurement probabilities, the objective function is

$$\min_{\theta} O(\theta, \{\langle H \rangle_{U(\theta)}\}). \quad (3)$$

The formulation in terms of expectation values of qubit Hamiltonians often allows for more compact definitions of the objective function. For the original VQE (Peruzzo *et al.*, 2014) and QAOA (Farhi *et al.*, 2014) it can, for example, be described as a single expectation value  $\min_{\theta} \langle H \rangle_{U(\theta)}$ , where the differences solely appear in the specific form and construction of the qubit Hamiltonian.

The choice of the objective function is crucial in a VQA to achieve the desired convergence. Vanishing gradient issues during the optimization, known as barren plateaus, are dependent on the cost function used (Cerezo *et al.*, 2020c) (see Sec. IV.B for details).

Following, we detail how to express the objective function in terms of expectation values of the problem Hamiltonian or other observables. Details on how to decompose and measure these expectations values are given in Sec. II.C.

### 1. Pauli strings

In order to be executable on a general quantum computer, it is sufficient to express the Hamiltonian as a linear combination of primitive tensor products of Pauli matrices  $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$  – so-called *Pauli strings*. For NISQ devices made up of qubits, Hamiltonians are commonly decomposed into a linear combination of tensor products of Pauli matrices  $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ . We refer to these tensor products as *Pauli strings*  $\hat{P} = \bigotimes_{j=1}^n \hat{\sigma}_j$ , where  $n$  is the number of qubits,  $\hat{\sigma} \in \{\hat{I}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$  and  $\hat{I}$  the identity operator. The Hamiltonian is given by

$$H = \sum_{k=1}^M c_k \hat{P}_k, \quad (4)$$

where  $c_k$  is a complex coefficient of the  $k$ -th Pauli string and the number of Pauli strings  $M$  in the expansion depends on the operator at hand. An expectation value in the sense of Eq. (2) then naturally decomposes into a set of expectation values, each defined by a single Pauli string

$$\langle H \rangle_{\mathcal{U}} = \sum_{k=1}^M c_k \langle \hat{P}_k \rangle_{\mathcal{U}}. \quad (5)$$

Examples of Hamiltonian objectives include molecules (by means of some fermionic transformation to Pauli strings, as detailed in Sec. V.A), condensed matter models written in terms of spin chains, or optimization problems encoded into a Hamiltonian form (see Sec. V.C).

### 2. Fidelity

Instead of searching for an unknown ground state of a given Hamiltonian, which is the canonical application of VQE, some algorithms have the goal of optimizing a specific target state  $|\Psi\rangle$  with the state obtained from the PQC  $\mathcal{U}(\theta)$ ,  $|\Psi\rangle_{\mathcal{U}(\theta)}$ . The goal is then to maximize the fidelity between the prepared state and the target state

$$F(\Psi, \Psi_{\mathcal{U}(\theta)}) \equiv |\langle \Psi | \Psi_{\mathcal{U}(\theta)} \rangle|^2, \quad (6)$$

which is equivalent to the expectation value over the projector  $\hat{\Pi}_{\Psi} = |\Psi\rangle\langle\Psi|$ . The state preparation objective is then the minimization of the infidelity  $1 - F(\Psi, \Psi_{\mathcal{U}(\theta)})$  or just the negative fidelity

$$\max_{\theta} F(\Psi, \Psi_{\mathcal{U}(\theta)}) = \min_{\theta} (-\langle \hat{\Pi}_{\Psi} \rangle_{\mathcal{U}(\theta)}). \quad (7)$$

To calculate the fidelity of the trial state with a state prepared by a known and efficiently implementable unitary, one can use the following approach to avoid the measurement of the projector

$$F(\Psi, \Psi_{\mathcal{U}(\theta)}) = \langle \Pi_0 \rangle_{\mathcal{U}_{\Psi}^{\dagger} \mathcal{U}(\theta)}. \quad (8)$$

Objective formulations over fidelities are prominent within state preparation algorithms in quantum optics (Kottmann *et al.*, 2020c; Krenn *et al.*, 2020a,b), excited state algorithms (Kottmann *et al.*, 2020b; Lee *et al.*, 2018) and quantum machine learning (Benedetti *et al.*, 2019a; Cheng *et al.*, 2018; Pérez-Salinas *et al.*, 2020a) (see also Sec. V.B for more references and details). In these cases, the fidelities are often defined in respect to computational basis states  $e_i$ , such that  $F_{e_i} = |\langle \Psi(\theta) | e_i \rangle|^2$ . They are then often referred to as Born's rule for the probability of basic measurements.

### 3. Other objective functions

Hamiltonian expectation values are not the only objective functions that are used in VQAs. Any cost function that is written in an operational form can constitute a good choice. One such example is the conditional value-at-risk (CVaR). Given the set of energy measurement samples  $\{E_1, \dots, E_M\}$  arranged in a non-decreasing order, instead of using the expectation value from Eq. (2) as the objective function, it was proposed to use (Barkoutsos *et al.*, 2020)

$$\text{CVaR}(\alpha) = \frac{1}{\lceil \alpha M \rceil} \sum_{k=1}^{\lceil \alpha M \rceil} E_k, \quad (9)$$

which measures the expectation value of the  $\alpha$ -tail of the energy distribution. Here,  $\alpha \in (0, 1]$  is the confidence level. The  $\text{CVaR}(\alpha)$  can be thought of as a generalization of the sample mean ( $\alpha = 1$ ) and the sample minimum ( $\alpha \rightarrow 0$ ).

Another proposal (Li *et al.*, 2020) is to use the Gibbs objective function

$$G = -\ln \langle e^{-\eta H} \rangle, \quad (10)$$

which is the cumulant generating function of the energy. The variable  $\eta > 0$  is a hyperparameter to be tuned. For small  $\eta$ , the Gibbs objective function reduces to the mean energy in Eq. (2). Since both the CVaR and the Gibbs objective function are more general than the mean energy, it is not surprising that they perform at least as well as the mean energy  $\langle H \rangle$ . Empirically, both measures have been shown to outperform  $\langle H \rangle$  for certain combinatorial optimization problems.

## B. Parameterized quantum circuits

Following the objective function, the next essential constituent of a VQA is the quantum circuit that prepares the state that best meets the objective. It is generated by means of a unitary operation that depends on a series of parameters, the PQC. In this subsection, we describe how this quantum circuit is defined and designed. Finding the parameter values of the PQC that deliver the optimal unitary, on the other hand, is the task of the classical optimization subroutine described in Sec. II.D.

The PQC is applied from an initial state  $|\Psi_0\rangle$  that can be the state at which the quantum device is initialized or a particular choice motivated by the problem at hand. Similarly, the initial parameters and the structure of the circuit are unknown *a priori*. However, knowledge about the particular problem can be leveraged to predict and postulate its structure.

We define the state after application of the PQC as

$$|\Psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\Psi_0\rangle, \quad (11)$$

where  $\boldsymbol{\theta}$  are the variational parameters.

Typically, the initial state  $|\Psi_0\rangle$  is a product state with all qubits in the same computational basis state  $|00\dots 0\rangle = |0\rangle^{\otimes n}$ , where  $n$  is the number of qubits. However, in some VQAs, it is convenient to prepare that state in a particular form before applying the PQC. The state preparation operation would then depend on some other unitary operation  $P$  that may depend on variational parameters  $\boldsymbol{\phi}$

$$|\Psi_0\rangle = P(\boldsymbol{\phi})|0\rangle^{\otimes n}. \quad (12)$$

This unitary  $P$  can be absorbed in a redefinition of the variational circuit unitary

$$|\Psi(\boldsymbol{\theta}, \boldsymbol{\phi})\rangle = U(\boldsymbol{\theta})P(\boldsymbol{\phi})|0\rangle^{\otimes n} = \mathcal{U}(\boldsymbol{\theta}, \boldsymbol{\phi})|0\rangle^{\otimes n}. \quad (13)$$

As for the design of the variational circuit ansatz, any known property about the final state can be used to obtain the initial guess. For instance, if we expect that the final state solution will contain all elements of the computational basis, or if we want to exploit a superposition state to seed the optimization, an initial state choice may be  $P|0\rangle^{\otimes n} = H_d^{\otimes n}|0\rangle^{\otimes n}$ , where  $H_d$  is the Hadamard gate. Applied to all qubits,  $H_d$  generates the even superposition of all basis states, i. e.

$$|D\rangle = H_d^{\otimes n}|0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{i=1}^n |e_i\rangle, \quad (14)$$

where  $|e_i\rangle$  are the computational basis states.

$P$  can also be used to encode the information about the problem, as is the case in many quantum machine learning algorithms. There are different strategies to feed the data into the quantum circuit, all of which rely on

encoding the data points  $\mathbf{x}$  into the angles of the quantum gates. Encoding strategies are discussed further in Sec. V.B.1.

In quantum chemistry algorithms, the initial state usually corresponds to the Hartree-Fock approximation, which simplifies the  $P$  operation (see Sec. V.A for details). The choice of a good initial state will allow the variational algorithm to start the search over the parameters  $\boldsymbol{\theta}$  in a region of the parameter space that is close to an optimum, helping the algorithm converge towards the solution.

The choice of the  $U$  ansatz greatly affects the performance of a VQA. From the perspective of the problem, the ansatz influences both the convergence speed and the closeness of the final state to a state that optimally solves the problem. On the other hand, the quantum hardware on which the VQA is executed has to be taken into account: Deeper circuits are more susceptible to errors, and some ansatz gates are costly to compose with native gates. Accordingly, most of the ansätze developed to date are classified either as more *problem-inspired* or more *hardware efficient*, depending on their structure and application. In the following paragraphs, we present common constructions of these ansätze.

### 1. Problem-inspired ansätze

An arbitrary unitary operation can be generated by an Hermitian operator  $\hat{g}$  which, physically speaking, defines an evolution in terms of the  $t$  parameter,

$$G(t) = e^{-i\hat{g}t}. \quad (15)$$

As an example, the generator  $\hat{g}$  can be a Pauli matrix  $\hat{\sigma}_i$  and thus,  $G(t)$  becomes a single-qubit rotation of the form

$$R_k(\theta) = e^{-i\frac{\theta}{2}\hat{\sigma}_k} = \cos(\theta/2)I - i\sin(\theta/2)\hat{\sigma}_k, \quad (16)$$

with  $t = \theta$  and  $\hat{g} = \frac{1}{2}\hat{\sigma}_k$ , corresponding to the spin operator.

From a more abstract viewpoint, those evaluations can always be described as time evolution of the corresponding quantum state, so that the generator  $\hat{g}$  is often just referred to as Hamiltonian. Note however, that the Hamiltonian generating evolution does not necessarily need to be the operator that describes the energy of the system of interest. In general, such generators can be decomposed into Pauli strings in the form of Eq. (4).

Within so-called *problem-inspired* approaches, evolutions in the form of Eq. (15), with generators derived from properties of the system of interest, are used to construct the parametrized quantum circuits. The unitary coupled-cluster approach (see below), mostly applied for quantum chemistry problems, is one prominent approach. The generators then are elementary fermionic excitations, as shown in Eq. (20).



The *Suzuki-Trotter* expansion or decomposition (Suzuki, 1976) is a general method to approximate a general, hard to implement unitary in the form of Eq. (15) as a function of the  $t$  parameter. This can be done by decomposing  $\hat{g}$  into a sum of non-commuting operators  $\{\hat{o}_k\}_k$ , with  $\hat{g} = \sum_k c_k \hat{o}_k$  and some coefficients  $c_k$ . The operators  $\hat{o}_k$  are chosen such that the evolution unitary  $e^{-i\hat{o}_k t}$  can be easily implemented, for example as Pauli strings. The full evolution over  $t$  can now be decomposed into integer  $m$  equal-sized steps

$$e^{-i\hat{g}t} = \lim_{m \rightarrow \infty} \left( \prod_k e^{-i \frac{c_k \hat{o}_k t}{m}} \right)^m. \quad (17)$$

For practical purposes, the time evolution can be approximated by a finite number  $m$ . When Pauli strings are used, this provides a systematic method to decompose an arbitrary unitary, generated by  $\hat{g}$ , into a product of multi-qubit rotations  $e^{-i \frac{\hat{P}_k t}{m}}$ , that can themselves be decomposed into primitive one and two qubit gates.

Knowledge about the physics of the particular Hamiltonian to be *trotterized* can reduce substantially the number of gates needed to implement this method. For instance, in (Kivlichan *et al.*, 2018), it is shown that by using fermionic swap gates, it is possible to implement a Trotter step for electronic structure Hamiltonians using first-neighbour connectivity circuits with  $N^2/2$  two-qubit gates width and  $N$  depth, where  $N$  is the number of spin orbitals. They also show that implementing arbitrary Slater determinants can be done efficiently with  $N/2$  gates of circuit depth.

*Unitary Coupled Cluster.* Historically, problem-inspired ansätze were proposed and implemented first. They arose from the quantum chemistry-specific observation that the unitary coupled cluster (UCC) ansatz (Taube and Bartlett, 2006), which adds quantum correlations to the Hartree-Fock approximation, is inefficient to represent on a classical computer (Yung *et al.*, 2014). Leveraging quantum resources, the UCC ansatz was instead realized as a PQC on a photonic processor (Peruzzo *et al.*, 2014). It is constructed from the parametrized cluster operator  $T(\theta)$  and acts on the Hartree-Fock ground state  $|\Psi_{\text{HF}}\rangle$ :

$$|\Psi(\theta)\rangle = e^{T(\theta)-T(\theta)^\dagger} |\Psi_{\text{HF}}\rangle. \quad (18)$$

The cluster operator is given by

$$\begin{aligned} T(\theta) &= T_1(\theta) + T_2(\theta) + \dots \\ T_1(\theta) &= \sum_{\substack{i \in \text{occ} \\ j \in \text{virt}}} \theta_i^j \hat{a}_j^\dagger \hat{a}_i \\ T_2(\theta) &= \sum_{\substack{i_1, i_2 \in \text{occ} \\ j_1, j_2 \in \text{virt}}} \theta_{i_1, i_2}^{j_1, j_2} \hat{a}_{j_2}^\dagger \hat{a}_{i_2} \hat{a}_{j_1}^\dagger \hat{a}_{i_1}, \end{aligned} \quad (19)$$

where higher-order terms follow accordingly (O'Malley *et al.*, 2016). The operator  $\hat{a}_k$  is the annihilation operator of the  $k$ -th Hartree-Fock orbital, and the sets *occ*

and *virt* refer to the occupied and unoccupied Hartree-Fock orbitals. Due to their decreasing importance, the series is usually truncated after the second or third term. The ansatz is termed *UCCSD* or *UCCSDT*, respectively, referring to the inclusion of single, double, and triple excitations from the Hartree-Fock ground state. The  $k$ -UpCCGSD approach restricts the double excitations to pairwise excitations but allows  $k$  layers of the approach (Lee *et al.*, 2018). After mapping to Pauli strings as described in Sec. II.A.1, the ansatz is converted to a PQC usually via the Trotter expansion Eq. (17).

In its original form, the UCC ansatz faces several drawbacks in its application to larger chemistry problems as well as to other applications: For strongly correlated systems, the widely proposed UCCSD ansatz is expected to have insufficient overlap with the true ground state and results typically in large circuit depths (Grimsley *et al.*, 2019b; Lee *et al.*, 2018). Consequently, improvements and alternative ansätze are proposed to mitigate these challenges. We restrict our discussion here to provide a short overview of alternative ansatz developments.

*Factorized Unitary Coupled-Cluster and Adaptive Approaches.* The non commuting nature of the fermionic excitation generators, given by the cluster operators Eq. (20) leads to difficulties in decomposing the canonical UCC ansatz Eq. (18) into primitive one- and two-qubit unitaries. First approaches employed the Trotter decomposition Eq. (17) using a single step (McClean *et al.*, 2016; Romero *et al.*, 2018). The accuracy of the so obtained factorized ansatz depends however on the order of the primitive fermionic excitations (Grimsley *et al.*, 2019a; Izmaylov *et al.*, 2020).

Alternative approaches propose to use factorized unitaries, constructed from primitive fermionic excitations, directly (Evangalista *et al.*, 2019; Izmaylov *et al.*, 2020). Adaptive approaches, are a special case of a factorized ansatz, where the unitary is iteratively grown by subsequently screening and adding primitive unitary operators from a predefined operator pool. The types of operator pools can be divided into two classes: Adapt-VQE (Grimsley *et al.*, 2019b), that constructs the operator pool from primitive fermionic excitations, and Qubit-Coupled-Cluster (Ryabinkin *et al.*, 2018b) that uses Pauli Strings.

In both original works, the screening process is based on energy gradients with respect to the prospective operator candidate. Since this operator is the trailing part of the circuit, the gradient can be evaluated through the commutator of the Hamiltonian with the generator of that operator. In contrast to commutator based gradient evaluation, direct differentiation, as proposed in (Kottmann *et al.*, 2021) allows gradient evaluations with similar cost as the original objective and generalizes the approach by allowing screening and insertion of operators at arbitrary positions in the circuit. This is, for

example, necessary for excited state objectives as discussed in Sec. V.A.4. Extended approaches include iterative approaches (Ryabinkin *et al.*, 2020), operator pool construction from involutory linear combination of Pauli strings (Lang *et al.*, 2020), Pauli string pools from decomposed fermionic pools (Tang *et al.*, 2019), mutual information based operator pool reduction (Zhang *et al.*, 2021), measurement reduction schemes based on the density matrix reconstruction (Liu *et al.*, 2020a), and external perturbative corrections (Ryabinkin *et al.*, 2021).

*Variational Hamiltonian Ansatz.* Motivated by adiabatic state preparation, the Variational Hamiltonian Ansatz (VHA) was developed to reduce the number of variational parameters and accelerate the convergence (McClean *et al.*, 2016; Wecker *et al.*, 2015). Instead of the Hartree-Fock operators, the terms of the fermionic Hamiltonian itself are used to construct the PQC. For this purpose, the fermionic Hamiltonian  $H$  is written as a sum of  $M$  terms  $H = \sum_i \hat{h}_i$ . Which parts of the Hamiltonian are grouped into each term  $\hat{h}_i$  depends on the problem and there is a degree of freedom in the design of the algorithm. The PQC is then chosen as

$$U_{\text{VHA}} = \prod_{i=1}^M \exp(i\theta_i \hat{h}_i), \quad (20)$$

with the operators in the product ordered by decreasing  $i$ . The unitary corresponds to  $n$  short time evolutions under different parts of the Hamiltonian, where the terms  $\hat{h}_a$  of the Hamiltonian can be repeated multiple times. The initial state is chosen so that it is easy to prepare yet it is related to the Hamiltonian, for example the eigenstate of the diagonal part of  $H$ . The Fermi-Hubbard model with its few and simple interaction terms is proposed as the most promising near-term application of the method. However, it is also shown that the VHA can outperform specific forms of the UCCSD ansatz for strongly correlated model systems in quantum chemistry.

In Sec. V.A we discuss some VQE-inspired algorithms that also use adiabatic evolution to improve the performance of the algorithm.

*Quantum Approximate Optimization Algorithm.* One of the canonical NISQ era algorithms, with a promise to provide approximate solutions in polynomial time for combinatorial optimization problems, is the Quantum Approximate Optimization Algorithm (QAOA) (Farhi *et al.*, 2014). While QAOA can be thought of as a special case of VQA, it has been studied in depth over the years both empirically and theoretically, and it deserves special attention.

The cost function  $C$  depends on the bit strings that form the computational basis and is designed to encode

a combinatorial problem in those strings. With the computational basis vectors  $|e_i\rangle$ , one can define the problem Hamiltonian  $H_P$  as (see Sec. V.C.1 for an example)

$$H_P \equiv \sum_{i=1}^n C(e_i) |e_i\rangle, \quad (21)$$

and the mixing Hamiltonian  $H_M$  as

$$H_M \equiv \sum_{i=1}^n \hat{\sigma}_x^i, \quad (22)$$

The initial state in the QAOA algorithm is conventionally chosen to be the uniform superposition state  $|D\rangle$  from Eq. (14). The final quantum state is given by alternately applying  $H_P$  and  $H_M$  on the initial state  $p$ -times:

$$|\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \equiv e^{-i\beta_p H_M} e^{-i\gamma_p H_P} \dots e^{-i\beta_1 H_M} e^{-i\gamma_1 H_P} |D\rangle. \quad (23)$$

A quantum computer is used to evaluate the objective function

$$C(\boldsymbol{\gamma}, \boldsymbol{\beta}) \equiv \langle \Psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) | H_P(\boldsymbol{\gamma}, \boldsymbol{\beta}) | \Psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle, \quad (24)$$

and a classical optimizer is used to update the  $2p$  angles  $\boldsymbol{\gamma} \equiv (\gamma_1, \gamma_2, \dots, \gamma_p)$  and  $\boldsymbol{\beta} \equiv (\beta_1, \beta_2, \dots, \beta_p)$  until the objective function  $C$  is maximized, i.e.  $C(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \equiv \max_{\boldsymbol{\gamma}, \boldsymbol{\beta}} C(\boldsymbol{\gamma}, \boldsymbol{\beta})$ . Here,  $p$  is often referred to as the QAOA level or depth. Since the maximization at level  $p-1$  is a constrained version of the maximization at level  $p$ , the performance of the algorithm improves monotonically with  $p$  in the absence of experimental noise and infidelities.

In quantum annealing (see Sec. III.A), we start from the ground state of  $H_M$  and slowly move towards the ground state of  $H_P$  by slowly changing the Hamiltonian. In QAOA, we instead alternate between  $H_M$  and  $H_P$ . One can think of QAOA as a Trotterized version of quantum annealing. Indeed, the adiabatic evolution as used in quantum annealing can be recovered in the limit of  $p \rightarrow \infty$ .

In the original proposal of QAOA, the alternating unitary operators acting on the state are taken to be the time evolution of some problem and mixing Hamiltonians (see Eq. (23)). Hard constraints which must be satisfied are usually imposed as a penalty to the cost function. This might not be an efficient strategy in practice as it is still possible to obtain solutions which violate some of the hard constraints. Thus, building on previous work in quantum annealing (Hen and Sarandy, 2016; Hen and Spedalieri, 2016), it was proposed to encode the hard constraints directly in the mixing Hamiltonian (Hadfield *et al.*, 2017). This approach yields the main advantage of restricting the state evolution to the feasible subspace where no hard constraints are violated, which consequently speeds up the classical optimization routine to find the optimal angles. This framework was later generalized as the *Quantum Alternating Operator*

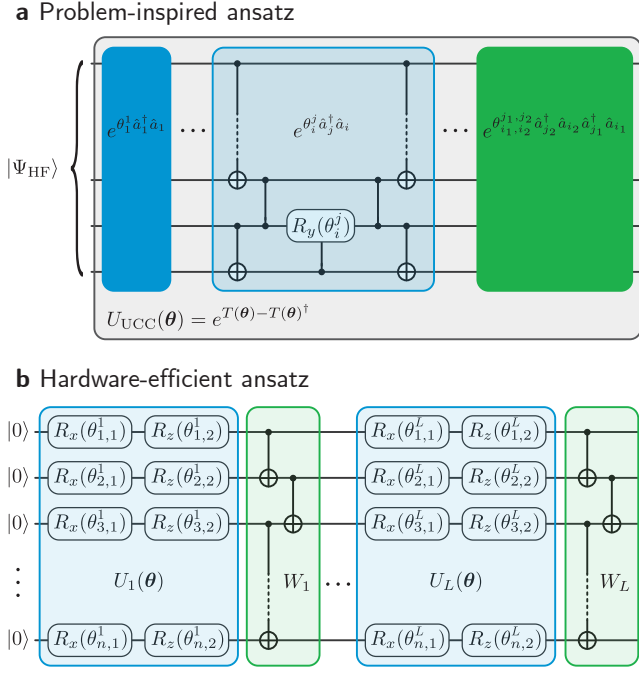


Figure 3 Exemplary problem-inspired and hardware-efficient ansätze. (a) Circuit of the Unitary Coupled Cluster ansatz with a detailed view of a fermionic excitation as discussed in (Yordanov *et al.*, 2020). (b) Hardware-efficient ansatz tailored to a processor that is optimized for single-qubit  $x$ - and  $z$ -rotations and nearest-neighbor two-qubit CNOT gates.

*Ansatz* to consider phase-separation and mixing unitary operators ( $U_P(\gamma)$  and  $U_M(\beta)$  respectively) which need not originate from the time-evolution of a Hamiltonian (Hadfield *et al.*, 2019). Eq. (23) is thus re-written as

$$|\Psi(\gamma, \beta)\rangle = U_M(\beta_p)U_P(\gamma_p)\cdots U_M(\beta_1)U_P(\gamma_1)|D\rangle. \quad (25)$$

## 2. Hardware-efficient ansätze

Thus far, we have described circuit ansätze constructed from the underlying physics of the problem to be solved. Although it has been shown computationally that such ansätze can ensure fast convergence to a satisfying solution state. However, they can be challenging to realize experimentally. Quantum computing devices possess a series of experimental limitations that include, among others, a particular qubit connectivity, a restricted gate set, and limited gate fidelities and coherence times. Therefore, existing quantum hardware is not suited to implement the deep and highly connected circuits required for the UCC and similar ansätze for applications beyond basic demonstrations such as the  $H_2$  molecule (Moll *et al.*, 2018).

A class of *hardware-efficient ansätze* has been proposed to accommodate device constraints (Kandala *et al.*,

2017). The common trait of these circuits is the use of a limited set of quantum gates as well as a particular qubit connection topology. The gate set usually consists of one two-qubit entangling gate and up to three single-qubit gates. The circuit is then constructed from blocks of single-qubit gates and entangling gates, which are applied to multiple or all qubits in parallel. One sequence of a single-qubit and an entangling block is usually called a *layer*, and the ansatz circuit generally has multiple such layers. In this way, it enables exploration of a larger portion of the Hilbert space.

The unitarity of a hardware-efficient ansatz with  $L$  layers is usually given by

$$U(\theta) = \prod_{k=1}^L U_k(\theta_k) W_k, \quad (26)$$

where  $\theta = (\theta_1, \dots, \theta_L)$  are the variational parameters,  $U_k(\theta_k) = \exp(-i\theta_k V_k)$  is a unitary derived from a Hermitian operator  $V_k$ , and  $W_k$  represents non-parametrized quantum gates. Typically, the  $V_k$  operators are single-qubit rotation gates, i.e.  $V_k$  are Pauli strings acting locally on each qubit. In those cases,  $U_k$  becomes a product of combinations of single-qubit rotational gates, each one defined as in Eq. (16).  $W_k$  is an entangling unitary constructed from gates that are native to the architecture at hand, for example CNOT or CZ gates for superconducting qubits or XX gates for trapped ions (Krantz *et al.*, 2019; Wright *et al.*, 2019). Following this approach, the so-called *Alternating Layered Ansatz* is a particular case of these Hardware-efficient ansätze which consists of layers of single qubit rotations, and blocks of entangling gates that entangle only a local set of qubits and are shifted every alternating layer.

The choice of these gates, their connectivity, and their ordering influences the portion of the Hilbert space that the ansatz covers and how fast it converges for a specific problem. Some of the most relevant properties of hardware-efficient ansätze, namely expressibility, entangling capability and trainability are studied in Refs. (Nakaji and Yamamoto, 2020a; Sim *et al.*, 2019; Woitzik *et al.*, 2020) and further discussed in Sec. IV.C.

Instead of making a choice between the problem-inspired and hardware-efficient modalities, some PQC designers have chosen an intermediate path. One example is the use of an exchange-type gate, which can be implemented natively in transmons, to construct a PQC that respects the symmetry of the variational problem (Ganzhorn *et al.*, 2019; Sagastizabal *et al.*, 2019b). Such an ansatz leads to particularly small parameter counts for quantum chemistry problems such as the  $H_2$  and LiH molecules (Gard *et al.*, 2020). Another intermediate approach, termed QOCA for its inspiration from quantum optimal control, is to add symmetry-breaking unitaries into the conventional VHA circuit (Choquette *et al.*, 2020). The symmetry-breaking part of the circuit

is akin to a hardware-efficient ansatz. This modification enables excursions of the variational state into previously restricted sections of the Hilbert space, which is shown to yield shortcuts in solving fermionic problems.

### C. Measurement

The main task in most quantum algorithms is to gain information about the quantum state that has been prepared on the quantum hardware. In VQA, this translates to estimate the expectation value of the objective function  $\langle \hat{O} \rangle_{\mathcal{U}_\theta}$ . The most direct approach to estimate expectation values is to apply a unitary transformation on the quantum state to the diagonal basis of the observable  $\hat{O}$ . Then, one can read out the expectation value directly from the probability of measuring specific computational states corresponding to an eigenvalue of  $\hat{O}$ . Here, one needs to determine whether a measured qubit is in the  $|0\rangle$  or  $|1\rangle$  state. For experimental details on this task we refer to existing reviews, such as for superconducting qubits (Krantz *et al.*, 2019) or ion traps (Häffner *et al.*, 2008).

However, on NISQ devices, the transformation to the diagonal basis mentioned before can be an overly costly operation. As a NISQ friendly alternative, most observables of interest can be efficiently parameterized in terms of Pauli strings, as shown in Eq. (4) and Eq. (5). One can transform into the diagonal basis of the Pauli string via simple single-qubit rotations as shown below.

*Measurement of Pauli strings.* The expectation value of the  $\hat{\sigma}_z$  operator on a particular qubit can be measured by reading out the probabilities of the computational basis state  $\{|0\rangle, |1\rangle\}$

$$\langle \hat{\sigma}_z \rangle = 2p_0 - 1, \quad (27)$$

where  $p_0$  is the probability to measure the qubit in state  $|0\rangle$ . Measurements defined by  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$  can be defined similarly by transforming them into the  $\hat{\sigma}_z$  basis first. The transformation is given by primitive single-qubit gates

$$\hat{\sigma}_x = R_y^\dagger \left( \frac{\pi}{2} \right) \hat{\sigma}_z R_y \left( \frac{\pi}{2} \right) = H_d \hat{\sigma}_z H_d, \quad (28)$$

$$\hat{\sigma}_y = R_x^\dagger \left( \frac{\pi}{2} \right) \hat{\sigma}_z R_x \left( \frac{\pi}{2} \right) = S H_d \hat{\sigma}_z H_d S^\dagger, \quad (29)$$

where  $S = \sqrt{\hat{\sigma}_z}$  and  $H_d = (\hat{\sigma}_x + \hat{\sigma}_z)/\sqrt{2}$  is the Hadamard gate. As an example, to measure  $\hat{\sigma}_y$  on a quantum state  $|\Psi\rangle$ , we have to apply  $H_d S^\dagger$  on  $|\Psi\rangle$  and then measure the probability of  $p_0$  that the state  $|0\rangle$  occurs

$$\langle \hat{\sigma}_y \rangle = \langle \Psi | \hat{\sigma}_y | \Psi \rangle = \langle \Psi | S H_d \hat{\sigma}_z H_d S^\dagger | \Psi \rangle = 2p_0 - 1. \quad (30)$$

Arbitrary Pauli strings  $\hat{P}$ , with primitive Pauli operations  $\hat{\sigma}_{f(k)}$  on qubits  $k \in K$ , can then be measured by the same procedure on each individual qubit in

$$\begin{aligned} \langle \hat{P} \rangle_{\mathcal{U}} &= \langle \prod_{k \in K} \hat{\sigma}_{f(k)}(k) \rangle_{\mathcal{U}} \\ &= \langle \prod_{k \in K} \hat{\sigma}_z(k) \rangle_{\tilde{U}\mathcal{U}} = \prod_{k \in K} (2p_0(k) - 1), \end{aligned} \quad (31)$$

where  $\tilde{U}$  is a product of single qubit rotations according to Eq. (28) and Eq. (29) depending on the Pauli operations  $\hat{\sigma}_{f(k)}$  at qubit  $k$ .

For many problems, such as quantum chemistry-related tasks, the number of terms in the cost Hamiltonian to be estimated can become very large. A naive way of measuring each Pauli string separately may incur a prohibitively large number of measurements. Recently, various more efficient approaches have been proposed (see (Bonet-Monroig *et al.*, 2020) for an overview). The common idea is to group different Pauli strings that can be measured simultaneously together such that a minimal number of measurements needs to be performed. Several methods to accomplish this have been proposed.

Pauli strings that commute qubit-wise, i.e. the Pauli operators on each qubit commute, can be measured at the same time (McClean *et al.*, 2016). Finding the minimal number of groups can be mapped to the minimum clique cover problem which is NP-hard in general, but good heuristics exist (Verteletskyi *et al.*, 2020). One can collect mutually commuting operators and transform them into a shared eigenbasis, which adds an additional unitary transformation to the measurement scheme (Crawford *et al.*, 2019b,b; Gokhale *et al.*, 2019; Yen *et al.*, 2020). Combinations of single qubit and Bell measurements have been proposed as well (Hamamura and Imamichi, 2020).

Alternatively, one can use a method called unitary partitioning to linearly combine different operators into a unitary, and use the Hadamard test (which is outlined in the next paragraph) to evaluate it (Izmaylov *et al.*, 2019a; Zhao *et al.*, 2020a).

In (Izmaylov *et al.*, 2019b), the observables can be decomposed into so-called mean-field Hamiltonians, which can be measured more efficiently if one measures one qubit after the other, and uses information from previous measurement outcomes.

For specific problems such as chemistry and condensed matter systems, it is possible to use the structure of the problem to reduce the number of measurements (Cade *et al.*, 2020; Cai, 2020; Gokhale and Chong, 2019; Huggins *et al.*, 2019). In particular, in (Cai, 2020), where a Fermi-Hubbard model is studied using VQE, the number of measurements is reduced by considering multiple orderings of the qubit operators when applying the Jordan-Wigner transformation. In the context of quantum chemistry, the up-to-date largest reduction could be achieved by the Cartan subalgebra approach of (Yen and Izmaylov, 2020). All those kind of optimizations require



an understanding of the underlying problem, and are not applicable for every use of the VQE.

*Measurement of overlaps.* Several variational algorithms require the measurement of an overlap of a quantum state  $|\psi\rangle$  with unitary  $U$  in the form of  $\langle\psi|U|\psi\rangle$ . This overlap is in general not an observable and has both real and imaginary parts, although the absolute square can be measured directly as a fidelity Eq. (6). The Hadamard test can evaluate such a quantity on the quantum computer using a single ancilla (Miquel *et al.*, 2002). The idea is to apply a controlled  $U$  operation, with control on the ancilla and target  $U$  on the quantum state. Then, one can measure from the ancilla state both real and imaginary part of the overlap. A downside of this method is the requirement to be able to implement a controlled unitary, which may require too many resources on current quantum processors. Alternative methods to measure the overlap without the use of control unitaries have been proposed (Mitarai and Fujii, 2019). One idea is to decompose  $U$  into a sum of Pauli strings, and then to measure the expectation value of each Pauli string individually. Another approach is possible if  $U = \bigotimes_q U_q$  can be rewritten into a product of unitaries  $U_q$  that act locally on only a few qubits. Then, one can find via classical means the diagonalization of  $U_q = V_q^\dagger D V_q$ , with diagonal matrix  $D$  and  $V_q$  being a unitary. The overlap can be found by applying the  $V_q$  unitaries on the state  $|\psi\rangle$ , measure the outcomes in the computational basis and do post-processing of the results with the classically calculated eigenvalues of  $D$ .

*Shadow tomography.* A powerful method to determine the expectation value of observables composed of many Pauli strings using only few measurements is shadow tomography (Huang *et al.*, 2020c). The idea of this method is to measure the *classical shadow* of a quantum state. This is achieved by preparing the desired quantum state, then apply a random circuit from the Clifford group and sample from the circuit. What follows is performed on a classical computer: The reversed Clifford circuit is applied on the sampled bit strings and then the desired observable is calculated. Since Clifford circuits can be calculated in polynomial time on a classical computer, this is easily possible even for large circuits. These shadows allow one to approximate the expectation value of Pauli strings from very few measurements.

## D. Parameter optimization

In many VQAs, a core task that is allocated to the classical computer is the optimization of quantum circuit parameters. In principle, this problem is not different

from any multivariate optimization procedure and standard classical methods can be applied (Lavrijsen *et al.*, 2020). However, in the NISQ era, the coherence time is short, which means that complicated analytical gradient circuits can not be implemented. At the same time, objective measurements take a long time, which means that algorithms with few function evaluations should be favored. As a last criterion, the optimizer should be resilient to noisy data coming from current devices and precision on expectation values that is limited by the number of shots in the measurement. These three requirements make that certain existing algorithms are better suited for PQC optimization and are more commonly used, and that new algorithms are being developed specifically for PQC optimization. In this section, we first review two classes of optimization, gradient-based and gradient-free.

We also consider a set of *resource-aware* optimization methods and strategies that additionally minimize quantities associated with the quantum cost of optimization. While we reserve more detailed descriptions to the respective references, we highlight the main features and advantages for each optimization method/strategy.

### 1. Gradient-based approaches

A common approach to optimise an objective function  $f(\boldsymbol{\theta})$  is via its gradient, i.e. the change of the objective function with respect to a variation of its  $M$  parameters  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_M)$ . The gradient indicates the direction in which the objective function shows the greatest change. This is a local optimization strategy as one uses information starting from some initial parameter value  $\boldsymbol{\theta}^{(0)}$ . Then, one iteratively updates  $\boldsymbol{\theta}^{(t)}$  over multiple discrete steps  $t$ . A common update rule for each  $\theta_i$  is

$$\theta_i^{(t+1)} = \theta_i^{(t)} - \eta \partial_i f(\boldsymbol{\theta}), \quad (32)$$

or  $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \eta \nabla f(\boldsymbol{\theta})$ , where  $\eta$  is a small parameter called learning rate and

$$\partial_i \equiv \frac{\partial}{\partial \theta_i}, \quad \nabla = (\partial_1, \dots, \partial_M) \quad (33)$$

are the partial derivative with respect to the parameter  $\theta_i$  and the gradient vector, respectively, using Einstein notation. We will drop the optimization variable symbol (in this case,  $\theta$ ) when possible.

There are various ways of estimating the gradient on a quantum computer, i.e.  $\partial_i f(\boldsymbol{\theta})$  with  $f(\boldsymbol{\theta}) = \langle 0|U^\dagger(\boldsymbol{\theta})HU(\boldsymbol{\theta})|0\rangle$  (Romero *et al.*, 2018). The most relevant of them are detailed in the following paragraphs.

*Finite difference.* The finite difference method approximates the gradient of a function  $f(\boldsymbol{\theta})$  as follows:

$$\partial_i f(\boldsymbol{\theta}) \approx \frac{f(\boldsymbol{\theta} + \epsilon \mathbf{e}_i) - f(\boldsymbol{\theta} - \epsilon \mathbf{e}_i)}{2\epsilon}, \quad (34)$$

where  $\epsilon$  is a small number and  $\mathbf{e}_i$  is the unit vector with 1 as its  $i$ -th element and 0 otherwise. The smaller  $\epsilon$ , the closer the right-hand side of above formula is to the true value of the gradient. However, for small  $\epsilon$  the difference of the objective function  $f(\boldsymbol{\theta} + \epsilon \mathbf{e}_i) - f(\boldsymbol{\theta} - \epsilon \mathbf{e}_i)$  becomes small as well. As the objective function is commonly an expectation value sampled from the quantum device and it is only estimated with limited accuracy, smaller  $\epsilon$  require more samples taken from the quantum hardware to achieve a good estimation of the gradient.

*Parameter-shift rule.* The analytical gradient can be calculated on quantum hardware using the parameter-shift rule, which was originally proposed in (Romero *et al.*, 2018) and developed in (Mitarai *et al.*, 2018; Schuld *et al.*, 2019). A key advantage is that the gradient is exact even if the difference parameter  $\epsilon$  is chosen to be a large number (commonly  $\epsilon = \pi/2$ ), avoiding the issues of the finite difference method. We assume that the unitary to be optimized can be written as  $\mathcal{U}(\boldsymbol{\theta}) = V(\boldsymbol{\theta}_{-i})G(\theta_i)W(\boldsymbol{\theta}_{-i})$ , where  $G = e^{-i\theta_i g}$  is the unitary affected by the parameter  $\theta_i$ ,  $g$  is the generator of  $G$  and  $V, W$  are unitaries independent of  $\theta_i$ . If  $g$  has a spectrum of two eigenvalues  $\pm\lambda$  only, the gradient can be calculated by measuring the observable at two shifted parameter values as follows:

$$\partial_i \langle f(\boldsymbol{\theta}) \rangle = \lambda (\langle f(\boldsymbol{\theta}_+) \rangle - \langle f(\boldsymbol{\theta}_-) \rangle), \quad (35)$$

where  $\boldsymbol{\theta}_\pm = \boldsymbol{\theta} \pm (\pi/4\lambda)\mathbf{e}_i$ .

This rule can be generalised to the case where the generator  $g$  does not satisfy the eigenspectrum condition by decomposing the unitary as  $G = G_1 G_2 \dots G_n = e^{-i\theta_i(g_1 + g_2 + \dots + g_n)}$ , where the generator  $g_m$  of  $G_m = e^{-i\theta_i g_m}$  satisfies it. We can then use the parameter-shift rule on each  $G_m = e^{-i\theta_i g_m}$  and calculate the analytical gradient using the product rule. This has been further developed for calculating analytical gradients for fermionic generators of Unitary Coupled-Cluster operators (Kottmann *et al.*, 2020b) and higher order derivatives (Mari *et al.*, 2020).

One can also use an ancilla qubit and controlled unitaries to evaluate the gradient of multi-qubit unitaries where the PARAMETER-SHIFT rule does not apply. This was originally proposed in the context of unitary coupled-cluster (Romero *et al.*, 2018) and later generalized for arbitrary gradients (Schuld *et al.*, 2019; Yuan *et al.*, 2019). A further alternative is the stochastic PARAMETER-SHIFT rule (Banchi and Crooks, 2020), which relies on stochastically sampling scaled evolutions of the generator.

*L-BFGS.* L-BFGS is a quasi-Newton method that efficiently approximates the inverse Hessian using a limited history of positions and gradients (Liu and Nocedal, 1989). It is a memory-efficient variant of the BFGS method, which stores dense approximations of the inverse

Hessian (Fletcher, 2000). While effective in simulations, recent studies observed BFGS methods do not perform well in experimental demonstrations of variational algorithms due to the level of noise in the cost function and gradient estimates (Lavrijsen *et al.*, 2020).

Two heuristics were proposed (Zhou *et al.*, 2020a) to find quasioptimal parameters for QAOA using BFGS: *i*) INTERP, where the optimized parameters at QAOA level  $p$  are linearly interpolated and used as initial parameters for the level  $p+1$  optimization; and *ii*) FOURIER, where instead of optimizing the  $2p$  QAOA parameters  $\gamma$  and  $\beta$  in Eq. (23), one can instead optimize  $2q$  new parameters  $\mathbf{u} \equiv (u_1, u_2, \dots, u_q)$  and  $\mathbf{v} \equiv (v_1, v_2, \dots, v_q)$  defined via the discrete sine and cosine transformations

$$\begin{aligned} \gamma_i &= \sum_{j=1}^q u_j \sin \left[ \left( i - \frac{1}{2} \right) \left( j - \frac{1}{2} \right) \frac{\pi}{p} \right] \\ \beta_i &= \sum_{j=1}^q v_j \cos \left[ \left( i - \frac{1}{2} \right) \left( j - \frac{1}{2} \right) \frac{\pi}{p} \right] \end{aligned} \quad (36)$$

Similarly to INTERP, the optimal parameters found at level  $p$  are used to initialize the parameters for level  $p+1$ . Note that these heuristic strategies can be easily extended to gradient-free optimization methods such as Nelder-Mead.

*Quantum natural gradient.* The update rule of standard gradient descent Eq. (32) has the implicit assumption that the underlying parameter space is a flat Euclidean space. However, in general this is not the case, which can severely hamper the efficiency of gradient descent methods. In classical machine learning, the natural gradient was proposed that adapts the update rule to the non-Euclidean metric of the parameter space (Amari, 1998). As an extension to the realm of parameterized quantum circuits, the quantum natural gradient (QNG) has been proposed (Stokes *et al.*, 2020). The update rule for this method is

$$\theta_i^{(t+1)} = \theta_i^{(t)} - \eta \mathcal{F}^{-1}(\boldsymbol{\theta}) \partial_i f(\boldsymbol{\theta}), \quad (37)$$

where  $\mathcal{F}(\boldsymbol{\theta})$  is the Fubini-Study metric tensor or quantum Fisher information matrix given by

$$\mathcal{F}_{ij} = \Re(\langle \partial_i \psi(\boldsymbol{\theta}) | \partial_j \psi(\boldsymbol{\theta}) \rangle - \langle \partial_i \psi(\boldsymbol{\theta}) | \psi(\boldsymbol{\theta}) \rangle \langle \psi(\boldsymbol{\theta}) | \partial_j \psi(\boldsymbol{\theta}) \rangle). \quad (38)$$

The QNG can show superior performance compared to other gradient methods (Stokes *et al.*, 2020; Yamamoto, 2019) and has been shown to be able to avoid becoming stuck in local minima (Wierichs *et al.*, 2020). It can be generalized to noisy quantum circuits (Koczor and Benjamin, 2019). While the full Fubini-Study metric tensor is difficult to estimate on quantum hardware, diagonal and block-diagonal approximations can be efficiently evaluated (Jones, 2020; Stokes *et al.*, 2020).

*Quantum imaginary time evolution.* Instead of using the standard gradient descent for optimization, a variational imaginary time evolution method was proposed in (McArdle *et al.*, 2019a) to govern the evolution of parameters. They focused on many-body systems described by a  $k$ -local Hamiltonian and considered a PQC that encodes the state  $|\phi(\tau)\rangle$  as a parameterized trial state  $|\psi(\boldsymbol{\theta}(\tau))\rangle$ . The evolution of  $\boldsymbol{\theta}(\tau)$  with respect to all the parameters can then be obtained by solving the following differential equation:

$$\sum_j A_{ij} \partial_t \theta_j = C_i, \quad (39)$$

where

$$A_{ij} = \Re(\partial_i \langle \psi(\boldsymbol{\theta}(\tau)) | \partial_j | \psi(\boldsymbol{\theta}(\tau)) \rangle), \quad (40)$$

$$C_i = \Re\left(-\sum_\alpha \lambda_\alpha \partial_i \langle \psi(\boldsymbol{\theta}(\tau)) | h_\alpha | \psi(\boldsymbol{\theta}(\tau)) \rangle\right), \quad (41)$$

and  $h_\alpha$  and  $c_\alpha$  are the Hamiltonian terms and coefficients. It was later shown in (Stokes *et al.*, 2020) that the matrix  $A_{ij}$  is related to the Fubini-Study metric tensor from Eq. (38), and the imaginary time evolution is analogous to the gradient descent via the QNG when considering infinitesimal small step sizes.

*Hessian-aided gradient descent.* A recent work (Huembeli and Dauphin, 2020) proposed computing the Hessian and its eigenvalues to help analyze the cost function landscapes of quantum machine learning algorithms. Tracking the numbers of positive, negative, and zero eigenvalues provides insight into where the current point is in the landscape and whether the optimizer is heading towards a stationary point. Using notations from (Huembeli and Dauphin, 2020), the Hessian can be computed by doubly applying the parameter shift rule (Mitarai and Fujii, 2019)

$$\begin{aligned} \partial_i \partial_j f(\boldsymbol{\theta}) = & \frac{1}{2} \left( \langle f(\boldsymbol{\theta}_{-i,j}, \theta_i + \alpha, \theta_j + \alpha) \rangle + \langle f(\boldsymbol{\theta}_{-i,j}, \theta_i - \alpha, \theta_j - \alpha) \rangle \right. \\ & \left. - \langle f(\boldsymbol{\theta}_{-i,j}, \theta_i - \alpha, \theta_j + \alpha) \rangle - \langle f(\boldsymbol{\theta}_{-i,j}, \theta_i + \alpha, \theta_j - \alpha) \rangle \right), \end{aligned} \quad (42)$$

where the shift parameter  $\alpha = \frac{\pi}{4\lambda}$  for gates generated by operators with eigenvalues  $\pm\lambda$ . Other parameters, i.e. parameters not at the  $i$ -th and  $j$ -th indices, denoted  $\boldsymbol{\theta}_{-i,j}$ , are fixed. To improve optimization, they propose setting the learning rate to the inverse of the largest eigenvalue of the Hessian. When numerically demonstrated alongside QNG and the standard gradient descent methods, the Hessian-based method and QNG both showed improvement over standard gradient descent in the ability

to escape flat regions of the landscape with the Hessian-based method requiring fewer training epochs than QNG. While a deeper analysis is necessary to more closely compare the performance, both QNG and Hessian-based methods accelerate optimization by leveraging local curvature information.

*Quantum Analytic Descent.* An alternative method of using a classical model of the local energy landscape was proposed in (Koczor and Benjamin, 2020). They proposed a hybrid approach where a quantum device is used to construct an approximate ansatz landscape, and the descent towards the minima of these approximate surfaces can be carried out efficiently on a classical computer. The method considers the ansatz circuit as a product of  $m$  gate operations as  $\Phi(\boldsymbol{\theta}) = \Phi_m(\theta_m) \cdots \Phi_2(\theta_2) \Phi_1(\theta_1)$ , which, without loss of generality, can be approximated around a reference point  $\theta_0$  as:

$$\begin{aligned} \Phi(\boldsymbol{\theta}) = & A(\boldsymbol{\theta}) \Phi^{(A)} + \sum_{k=1}^v [B_k(\boldsymbol{\theta}) \Phi_k^{(B)} + C_k(\boldsymbol{\theta}) \Phi_k^{(C)}] \\ & + \sum_{l>k}^v [D_{kl}(\boldsymbol{\theta}) \Phi_{kl}^{(D)}] + O(\sin^3 \delta), \end{aligned} \quad (43)$$

where  $A, B_k, C_k, D_{kl} : \mathbb{R}^v \rightarrow \mathbb{R}$  are products of simple univariate trigonometric functions,  $\Phi^{(A)}, \Phi_k^{(B)}, \Phi_k^{(C)}, \Phi_{kl}^{(D)}$  are discrete mappings of the gates and  $\delta$  is the absolute largest entry of the parameter vector. Using this approximate ansatz landscape, the full energy surface, gradient vector and metric tensor can be expressed in term of the ansatz parameters. The analytic descent has been shown to achieve faster convergence as compared to the QNG (Koczor and Benjamin, 2020).

*Stochastic gradient descent.* A major drawback of gradient-based methods is the high number of measurements. The stochastic gradient descent (SGD) algorithm (Harrow and Napp, 2019; Sweke *et al.*, 2020) addresses this issue by replacing the normal parameter update rule with a modified version

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \alpha \mathbf{g}(\boldsymbol{\theta}^{(t)}), \quad (44)$$

where  $\alpha$  is the learning rate and  $\mathbf{g}$  is an unbiased estimator of the gradient of the cost function. There are many choices for this estimator, for instance a measurement of the gradient with a finite number of shots (Harrow and Napp, 2019). It was also shown that it is not necessary to include all Pauli terms in the evaluation of the cost function; sampling from a subset still results in well-behaved gradient estimator. It is however possible to go even further by combining this technique with sampling of the parameter-shift rule terms (Sweke *et al.*, 2020).

In the doubly stochastic gradient, finite measurements are performed for only a subset of the expectation values

of a subset of the Hamiltonian terms. This sampling can be performed in the extreme situation where only a single Pauli-term is evaluated at a single point in the quadrature. This is a very powerful method that reduces the number of measurements drastically (Anand *et al.*, 2020b).

This method can be extended beyond circuits that allow the parameter-shift rule by expressing the gradient as an integral (Banchi and Crooks, 2020). The integral can be seen as an infinite sum of terms that can be sampled. To accelerate the convergence of SGD for VQA, two optimization strategies were proposed (Lyu *et al.*, 2020): *i*) Qubit-recursive, where the optimization is first done for a smaller quantum system and the optimization is then used as initial parameter guesses for a larger quantum system; and *ii*) Layer-recursive, which is similar to a greedy approach, where the parameters are sequentially updated layer-by-layer in the quantum circuit.

## 2. Gradient-free approaches

In this section, we discuss optimization approaches for variational quantum algorithms that do not rely on gradients measured on the quantum computer.

*Evolutionary algorithms.* Evolutionary strategies (Rechenberg, 1978; Schwefel, 1977) are black-box optimization tools for high dimensional problems. They use a search distribution, from which they sample data to estimate the gradient of the expected fitness, to update the parameters in the direction of steepest ascent. More recently, natural evolutionary strategies (NES) (Wierstra *et al.*, 2014) have demonstrated considerable progress in solving these high dimensional optimization problems, and they use natural gradient estimates for parameter updates instead of the standard gradients. They have been adapted for optimization of variational algorithms (Anand *et al.*, 2020a; Zhao *et al.*, 2020b) and have been shown to have similar performance as the state-of-the-art gradient based method.

The search gradients used in NES can be estimated as

$$\nabla J(\boldsymbol{\theta}) \approx \frac{1}{k} \sum_{n=1}^k f(\mathbf{z}_n) \nabla \log \pi(\mathbf{z}_n | \boldsymbol{\theta}), \quad (45)$$

where  $J(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}[f(\mathbf{z})]$  is the expected fitness,  $\pi(\mathbf{z} | \boldsymbol{\theta})$  is the density of the search distribution with parameter  $\boldsymbol{\theta}$ ,  $f(\mathbf{z})$  is the fitness for the corresponding sample  $\mathbf{z}$  drawn from the search distribution and  $k$  is the different number of samples drawn from the distribution. The (classical) Fisher matrix  $\mathcal{F}_C$  for the natural gradient can be estimated as

$$\mathcal{F}_C \approx \frac{1}{k} \sum_{n=1}^k \nabla \log \pi(\mathbf{z}_n | \boldsymbol{\theta}) \nabla \log \pi(\mathbf{z}_n | \boldsymbol{\theta})^T \quad (46)$$

and the parameter update can then be carried out as  $\boldsymbol{\theta} = \boldsymbol{\theta} + \eta \cdot \mathcal{F}_C^{-1} \nabla J(\boldsymbol{\theta})$ . It has been shown in (Anand *et al.*, 2020a) that NES along with techniques like Fitness shaping, local natural coordinates, adaptive sampling and batch optimization, can be used for optimization of deep quantum circuits.

*Reinforcement learning.* Several authors have used reinforcement learning (RL) to optimize the QAOA parameters (Khairy *et al.*, 2019; Wauters *et al.*, 2020b; Yao *et al.*, 2020a,b). The RL framework consists of a decision-making agent with policy  $\pi_{\boldsymbol{\theta}}(a|s)$  parameterized by  $\boldsymbol{\theta}$ , which is a mapping from a state  $s \in$  state space  $S$  to an action  $a \in$  action space  $A$ . In response to the action, the environment provides the agent with a reward  $r$  from the set of rewards  $R$ . The goal of RL is to find a policy which maximizes the expected total discounted reward. For more details, refer to Sec. V.B.3. In the context of QAOA, for example, the state space  $S$  can be the set of QAOA parameters  $(\boldsymbol{\gamma}, \boldsymbol{\beta})$  used, the action  $a$  can be the value of  $\gamma$  and  $\beta$  for the next iteration, and the reward can be the finite difference in the QAOA objective function between two consecutive iterations. The policy can be parameterized by a deep neural network, for which  $\boldsymbol{\theta}$  represents the weights of the neural network. The policy parameters  $\boldsymbol{\theta}$  can be optimized using a variety of algorithms such as Monte-Carlo methods, Q-Learning and policy gradient methods.

*Sequential minimal optimization.* In machine learning, the sequential minimal optimization (SMO) method (Platt, 1998) has proven successful in optimizing the high-dimensional parameter landscape of support vector machines. The method breaks the optimization into smaller components for which the solution can be found analytically. This method has been applied to variational circuit optimization (Nakanishi *et al.*, 2020), circuit optimization with classical acceleration (Parrish *et al.*, 2019b) and circuit optimization and learning with Rotosolve and Rotoselect (Ostaszewski *et al.*, 2019). Although these algorithms heavily rely on the parameter-shift rule, they can be considered gradient-free methods. They exploit the sinusoidal nature of the expectation value of a specific operator  $\hat{O}$  when all but one parameters in the variational circuit are fixed:

$$\langle \hat{O} \rangle(\boldsymbol{\theta}) = A * \sin(\boldsymbol{\theta} + B) + C, \quad (47)$$

where  $A$ ,  $B$  and  $C$  are parameters that can be found analytically. This means that only three well-chosen circuit evaluations are needed to exactly determine these coefficients and the optimal value  $\theta^*$  of the parameter  $\theta$  for



this operator is given by

$$\theta^* = -\arctan2 \left[ 2\langle \hat{O} \rangle(\varphi) - \langle \hat{O} \rangle\left(\varphi + \frac{\pi}{2}\right) - \langle \hat{O} \rangle\left(\varphi - \frac{\pi}{2}\right), \right. \\ \left. \langle \hat{O} \rangle\left(\varphi + \frac{\pi}{2}\right) - \langle \hat{O} \rangle\left(\varphi - \frac{\pi}{2}\right) \right] + 2\pi k - \varphi - \frac{\pi}{2}, \quad (48)$$

where  $\arctan2$  is the 2-argument arctangent and for any integer  $k$  and angle  $\varphi$ . The most straightforward choice is to set  $\varphi = 0$  and choose  $k$  such that  $\theta^* \in (-\pi, \pi]$ . The algorithm proceeds by looping over all the variational parameters until convergence.

The method can be generalized to optimize more than one parameter at a time (Nakanishi *et al.*, 2020; Parrish *et al.*, 2019b) but no general analytical expression can be found here. One has to resort to numerical methods to find the solutions for the free parameters. SMO offers a versatile starting point that can be combined with more advanced search acceleration algorithms (Parrish *et al.*, 2019b) like Anderson acceleration (Anderson, 1965) or direct inversion of the iterative subspace (DIIS) (Pulay, 1980). The same tools have also been used to optimize categorical variables like rotation axes in the Rotoselect algorithm (Ostaszewski *et al.*, 2019). While cost efficient, sequential parameter optimization only takes into account local information (albeit exactly), which often causes the optimization to get stuck in local minima (Koczor and Benjamin, 2020). One has to balance the speed of a local method like SMO with the global approximate information of methods like the quantum analytic descent for specific problems.

*Surrogate model-based optimization.* When function evaluations are costly, it pays off to not only use the current function value to inform a next parameter value, but to use all previous evaluations to extract information about the search space. The function values in memory are used to build a surrogate model, an auxiliary function that represents the full expensive cost function based on the current information. The surrogate model of the cost function can be evaluated cheaply and many of these evaluations can be used to inform the next parameter value at which to evaluate the cost function. The new evaluation of the true objective function is added to the set of function values and the surrogate model is iteratively refined until convergence. All optimization happens on the surrogate cost landscape, so no explicit derivatives of the cost function are needed. Through the use of a fitted cost function, these methods are also expected to be more resilient to noise.

Several classical surrogate models have been included in the scikit-quant package (Lavrijsen and the scikit-quant contributors, 2020; Lavrijsen *et al.*, 2020). In the Bound optimization by quadratic approximation

(BOBYQA) algorithm (Powell, 2009), a local quadratic model is formulated from the previous function values. It is then minimized in the trust region to obtain a new parameter value. When the evaluation at this new parameter value does not result in a lower function value, the trust region is altered and the quadratic model is optimized in this new parameter space. It was shown that this method works well when the PQC is initialized close to the optimal parameters but has more problems with shallow optimization landscapes and gets stuck in local minima (Lavrijsen *et al.*, 2020). The stable noisy optimization by branch and fit (SnobFit) (Huyer and Neumaier, 2008) algorithm uses a branching algorithm to explore new areas in parameter space. In these areas it proposes several evaluation points and from the function values it fits a quadratic model. The algorithm combines this local search with an explorative generation of points in new areas of parameter space. SnobFit performs well when tight bounds on the parameters are available (Lavrijsen *et al.*, 2020). This can be achieved by combining the method with other optimizers that limit the size of the parameter space.

Sung *et al.* (2020) have used a trust region with a least-squares fit to a quadratic function. They find that the minimum of the quadratic function often lies outside of the trust region, which causes the algorithms to constantly readjust the trust region. Instead of using standard trust region optimization, they use either standard gradient descent or policy gradient descent optimization on the quadratic function to define a search direction in parameter space. Both the model gradient descent (MGD) as the model policy gradient (MPG) have a hyperparameter that gradually shrinks the set of points used for the fit around the current minimum as the optimization progresses. The authors show that their algorithms are well suited for realistic conditions on near-term hardware. In particular, they study the ability of the algorithm to take into account a cloud access situation where circuits need to be uploaded in batches with a certain latency. The MPG outperforms MGD in the case of gate errors due to the ability to handle a large level of uncertainty while learning its policy (Sung *et al.*, 2020).

### 3. Resource-aware optimizers

Optimization methods and strategies adopted for early demonstrations of variational quantum algorithms were largely general-purpose and black-box, with minimal emphasis on reducing *quantum resources* used in the optimization, which are more costly and prone to errors than their classical counterparts. Optimizers developed in more recent years were tailored to additionally minimize quantities associated with the quantum cost of the optimization, e.g. circuit depth, two-qubit gate count, or the number of measurements.

**ADAPT-VQE.** Early variational algorithms employed a fixed ansatz design, with its parameters tuned using some classical optimizer. The “Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver” (ADAPT-VQE) was introduced as a more scalable and efficient way to simultaneously design and optimize a parameterized ansatz (Grimsley *et al.*, 2019b). At each iteration, ADAPT-VQE constructs an ansatz by adding an operator corresponding to the largest gradient from a carefully designed operator pool. That is, given an operator  $\hat{\tau}_i$  from the operator pool, the gradient of the energy with respect to the corresponding parameter  $\theta_i$  is defined as

$$\partial_i E = \langle \psi | [H, \hat{\tau}_i] | \psi \rangle, \quad (49)$$

where  $|\psi\rangle$  is the ansatz at the current iteration to be updated. After computing the gradient components and choosing the operator corresponding to the largest gradient, the gate operation implementing  $\hat{\tau}_i$  is added to the ansatz with its parameter value initialized at 0. The ansatz is then optimized before growing another operator. The ADAPT-VQE algorithm terminates when the norm of the gradient vector falls below a defined threshold. In the case of fermionic ADAPT-VQE, the operator pool consists of fermionic operators that are transformed into quantum gate operations through e.g. the Jordan-Wigner mapping. A more “hardware-efficient” variant of the ADAPT-VQE algorithm is the qubit ADAPT-VQE, in which the pool consists of gate operators acting directly on qubits (Tang *et al.*, 2019). Both versions of ADAPT-VQEs were able to generate optimized circuits with reduced depths and CNOT counts compared to previous ansatz construction and optimization methods.

**MI-ADAPT-VQE.** The mutual information-assisted ADAPT-VQE (MI-ADAPT-VQE), introduced by (Zhang *et al.*, 2021), leverages the density matrix renormalization group (DMRG) (Hallberg, 2006; White, 1992) method to accelerate the circuit constructions for the ADAPT-VQE routine. At the start of the algorithm, the pair-wise quantum mutual information is approximated using DMRG, which is then applied to construct a reduced pool of entangling gates. In each iteration of the method, new circuits are generated in which quantum gates are mainly distributed among pairs of qubits corresponding to large mutual information. This avoids allocating quantum resources on pairs of qubits that are less important to entangle. Numerical experiments suggest that using MI-ADAPT-VQE, which uses the mutual information to guide circuit constructions rather than gradients, the number of new circuits needed in each step of the adaptive construction can be significantly reduced, saving in both time and quantum resources. The paper showed that the number

of trial circuits in certain cases can be reduced to about 5% for  $H_2$  and 10% for  $H_2O$  as compared to ADAPT-VQE using an operation pool based on the qubit coupled-cluster method (Ryabinkin *et al.*, 2020, 2018b).

**ROSALIN.** While variational algorithms leverage low-depth circuits to execute on near-term quantum processors, a significant challenge in implementing these algorithms is the prohibitive number of measurements, or shots, required to estimate each expectation value that is used to compute the cost function value. To address the challenge, (Arrasmith *et al.*, 2020b) developed a “shot-frugal” optimizer called ROSALIN (Random Operator Sampling for Adaptive Learning with Individual Number of shots) that effectively distributes fractions of a predefined number of shots to estimate each term of the Hamiltonian as well as each partial derivative. Given the expectation value of the Hamiltonian decomposed into the  $h_i$  terms as in Eq. (4), the authors note several strategies for allocating shots for estimating each term  $\langle h_i \rangle$ . While a naive strategy would allocate equal numbers of shots per term, the authors observed lower variance in the energies using weighted approaches in which the number of shots allocated to the  $i$ -th term  $b_i$  is proportional to the corresponding Hamiltonian coefficient  $c_i$ . For instance, in the *weighted deterministic sampling* method,  $b_i = b_{\text{tot}} \frac{|c_i|}{\sum_i |c_i|}$  where  $b_{\text{tot}}$  is the total number of shots. In the *weighted random sampling* method,  $b_i$  is drawn from a multinomial distribution with the probability of measuring the  $i$ -th term weighted by  $|c_i|$ , i.e.  $p_i = \frac{|c_i|}{\sum_i |c_i|}$ .

In addition, ROSALIN employs iCANS, an optimizer that allocates shots for partial derivatives, as a subroutine. The iCANS (individual Coupled Adaptive Number of Shots) optimizer was originally developed by (Kübler *et al.*, 2020). As a brief overview, the iCANS algorithm allocates measurements for each partial derivative such that the *expected gain per shot*, which depends on quantities including the learning rate, Lipschitz constant of the cost function, and estimates of gradient components and their variances, is maximized. Through VQE optimizations, ROSALIN was shown to outperform other optimizers such as iCANS and Adam especially in the presence of shot noise.

**SPSA.** In experimental realizations of variational algorithms, the optimizer is often hindered by statistical noise. (Kandala *et al.*, 2017) circumvented this issue by applying the simultaneous perturbation stochastic approximation (SPSA) algorithm (Spall, 1992), in which the algorithm hyperparameters are determined by experimental data on the level of statistical noise. Compared to the finite-difference gradient approximation, which requires  $O(p)$  function evaluations for  $p$  parameters, SPSA

requires only two evaluations. That is, for a small positive  $b_k$ , the gradient at  $k$ -th iteration is approximated as

$$\mathbf{g}_k(\boldsymbol{\theta}_k) = \frac{\langle \psi(\boldsymbol{\theta}_k^+) | H | \psi(\boldsymbol{\theta}_k^+) \rangle - \langle \psi(\boldsymbol{\theta}_k^-) | H | \psi(\boldsymbol{\theta}_k^-) \rangle}{2b_k} \boldsymbol{\Delta}_k^{-1}, \quad (50)$$

where  $\boldsymbol{\Delta}_k$  is a random perturbation vector and  $\boldsymbol{\theta}_k^\pm = \boldsymbol{\theta}_k \pm b_k \boldsymbol{\Delta}_k$ . After computing the gradient estimate, the following parameter settings are updated with learning rate  $a_k$ . In general,  $b_k$  and  $a_k$  decrease over iterations, e.g.  $b_k = \frac{b}{k^\gamma}$  and  $a_k = \frac{a}{k^\alpha}$  for some fixed  $\gamma$  and  $\alpha$ . In (Kandala *et al.*, 2017), values of  $b$  and  $a$  are carefully estimated to be robust against statistical noise based on samples of energy differences  $|\langle \psi(\boldsymbol{\theta}_k^+) | H | \psi(\boldsymbol{\theta}_k^+) \rangle - \langle \psi(\boldsymbol{\theta}_k^-) | H | \psi(\boldsymbol{\theta}_k^-) \rangle|$ . The convergence of SPSA with various types of PQCs has been studied (Woitzik *et al.*, 2020).

**MoG-VQE.** To reduce two-qubit gate counts for near-term experiments, the multiobjective genetic variational quantum eigensolver (MoG-VQE) optimizes for both the energy and the number of CNOTs in the quantum circuit (Chivilikhin *et al.*, 2020). The MoG-VQE algorithm combines two evolutionary strategies: (1) NSGA-II (Deb *et al.*, 2000), a multiobjective genetic algorithm, to propose a circuit structure to minimize both the energy and CNOT count, and (2) CMA-ES (Hansen *et al.*, 2003) to tune parameters and evaluate optimized energies for the qubit topologies suggested by the NSGA-II algorithm. MoG-VQE initializes a diverse population by sampling a checkerboard pattern of two-qubit circuit blocks. To vary the populations over different generations, the three possible mutation operators are: *i*) inserting a two-qubit circuit block in a random position; *ii*) removing a two-qubit circuit block in a random position; and *iii*) adding or removing 10 circuit blocks to help escape from local minima. The authors note that *iii*) is selected with a lower probability than mutation operators *i*) and *ii*). Parents are selected using the tournament selection method. For each circuit topology, its corresponding energy is evaluated using the CMA-ES optimizer. These steps repeat until some termination criteria are satisfied. Using MoG-VQE, the authors reported significant reductions in the CNOT counts compared to those of other hardware-efficient circuits when estimating ground state energies of several molecules. For example, for a 12-qubit LiH Hamiltonian, MoG-VQE generated a circuit corresponding to estimating the ground state energy to within chemical precision using only 12 (non-nearest-neighbor) CNOTs.

**PECT.** An alternative approach for adaptively constructing and optimizing an ansatz was introduced by the “Parameter-efficient circuit training” (PECT) scheme

(Sim *et al.*, 2020). PECT enables optimizations of pre-defined ansätze, such as unitary coupled-cluster or the low-depth circuit ansatz (LDCA) (Dallaire-Demers *et al.*, 2019), by dynamically pruning and adding back parameterized gates during an optimization. After selecting an ansatz  $\mathcal{U}$ , in PECT, a subset of gate operations from  $A$  is chosen while other parameterized gate operations are tuned to identity operations. This results in an ansatz substructure  $A'$  with reduced circuit depth and gate count. Parameters of  $A'$  are then optimized, which the authors call a “local optimization” step. After local optimization, to refine or reparameterize the ansatz substructure, parameters with small magnitudes are pruned or removed. A heuristic growth rule is used to grow back the same number of parameters that was pruned. Steps of local optimization and ansatz reparameterization are repeated until termination criteria are met. Because PECT optimizes parameter subsets at any iteration, circuits that are executed on the quantum computer have reduced depths and CNOT counts than those of the original ansatz. Using PECT, Sim *et al.* were able to optimize 12-qubit LDCA circuits, naively equipped with hundreds to low thousands of parameters, to estimate ground state energies of LiH and H<sub>2</sub>O. Previous optimizations of LDCA were limited to 8 qubits.

Some intuitive concepts of the mechanisms behind optimisation of quantum problems have been investigated (McClean *et al.*, 2020b). Despite the significant progress in development of optimization algorithms and strategies, many of them are limited to simulation studies. To continue building effective optimizers for variational algorithms, it is important for these methods to be more “quantum aware” and directly address current limitations and challenges of variational algorithms, e.g. the large number of measurements required to evaluate a cost function to high precision and/or the need to further reduce circuit depth.

### III. OTHER NISQ APPROACHES

We proceed to review some of the notable NISQ algorithms, besides VQA. These algorithms do not require tuning the parameters of a PQC in an adaptive feedback manner and often exploit analog or hybrid paradigms that constitute alternatives to the digital quantum computation.

#### A. Quantum annealing

Quantum annealing (QA) derives its inspiration from simulated annealing (SA), a classical global optimization technique, usually employed to solve combinatorial optimization problems. SA can be valuable in discovering global optima in situations involving many local

optima. The word “annealing” comes from metallurgy, which represents heating and slow cooling. In SA, one identifies the objective function with the energy of a statistical-mechanical system. The system is assigned an artificially-induced control parameter, called temperature. Like annealing, SA starts with some high temperature  $T$ , and then the value of  $T$  is brought down following some temperature variation function called “annealing schedule” such that the final temperature is  $T = 0$ . The algorithm chooses a candidate state close to the current state randomly. If it improves the solution, it is always accepted with probability 1. If it does not, then the acceptance is determined based on a temperature-dependent probability function. The idea of tolerating worse solutions can be considered as a virtue of the algorithm. In SA, the probability that a bad solution is accepted is slowly decreased as the solution space is explored. This relates to the notion of “slow cooling” in annealing.

In QA, one utilises quantum-mechanical fluctuations, like quantum tunnelling, to explore the solution space. This is analogous to the idea of using thermal fluctuations in SA to explore the solution space. In QA, artificial degrees of freedom of quantum nature via non-commutative operators are introduced, which induces quantum fluctuations. The strength of these quantum fluctuations is controlled using an annealing schedule (similar to SA, where we decrease the temperature). The physical idea behind annealing schedule in QA is to move the system from some initial Hamiltonian ground state to the ground state of the problem Hamiltonian. The concept of QA is related to the notion of quantum adiabatic evolution, which is being used for adiabatic quantum computation (Albash and Lidar, 2018).

We proceed to a formal discussion now. Adiabatic quantum computation is model of computation based on quantum mechanical processes operating under adiabatic conditions (Albash and Lidar, 2018). Before understanding adiabatic quantum computation, one needs to grasp the concept of  $k$ -local Hamiltonian.

**Definition 1  $k$ -local Hamiltonian:** A  $k$ -local Hamiltonian is a Hermitian matrix of the form  $H = \sum_{i=1}^r \hat{h}_i$  where each term is a Hermitian operator acting non-trivially on at-most  $k$  qudits, i.e.,  $\hat{h}_i = h \otimes I$  where  $h$  is a Hamiltonian acting on at-most  $k$  qudits and  $I$  is an identity operator.

Let us consider a time-dependent hamiltonian  $H(s)$ , for  $s \equiv \frac{t}{T} \in [0, 1]$  and a quantum system initialized in the ground state of  $H(0)$ . We assume that  $H(s)$  varies smoothly as a function of  $s$  and  $H(s)$  has a unique ground state for  $s \in [0, 1]$ . A quantum state initialized in  $|\psi(t=0)\rangle$  evolves according to the following Schrödinger equation (setting  $\hbar = 1$ ),

$$i \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle. \quad (51)$$

The above equation can be further, equivalently, written as

$$i \frac{d}{ds} |\psi(s)\rangle = T H(s) |\psi(s)\rangle. \quad (52)$$

Assuming  $|\psi(0)\rangle$  is a ground state of  $H(0)$ , then in the limit  $T \rightarrow \infty$ ,  $|\psi(t)\rangle$  is a ground state of  $H(1)$  obtained via evolution Eq. (51). Such an evolution will be, henceforth, referred as adiabatic evolution according to  $H$  for time  $T$ .

Now we proceed to define adiabatic quantum computation.

**Definition 2 Adiabatic quantum computation** (adapted from (Aharonov et al., 2008)): An adiabatic quantum computation is specified by two  $k$ -local Hamiltonians  $H_0$  and  $H_1$  acting on  $n$  qudits and a map  $s(t) : [0, T] \rightarrow [0, 1]$ . The input of the computation is the ground state of  $H_0$ , which is unique and is a product state. The desired output is given by a quantum state which is  $\epsilon$ -close in  $l_2$ -norm to the ground state of  $H_1$ . Furthermore,  $T$  is the smallest time such that the adiabatic evolution generated via  $H(s) = (1-s)H_0 + sH_1$  for time  $T$  yields the desired output. The running time of the algorithm is given by  $T \cdot \max_s \|H(s)\|$ , where  $\|\cdot\|$  denotes the spectral norm.

QA relaxes the strict requirement of adiabatic evolution, thus allowing diabatic transitions from fast changes of Hamiltonian parameters and noisy environment (Hauke et al., 2020). Because of diabatic transitions, QA is prone to getting trapped in excited states.

QA has been investigated for problems in diverse areas including machine learning (Benedetti et al., 2017; Li et al., 2018), protein folding (Perdomo-Ortiz et al., 2012), compressive sensing (Ayanzadeh et al., 2019), finance (Bouland et al., 2020; Cohen et al., 2020; Marzec, 2016; Orus et al., 2019; Rosenberg et al., 2016; Venturelli and Kondratyev, 2019), fermionic simulation (Babbush et al., 2014) and high energy physics (Das et al., 2019; Mott et al., 2017). QA has been one of the prominent approaches in the NISQ era in the search for quantum advantage (Bouland et al., 2020; Hauke et al., 2020).

A major experimental implementation of QA is the D-Wave machine. It attempts to solve problems in a particular form called Quadratic Unconstrained Binary Optimization (QUBO) (Bouland et al., 2020; Lucas, 2014). Optimization problems can be cast as a polynomial unconstrained binary optimization (PUBO) expressed in the form of a  $k$ -local interaction with  $k \geq 3$  over binary variables  $x_i \in \{0, 1\}$  (Hauke et al., 2020). QUBO is a special case of PUBO with  $k = 2$ . For a vector of  $n$  binary variables  $\mathbf{x} \in \{0, 1\}^n$  and problem specified values of  $Q \in \mathbb{R}^{n \times n}$  and  $\mathbf{c} \in \mathbb{R}^n$ , QUBO is defined as

$$\arg \min \mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x}. \quad (53)$$



Using the map  $x_i \rightarrow \frac{1-\sigma_z^i}{2}$ , one can convert the problem in expression 53 to ground state finding problem of the following diagonal  $n$ -qubit Ising Hamiltonian (up to a constant),

$$H_{\text{QUBO}} = -\sum_{i,j} J_{i,j} \hat{\sigma}_z^i \hat{\sigma}_z^j - \sum_i h_i \hat{\sigma}_z^i, \quad (54)$$

where  $\hat{J}_{i,j} = -\frac{Q_{i,j}}{4}$  and  $h_i = \frac{-c_i + \sum_j Q_{i,j}}{2}$ .

Starting with the ground state of the base Hamiltonian  $H_0 = -\sum_i \hat{\sigma}_x^i$ , solving the QUBO problem on a quantum annealer corresponds to implementing the annealing schedule  $A(t)$  and  $B(t)$  for the Hamiltonian

$$H(t) = A(t)H_0 + B(t)H_{\text{QUBO}}. \quad (55)$$

Here,  $A(0) = B(T) = 1$  and  $A(T) = B(0) = 0$ , where  $T$  is computation time. Because annealing does not necessarily satisfy the constraints of adiabatic evolution, one can end up in excited states as mentioned earlier. However, one can run the annealing schedule multiple times and take the best answer i.e., the one corresponding to lowest energy. The qubits in an annealer are not necessarily all-to-all connected, necessitating additional engineering restrictions, such as the minor embedding problem (Bouland *et al.*, 2020).

The potential of QA has been studied extensively (Brady and van Dam, 2016; Denchev *et al.*, 2016; Farhi *et al.*, 2002; Hastings, 2020; Hauke *et al.*, 2020). The performance of D-Wave annealers have also been explored comprehensively (Albasha *et al.*, 2015; Cohen *et al.*, 2020; Shin *et al.*, 2014). For the details of QA, refer to (Hauke *et al.*, 2020) and the references therein. A review on Adiabatic Quantum Computation is presented in (Albasha and Lidar, 2018).

## B. Gaussian boson sampling

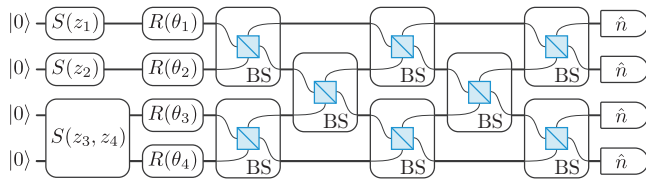


Figure 4 Gaussian boson sampling circuit for a photonic setup. The qumodes are prepared in gaussian states from the vacuum by squeezing operations  $S(z_i)$ , followed by an interferometer consisting of phaseshifters  $R(\theta) = e^{i\theta_j}$  and beam-splitters BS. At the end, photon number resolving measurements are made in each mode.

Boson sampling was first proposed as a candidate for quantum computational supremacy by (Aaronson and Arkhipov, 2011). The scenario consists of having  $n$  photons that enter an optical circuit consisting of  $m$

modes. This state is then acted upon by a series of phase-shifters and beam-splitters. A phase-shifter adds a phase  $R(\theta) = e^{i\theta_j}$  with some angle  $\theta_j$  to the amplitude in mode  $j$ , and acts as the identity in the other  $m-1$  modes. A beam-splitter acts on two modes with a rotation  $\begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$  for some angle  $\phi$  and as the identity in the other  $m-2$  modes. Finally, a measurement is made where the number of photons in each mode is found. An optical circuit with these elements is shown in Figure 4. Each of these measurement outcomes represent a sample from the symmetric wavefunction that bosonic systems have. Aaronson and Arkhipov found that the existence of an efficient classical algorithm for sampling from the distribution implies the existence of a classically efficient algorithm for the calculation of the permanent of a related matrix. Such an algorithm would imply the collapse of the polynomial hierarchy (see I.A) to the third order, which is believed to be unlikely (Arora and Barak, 2009). Consequently, such an algorithm is unlikely to exist.

Gaussian boson sampling (GBS) is a variant of boson sampling, where instead of photon states as inputs into the optical circuit, Gaussian states are used as inputs (Hamilton *et al.*, 2017). Gaussian states are those whose Wigner quasi-probability distributions  $W(q, p)$  have Gaussian shape. They have the advantage that they can be created deterministically (Hamilton *et al.*, 2017; Kruse *et al.*, 2019). They also provide additional degrees of freedom in comparison to boson sampling. Where boson sampling is equivalent to sampling from the permanent of a matrix, GBS is computationally equivalent to sampling from the Hafnian function of a matrix. Given a graph  $G$  with adjacency matrix  $E$ , the Hafnian of  $E$  is the number of perfect matchings of the graph  $G$ . A matching of a graph  $G$  is a subset of edges  $M$  such that no two edges in  $M$  have a vertex in common. A matching  $M$  is perfect if every vertex is incident to exactly one edge in  $M$ . While the Permanent gives the number of perfect matchings for a bipartite graph; the Hafnian gives perfect matching for any graph. Thus, the Hafnian can be thought of as a generalization of the Permanent. Using the adjacency matrix  $E$ , the relation between the Hafnian and the Permanent is given by

$$\text{Haf} \begin{pmatrix} 0 & E \\ E^T & 0 \end{pmatrix} = \text{Per}(E). \quad (56)$$

GBS has recently become the second platform to show quantum computational supremacy (Zhong *et al.*, 2020).

### 1. The protocol

In GBS we consider  $m$  quantum modes (*qumodes*), represented by harmonic oscillators with canonically conjugate variables  $q$  and  $p$ . Gaussian states of the qumodes

are those represented by a Wigner-function  $W(q, p)$  that has a Gaussian form. These states can be efficiently represented by the complex amplitude  $\alpha = \frac{1}{\sqrt{2\hbar}}(q + ip)$  and a covariance matrix  $\Sigma \in \mathbb{C}^{2m \times 2m}$ . A general pure Gaussian state can be generated from a vacuum with three steps: *i*) Single mode squeezing; *ii*) multi-mode linear interferometry; and *iii*) single-mode displacements. In the GBS protocol the state is then measured in the Fock-basis, performed in practice using photon number resolving detectors. The optical circuit in Figure 4 shows how the system is initialized in the vacuum state, followed by single- and multi-mode squeezing operators  $S(z_i)$  and  $S(z_i, z_j)$ , respectively, and an interferometer with phase-shifters  $R(\theta_j)$  and beamsplitters BS. At the end of the protocol, the photon number in each mode is measured.

For a Gaussian state with zero mean (of the Wigner-function), the probability of detecting  $s_i$  photons in the  $i$ -th qumode is given by: (Hamilton *et al.*, 2017; Kruse *et al.*, 2019)

$$P(s_1, s_2 \dots s_m) = \frac{1}{\det(Q)} \frac{\text{Haf}(A_s)}{\sqrt{s_1! s_2! \dots s_m!}} \quad (57)$$

where all the matrices are defined in terms of the covariance matrix  $\Sigma$ :

$$\begin{aligned} Q &= \Sigma + \mathbb{1}/2 \\ A &= X(\mathbb{1} - Q^{-1}) \\ X &= \begin{bmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix} \end{aligned}$$

The  $A_s$  matrix is a matrix created from  $A$  such that if  $s_i = 0$ , we delete the rows and columns  $i$  and  $i + m$  of the matrix, and if  $s_i \neq 0$ , we repeat the rows and columns  $s_i$  times.

This means that by manipulating the covariance matrix  $\Sigma$ , we control the matrix from which we sample the Hafnian. For a pure Gaussian state, it can be shown that the  $A$  matrix is symmetric (Bromley *et al.*, 2020).

A simpler form of the experiment where instead of counting the number of photons in each mode, we only detect if there are photons or not in each mode, can be used to sample from the so-called Torontonian function of a matrix (Quesada *et al.*, 2018). A more general experiment instead, where the mean of the Gaussian states is non-zero, can be used to sample from the loop Hafnian (Björklund *et al.*, 2019).

## 2. Applications

A number of algorithms for applications of GBS have been investigated, and are reviewed by (Bromley *et al.*, 2020). Here we only briefly summarise that work. Typically GBS algorithms are based on heuristics, and GBS devices are often used to provide a seed for starting points

of classical algorithms. GBS can also be viewed as directly giving access to a statistical distribution, such as in the case of point processes (Jahangiri *et al.*, 2020b).

Problems in chemistry have been tackled using GBS. Vibrational spectra of molecules have been computed using GBS by mapping the phononic modes of the molecule to the qumodes of the GBS device (Huh *et al.*, 2015), and by extension electron-transfer reactions have been studied (Jahangiri *et al.*, 2020a). The technique of sampling high-weight cliques has also been applied to predict molecular docking configurations (Banchi *et al.*, 2020a).

The largest number of GBS algorithms are for graph problems, since the adjacency matrix of a graph is a natural fit as the symmetric  $A$  matrix. The Hafnian function computes the number of perfect matchings of a graph, so the samples from the GBS device are with high likelihood from sub-graphs with high density. This is how GBS is used to identify dense subgraphs (Arrazola and Bromley, 2018), and to get good initial guesses for classical search algorithms to compute the max-clique of a graph (Banchi *et al.*, 2020a).

GBS can also be used to build succinct feature vectors, or “fingerprints”, of larger graphs via coarse-graining techniques. These feature vectors can then be used as inputs to statistical methods or machine learning to classify graphs. One such problem is to measure the similarity between graphs (Schuld *et al.*, 2020b), which has applications in tasks such as checking fingerprint comparison or detecting mutations of molecules.

GBS can also be used as a type of importance sampling device to speed up algorithms requiring randomness. This is how stochastic search algorithms have been sped up by sampling from a GBS device encoding the graph to be searched, instead of sampling uniformly (Arrazola *et al.*, 2018).

Recently variational methods have been used within the GBS framework (Banchi *et al.*, 2020b) and applied to stochastic optimization and unsupervised learning. The method is based on varying the squeezing and interferometer parameters in the device and updating based on the measurement outcomes.

## C. Analog quantum simulation

Simulating a quantum system is a hard problem for classical computers as the Hilbert space increases exponentially with the size of the system. As a solution to this long-standing problem, Feynman suggested the ground breaking idea to harness that physical systems given us by nature are quantum-mechanical. He proposed to use quantum systems that are well-controlled in the lab to simulate other quantum systems of interest (Feynman, 1982). This concept has spurred the field of analog quantum simulation (Georgescu *et al.*, 2014; Trabesinger, 2012).

The core idea differs from digital quantum simulation (Lloyd, 1996). Digital quantum simulators decompose the quantum dynamics to be simulated into a circuit of discrete gate operations that are implemented on a quantum processor. The quantum processor is a well controlled quantum system, that is engineered to be able to efficiently apply a set of specific quantum gates that are universal, i.e. a sequential application of those gates can realize arbitrary unitaries (see Sec. IV.F.1).

With this universal approach, a wide range of quantum problems can be simulated to a desired accuracy with a polynomially increase in quantum resources only (Lloyd, 1996). However, current quantum processors have only limited coherence time and lack the capability to correct errors that inevitably appear during the computation, severely limiting the range of dynamics that can be simulated. In contrast, the idea of analog quantum simulators is to map the problem Hamiltonian to be simulated  $\hat{H}_{\text{sys}}$  to the Hamiltonian of the quantum simulator  $\hat{H}_{\text{sim}}$ , which can be controlled to some degree,  $\hat{H}_{\text{sys}} \leftrightarrow \hat{H}_{\text{sim}}$ . One then runs the quantum simulator, and maps the results back to the problem.

The range of problems that can efficiently mapped to the simulator is limited, however as one uses the native quantum dynamics of the simulator, the accessible system size, coherence length and errors is often more favorable compared to current digital quantum simulators.

## 1. Implementations

A wide-range implementations in various controlled quantum systems has been achieved, ranging from solid state superconducting circuits (Houck *et al.*, 2012), quantum dot arrays (Hensgens *et al.*, 2017), nitrogen-vacancy centers (Yao *et al.*, 2012), atomic and molecular physics based platforms such as trapped ions (Blatt and Roos, 2012), interacting photons (Chang *et al.*, 2014; Hartmann, 2016), Rydberg atoms (Adams *et al.*, 2019), and cold atoms (Amico *et al.*, 2020; Bloch *et al.*, 2012; Gross and Bloch, 2017).

Concepts of analog quantum simulation have been used within VQA as well, such as physically inspired ansätze (see Sec. II.B.1) or protocols inspired by quantum control (Meitei *et al.*, 2020; Yang *et al.*, 2017). Experimental results for a quantum many-body problem beyond current classical computational capabilities have been reported for 2-D systems (Choi *et al.*, 2016).

## 2. Programmable quantum simulators

An analogue quantum system, such as a superconducting circuit, can be adapted to simulate arbitrary dynamics (Bastidas *et al.*, 2020). The idea is to drive the parameters of the Hamiltonian  $H(t)$  that describes the analog

quantum simulator in time  $t$ . This can be done by adjusting the physical parameters of the quantum simulator in time. The driving protocol is engineered via machine learning methods (Haug *et al.*, 2021) such that the effective dynamics of the driven system over a time  $T$  corresponds to the evolution of a problem Hamiltonian one wants to simulate. The effective dynamics that is generated can realize long-range interactions as well as complicated many-body terms, which are natively not supported by the quantum simulator and are often hard to simulate on digital quantum simulators. By periodically driving the analog quantum simulator with the aforementioned driving protocol, various problem Hamiltonians can be simulated (Oka and Kitamura, 2019). One can realize complicated many-body dynamics or chemistry problems, as well as solve combinatorial tasks such as SAT-3. Trapped ion based analog quantum simulators have been recently used for the implementation of quantum approximate optimization algorithm (Pagano *et al.*, 2020).

Highly controllable analog quantum simulators have also been proposed for engineering quantum chemistry Hamiltonians by combining different cold atom species embedded within cavity modes, which mediate long-range interactions required to simulate Coulomb repulsion. Optical fields can be used to modify the potential and interaction parameters in order to simulate large scale chemistry problems (Argüello-Luengo *et al.*, 2019) as well as quantum spin model with tunable interactions for system sizes ranging from 64 to 256 qubits (Ebadi *et al.*, 2020). For ion traps, a programmable quantum simulator can be designed by light fields, that are applied to manipulate the internal degrees of freedom as well as the interaction between different ions. This allows one to simulate various types of spin Hamiltonians with a high degree of control over the parameters (Monroe *et al.*, 2019).

## D. Digital-analog quantum simulation and computation

As mentioned previously, analog quantum simulation is limited in terms of system Hamiltonians one can simulate (Goldman and Dalibard, 2014; Kyriienko and Sørensen, 2018). On the other hand, digital quantum simulators can simulate any system Hamiltonian, but with sometimes costly quantum resources. To benefit from a combination of the two approaches, the digital-analog method to quantum computation (Dodd *et al.*, 2002; Parra-Rodriguez *et al.*, 2020) and simulation (Mezzacapo *et al.*, 2014; Yung *et al.*, 2014) have been proposed.

These schemes combine the application of digital single-qubit gates with the underlying analog Hamiltonian of the quantum processor. This approach allows for universal simulation of quantum dynamics while reducing the number of gates needed due to access to the analog

quantum Hamiltonian as a powerful resource.

Digital-analog quantum simulation has been proposed to simulate the Rabi model (Mezzacapo *et al.*, 2014), Dicke model (Lamata, 2017; Mezzacapo *et al.*, 2014), and fermion-fermion interactions (García-Álvarez *et al.*, 2015). Digital-analog quantum simulation has been reviewed in (Lamata *et al.*, 2018b), whereas digital-analog quantum computing is more recent. So far the computing framework has been used to simulate Ising models (Parra-Rodríguez *et al.*, 2020), where the analog blocks can be used to enhance the effective connectivity of the qubits to simulate graphs that have a different connectivity from the native connectivity of the quantum device (Galicía *et al.*, 2020). The analog blocks have also been applied to reduce the operation count required to perform the quantum Fourier transform (Martin *et al.*, 2020).

The digital-analog approach has also been combined with variational quantum algorithms (see Sec. II) resulting in a digital-analog QAOA algorithm, where the two-qubit gates have been replaced by analog blocks (Headley *et al.*, 2020). This also has two versions: *i*) where a layer of entangling gates is replaced by an analog block; and *ii*) where a continuous analog block is applied continuously with single-qubit operations overlaid.

### E. Iterative quantum assisted eigensolver

Almost all of the VQAs update a PQC's parameters in a feedback loop. However, there exist alternative algorithms that can circumvent this approach with the ansatz given by (Bharti and Haug, 2020a; Huang *et al.*, 2019; McClean *et al.*, 2017)

$$|\phi(\alpha(t), \theta)\rangle = \sum_{i=0}^{m-1} \alpha_i(t) |\psi_i(\theta_i)\rangle, \quad (58)$$

where  $\alpha_i \in \mathbb{C}$  and  $\theta_i \in \mathbb{R}^{k_i}$  for non-negative integers  $k_i$ . This ansatz is a linear combination of quantum states, where the  $\alpha_i$  parameters are stored on a classical device. In the special case  $m = 1$  it corresponds to the usual PQC, whereas for  $m > 1$  this ansatz subsumes it. This ansatz has been used for finding the ground state of Hamiltonians (Bharti, 2020; Bharti and Haug, 2020a), excited state (Huggins *et al.*, 2020; Parrish *et al.*, 2019a; Parrish and McMahon, 2019; Stair *et al.*, 2020) the simulation of quantum dynamics (Bharti and Haug, 2020b; Haug and Bharti, 2020), error mitigation (McClean *et al.*, 2017), nonlinear dynamics (Bharti and Haug, 2020b; Haug and Bharti, 2020) and linear systems (Huang *et al.*, 2019). If one keeps the parameters of the PQC  $\theta_i$  fixed and only varies the  $\alpha_i$ , the algorithm can be considered a borderline non-VQA algorithm. Update of  $\theta_i$  parameters has been shown to cause trainability issues in variational algorithms (see Sec. IV.B) and thus by fixing  $\theta_i$  one can by construction circumvent these issues. We present

here the iterative quantum assisted eigensolver algorithm (IQAE) as an illustration, and in the applications subsection the quantum assisted simulator for closed systems (see Sec. V.A.5), open systems (see Sec. V.A.7) and Gibbs state preparation (see Sec. V.A.9).

The IQAE algorithm provides an approximation to the ground state of a Hamiltonian  $H$ . Without loss of generality, the  $N$ -qubit Hamiltonian  $H$  is assumed to be a linear combination of unitaries

$$H = \sum_{i=1}^m \beta_i U_i. \quad (59)$$

Here,  $\beta_i \in \mathbb{C}$  and  $U_i \in SU(2^N)$  for  $i \in \{1, 2, \dots, m\}$ . The unitaries  $U_i$  act on at most  $\mathcal{O}(\text{poly}(\log N))$  qubits. This condition can be relaxed if the unitaries Pauli strings (see Sec. II.A.1). The ansatz state is taken as linear combination of “cumulative  $K$ -moment states”  $\mathbb{CS}_K$ , which is generated using some efficiently preparable quantum states and the unitaries defining the Hamiltonian in Eq. (59). For pedagogical reasons, we present the definition of  $K$ -moment states and cumulative  $K$ -moment states.

**Definition 3** (adapted from (Bharti and Haug, 2020a))  
For a given positive integer  $K$ , a set of unitaries  $\mathbb{U} \equiv \{U_j\}_{j=1}^m$  and a quantum state  $|\psi\rangle$ ,  $K$ -moment states is the set of quantum states of the form  $\{U_{j_K} \dots U_{j_2} U_{j_1} |\psi\rangle\}_{j_l \in \mathbb{U}}$  for  $U_{j_l} \in \mathbb{U}$ . Let us denote the aforementioned set by  $\mathbb{S}_K$ . We define the singleton set  $\{|\psi\rangle\}$  as the 0-moment state (denoted by  $\mathbb{S}_0$ ). Finally, we define the cumulative  $K$ -moment states  $\mathbb{CS}_K$  as  $\mathbb{CS}_K \equiv \cup_{i=0}^K \mathbb{S}_i$ .

As instructive example, note that the set of 1-moment states is  $\{U_j |\psi\rangle\}_{j=1}^m$ , where the unitaries  $\{U_j\}_{j=1}^m$  make up the Hamiltonian  $H$ . The set of cumulative 1-moment states is  $\mathbb{CS}_1 = \{|\psi\rangle\} \cup \{U_j |\psi\rangle\}_{j=1}^m$ , and the set of cumulative  $K$ -moment states is  $\mathbb{CS}_K = \{|\psi\rangle\} \cup \{U_{j_1} |\psi\rangle\}_{j_1=1}^m \cup \dots \cup \{U_{j_K} \dots U_{j_1} |\psi\rangle\}_{j_l=1, \dots, j_K=1}^m$ .

Now, the ansatz is given by  $|\xi(\alpha)\rangle^{(K)} = \sum_{|\chi_j\rangle \in \mathbb{CS}_K} \alpha_j |\chi_j\rangle$ . The ground state problem reduces to the following optimization program

$$\begin{aligned} & \min_{\alpha} \alpha^\dagger \mathcal{D}^{(K)} \alpha \\ & \text{subject to } \alpha^\dagger \mathcal{E}^{(K)} \alpha = 1. \end{aligned} \quad (60)$$

Here, the overlap matrices  $\mathcal{D}^{(K)}$  and  $\mathcal{E}^{(K)}$  are given by  $\mathcal{D}_{nm}^{(K)} = \sum_i \beta_i \langle \chi_n | U_i | \chi_m \rangle$  and  $\mathcal{E}_{nm}^{(K)} = \langle \chi_n | \chi_m \rangle$ . These overlap matrices can be computed on a quantum computer without the requirement of any complicated measurement involving multi-qubit controlled unitaries. For example, for a Hamiltonian composed of Pauli strings, the product of Pauli strings is a Pauli string up to a phase factor  $\pm 1$  or  $\pm i$ . Thus, the overlap matrices are simply expectation values  $\langle \psi | \hat{P} | \psi \rangle$  of some Pauli string  $\hat{P}$ , which can be easily measured (see Sec. II.C). The optimization program 60 is a quadratically constrained quadratic



program (QCQP) with single equality constraint. The Algorithms proceeds in three serial and disjoint steps.

- (1) Select ansatz, which can be done on paper
- (2) Estimate overlap matrices on quantum computer, which can be done efficiently in a parallel fashion.
- (3) Post-processing on a classical computer to solve the QCQP based on the overlap matrices from step 2.

As a major speedup compared to standard variational algorithms, there is no feedback loop between classical and quantum computer such that the calculations can be easily parallelized. The ansatz can be improved by changing  $K$  to  $K + 1$ . The ansatz construction is systematic and there is no trainability issue such as the barren plateau problem (see Sec. IV.B). For the QCQP, there exist conditions which tell whether a local minima is a global minima as a stopping criteria for the classical solver. Moreover, the Lagrangian relaxation of the program 60 is a semidefinite program and efficiently solvable.

#### IV. MAXIMIZING NISQ UTILITY

Current NISQ devices have a limited number of qubits ( $\sim 50 - 100$ ) available. In addition, due to their noisy nature and short coherence time, one can only perform ( $\sim 20 - 30$ ) gate operations. In order to make maximal use of the currently available quantum resources, there are two approaches from the operational point-of-view: bottom-up and top-down. The bottom-up approach refers to the scenario where one has full control over the design of any quantum computing platform to keep pushing the performance quality such as gate fidelity and coherence life time for given hardware design constraints. The top-down approach means one does not get involved in hardware design, and one simply makes do with what has already been made or fabricated in the experimental labs. In this section, we focus on the latter approach, i.e. algorithmically extending or maximizing the utility of current and near-term quantum devices.

##### A. Quantum error mitigation (QEM)

Sensitivity to errors and noise are the two most prominent roadblocks towards scalable universal quantum computers. Fault-tolerant quantum computing can be attained by encoding non-Abelian anyons in topological materials (Kitaev, 2003) or applying quantum error correction codes (Raussendorf and Harrington, 2007). While the former is still in its infancy, the latter mandates physical resources exceeding our current experimental capabilities. In the NISQ era of running hybrid quantum/classical algorithms, it is desirable to use all the restricted and available qubits as logical qubits without

applying QEC techniques. As we discuss throughout this review article, the hybrid quantum/classical algorithms rely on computing the expectation value of some physical observables using quantum processors. Quantum error mitigation (QEM) techniques discussed in this subsection need no extra qubit and can suppress errors in finding expectation values with simple classical post-processing and different runs of quantum circuits. To be precise, with QEM, we are not interested to recover the ideal quantum output state  $\hat{\rho}^{(0)}$ , but to estimate the ideal observables  $\hat{A}$  expectation value:  $E[\mu^{(0)}] = \langle \hat{A}^{(0)} \rangle = \text{Tr}(\hat{\rho}^{(0)} \hat{A})$  (Kandala *et al.*, 2019; Li and Benjamin, 2017; Temme *et al.*, 2017), sometimes surpassing the break-even point, where the effective gates are superior to their physical building blocks, at an affordable cost with respect to near-term quantum hardware (Zhang *et al.*, 2020a). Here,  $\mu$  is the outcome of a measurement and we use superscript (0) to denote an ideal noise-free realization of a state, operation or observable quantity. Recently, it was also shown how to achieve stochastic error mitigation for a continuous time evolution (Sun *et al.*, 2020a). For a comprehensive treatment on quantum error mitigation, refer to (Endo *et al.*, 2020a).

##### 1. Zero-noise extrapolation

(Li and Benjamin, 2017), and (Temme *et al.*, 2017) independently and concurrently proposed the Richardson extrapolation QEM, namely zero-noise extrapolation (ZNE), where a quantum program is to operate at various effective noise levels of a quantum processor. It is then extrapolated to an estimated value at a noiseless level.

Formally, a quantum circuit/system in the presence of noise can be modelled as an open quantum system (Breuer *et al.*, 2002) using the Gorini-Kossakowski-Sudarshan-Lindblad equation or in short the Lindblad master equation (setting  $\hbar = 1$ ):

$$\frac{d}{dt} \hat{\rho}(t) = -i [\hat{K}(t), \hat{\rho}(t)] + \hat{\mathcal{L}}[\hat{\rho}(t)], \quad (61)$$

where  $\hat{K}(t)$  acts as time-dependent driving Hamiltonian, and  $\hat{\mathcal{L}}[\cdot] = \sum_k \Gamma_k (\hat{\mathcal{O}}_k[\cdot] \hat{\mathcal{O}}_k^\dagger - \frac{1}{2} \{ \hat{\mathcal{O}}_k \hat{\mathcal{O}}_k^\dagger, [\cdot] \})$  is a superoperator. The above equation in general describes Markovian dynamics for  $\Gamma_k \geq 0$ . Whenever loss rate  $\Gamma_k$  become negative (Fleming and Hu, 2012), the above equation would also describe non-Markovian dynamics (Bastidas *et al.*, 2018; Kyaw *et al.*, 2020a). In general,  $\Gamma_k$  are fixed by the nature of the noise experienced by a quantum system. Mathematically, one can parametrize  $\Gamma_k$  with a dimensionless scalar  $\lambda$ , i.e.,  $\Gamma_k \rightarrow \lambda \Gamma_k$ . When  $\lambda = 0$ , there is no noise and the second term (loss term) in Eq. (61) is zero, resulting in pure unitary dynamics. When  $\lambda = 1$ , the actual quantum device loss rate is matched. In summary, ZNE involves two steps.

- (1) Noise-scaling: we make a number of measurements  $E[\mu^{(\lambda_j)}]$  for  $\lambda \geq 1$ .
- (2) Extrapolation: we estimate  $E[\mu^{(0)}]$  from the previous step.

*Noise-scaling* can be accomplished in three ways. Firstly, in Ref. (Temme et al., 2017) it was proposed to use a *time-scaling approach* by taking  $\lambda > 1$ , which means that the time-dependent driving Hamiltonian  $\hat{K}(t)$  is now rescaled by  $\frac{1}{\lambda}\hat{K}(t/\lambda)$ . This approach is only possible if the user has full control over back-end quantum processor. Control pulses for each quantum gate have to be recalibrated and be applied for longer duration. Secondly, one can apply a technique called *circuit folding* (Giurgica-Tiron et al., 2020). Suppose that a quantum circuit is composed of  $d$  unitary layers such that  $\mathcal{U} = L_d \cdots L_2 L_1$  where  $d$  refers to the circuit depth and each  $L_j$  either represents a single layer of gate operations or just a single quantum gate. The circuit folding is then achieved by

$$\mathcal{U} \rightarrow \mathcal{U}(\mathcal{U}^\dagger \mathcal{U})^n, \quad (62)$$

where  $n$  is some positive integer. Since  $\mathcal{U}^\dagger \mathcal{U}$  is an identity, this action has no effect on an ideal circuit. However, in noisy circuit,  $\mathcal{U}$  is imperfect and the above  $1 + 2n$  circuit operations would increase the noise level. Thirdly, instead of entire circuit folding, one can also use *gate folding* technique (Giurgica-Tiron et al., 2020):

$$L_j \rightarrow L_j(L_j^\dagger L_j)^n. \quad (63)$$

The second and third techniques do not require users to have full control of quantum computer back-end and thus we expect to be of greater use in software level control of quantum circuits.

*Extrapolation step* of the ZNE method can be considered as a regression problem if we choose to consider a generic model for calculating the expectation value  $E_{\text{model}}[\mu^{(\lambda; \Upsilon)}]$ , where the meaning of *model* would become clear shortly and  $\Upsilon$  corresponds to the model parameters. We note that the expectation value  $E$  is a real number that can only be obtained in the infinite measurement limit. With limited number of measurement samples  $N$ , statistical estimation is  $\hat{E}[\mu^{(\lambda)}] = E[\mu^{(\lambda)}] + \delta$ <sup>1</sup>, where  $\delta$  is a random variable with zero mean and variance  $\sigma^2 = \mathbb{E}(\delta^2) = \sigma_0^2/N$ . Here,  $\sigma_0^2$  is the single-shot variance. Given a set of  $m$  scaling parameters

$\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$  with  $\lambda_j \geq 1$ , and the corresponding measurement outcomes  $\mu = \{\mu_1, \mu_2, \dots, \mu_m\}$ , the ZNE is nothing but to build a good estimator  $\hat{E}[\mu^{(0)}]$  for  $E[\mu^{(0)}]$  such that its bias  $\mathbb{E}(\hat{E}[\mu^{(0)}] - E[\mu^{(0)}])$ , and its variance  $\mathbb{E}(\hat{E}[\mu^{(0)}]^2) - \mathbb{E}(E[\mu^{(0)}])^2$  are both reasonably small. Onwards, let us adopt a simplified notation of  $E[\mu^{(\lambda)}] = E(\lambda)$ . Now let us mention briefly the statistical models. The expectation value  $E(\lambda)$  cannot be of any arbitrary function, which would make ZNE impossible to extrapolate back to  $E(0)$ . Depending on some underlying noise model assumption, one can apply various statistical models.

*The polynomial extrapolation* is based on the polynomial model of degree  $d$  such that

$$E_{\text{poly}}^{(d)}(\lambda) = c_0 + c_1\lambda + \dots + c_d\lambda^d, \quad (64)$$

where  $c_j$  are  $d+1$  unknown real parameters. This extrapolation is justified in weak noise limit and we need the number of data points  $m$  to be equal or larger than  $d+1$ . Consequently, we can obtain two other variants: *the linear extrapolation* ( $d=1$ ) and *the Richardson extrapolation* ( $d=m-1$ ) (Temme et al., 2017). By construction, the error with respect to the true expectation value is  $O(m)$  when we have large sample size  $N \rightarrow \infty$ . By using the interpolating *Lagrange polynomial*, the estimator is explicitly given by:

$$\hat{E}_{\text{Rich}}(0) = \hat{c}_0 = \sum_{k=1}^m \mu_k \prod_{i \neq k} \frac{\lambda_i}{\lambda_i - \lambda_k}, \quad (65)$$

with the assumption that all  $\lambda_j$  are different. One important observation is that the Richardson model based ZNE is dictated by a statistical uncertainty which is exponentially scaling with the number of data points. There are also other statistical models such as *poly-exponential extrapolation* (Giurgica-Tiron et al., 2020) and *exponential extrapolation* (Endo et al., 2018).

The ZNE scheme suffers from a few limitations. The scheme works by extrapolation, and hence it is challenging to obtain result guarantees in general. The number of measurement shots required to obtain the mitigated expectation value can be relatively high compared to the unmitigated case as seen above. For large enough circuit size, the expectation value for a given observable is almost constant as the noise level increases. In such cases, the extrapolation scheme would fail. More importantly, the fundamental drawback of both ZNE and probabilistic error cancellation (PEC) (Temme et al., 2017), or quasi-probability method (which is discussed next) is that one needs to know precise physical noise model in advance (see Eq. (61)), which in itself is a difficult problem. Experimentalists in the lab will have imperfect knowledge about the real noise, which will typically differ from the canonical ones. Hence, instead of discussing PEC, we

<sup>1</sup> We note that the hat notation used is in accordance with statistics notation and it should not be confused with a quantum operator.

would discuss gate set tomography (GST)-based PEC proposed in Ref. (Endo *et al.*, 2018), which does not require explicit knowledge of the noise model and mitigates any localized Markovian errors, so that the error in the final output is only due to unbiased statistical fluctuation.

Let us familiarize ourselves with the notations used in quantum tomography (Greenbaum, 2015; Merkel *et al.*, 2013), which we adopt in the following Sec. IV.A.2. A quantum state is represented by a density matrix  $\hat{\rho}$ , and a physical observable is denoted by a Hermitian  $\hat{A}$  operator. An operation is a map on the states space such that one can use the Kraus representation to denote it as:  $\hat{\mathcal{L}}[\hat{\rho}] = \sum_j \hat{K}_j \hat{\rho} \hat{K}_j^\dagger$ . We note that this equivalence with Eq. (61) is only valid when we have Markovian dynamics. Here,  $\hat{K}_j$  are Kraus operators. In terms of the Pauli transfer matrix representation,  $\hat{\rho}$  in Eq. (61) can be written as a column vector (Navarrete-Benlloch, 2015), denoted as  $|\rho\rangle$ . Similarly, the Lindblad superoperator  $\hat{\mathcal{L}}$  can be recast as a square matrix, since it is a linear map. For simplicity and without loss of generality, we may absorb the unitary dynamics (the first term in Eq. (61)) into  $\hat{\mathcal{L}}$  onwards. A physical observable  $\hat{A}$  is now written as a row vector  $\langle\langle A|$ . Consequently, the expectation value is  $\langle\hat{A}\rangle = \text{Tr}[\hat{A}\hat{\rho}] = \langle\langle A|\rho\rangle\rangle$ . Likewise, the expectation of  $\hat{A}$  after the state  $\hat{\rho}$  passing through a series of linear maps is read as:  $\text{Tr}[\hat{A}\hat{\mathcal{L}}_N \circ \dots \circ \hat{\mathcal{L}}_1(\hat{\rho})] = \langle\langle A|\mathcal{L}_N \dots \mathcal{L}_1|\rho\rangle\rangle$ .

## 2. Probabilistic error cancellation

The central theme of probabilistic error cancellation (PEC) or quasiprobability decomposition introduced by the IBM team in Ref. (Temme *et al.*, 2017) is that one can mitigate effect of errors by sampling from a set of erroneous circuits, labelled by  $\mathcal{L}_{\text{tot}}^{(l)}$  for  $l = 1, 2, \dots$ , such that

$$\langle\hat{A}^{(0)}\rangle = \sum_l q_l \langle\langle A^{(l)}|\mathcal{L}_{\text{tot}}^{(l)}|\rho^{(l)}\rangle\rangle. \quad (66)$$

Given specific error models (assuming experimentalist has full and correct knowledge about them), the real numbers  $q_l$ , which represent quasiprobabilities, can be efficiently derived. Here, each  $\mathcal{L}_{\text{tot}}^{(l)}$  represents the total sequence of noisy gates in the  $l$ th circuit.

GST-based PEC (Endo *et al.*, 2018) can be implemented in three steps.

- (1) Implement gate set tomography.
- (2) Compute the quasiprobability decomposition.
- (3) Implement the computed decomposition using the Monte Carlo approach.

*Implementing GST* refers to measuring all imperfect gates used in the quantum computation: single-qubit gates and two-qubit ones (see Fig. 5). In short, the purpose of GST is to measure noisy individual quantum circuit output a priori. For a single-qubit gate, one prepares initial states  $|0\rangle, |1\rangle, |+_x\rangle$ , and  $|+_y\rangle$ , where  $|+_x\rangle$  and  $|+_y\rangle$  are the eigenstates of Pauli operators  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$  with +1 eigenvalue, respectively. Since we are dealing with noisy devices, these four states are denoted as  $\bar{\rho}_1, \bar{\rho}_2, \bar{\rho}_3$  and  $\bar{\rho}_4$ , accordingly. We also use  $\bar{\mathcal{L}}$  (superoperator) to denote a noisy/imperfect gate to be measured. Since what we care about are expectation value of physical observables, for single-qubit case, we have observables  $\hat{\mathbb{I}}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ , denoted as  $\bar{A}_1, \bar{A}_2, \bar{A}_3, \bar{A}_4$ . The mean value of observables, the  $4 \times 4$  matrix  $\bar{A}$ , is nothing but  $\bar{A}_{j,k} = \text{Tr}[\bar{A}_j \bar{\mathcal{L}} \bar{\rho}_k]$ . Similarly, we can also construct the  $4 \times 4$  matrix  $g$  without applying any gate to the initial states as  $g_{j,k} = \text{Tr}[\bar{A}_j \bar{\rho}_k]$ . This is repeated for each qubit and each single-qubit gate. Statistical estimation of the initial states  $\bar{\rho}_k$  and the observables  $\bar{A}_j$  are then given by

$$|\hat{\rho}_k\rangle = T_{\bullet,k}, \quad (67)$$

$$\langle\langle \hat{A}_j| = (gT^{-1})_{j,\bullet}, \quad (68)$$

where we note that the hat symbol represents the statistical estimate and  $T_{\bullet,k}(T_{j,\bullet})$  denotes the  $k$ th column ( $j$ th row) of the matrix  $T$ , where  $T$  is an invertible  $4 \times 4$  matrix with the following relationship  $\hat{\mathcal{L}} = Tg^{-1}\bar{\mathcal{L}}T^{-1}$ . The same procedure applies for the two-qubit case with the only difference being that there are total of 16 initial states:  $\bar{\rho}_{k_1} \otimes \bar{\rho}_{k_2}$  and 16 observables:  $\bar{A}_{j_1} \otimes \bar{A}_{j_2}$  to be measured. Similarly, we have  $g = g_1 \otimes g_2$  and  $T = T_1 \otimes T_2$ . We have to implement two-qubit gate GST for each qubit pair involved in a quantum program run.

*Quasiprobability decomposition* is computed based on GST results above. From GST, we have estimation of initial states  $|\hat{\rho}_k\rangle$ , observables to be measured  $\langle\langle \hat{A}_j|$ , and gates  $\hat{\mathcal{L}}$ . Let's denote  $\mathcal{L}^{(0)}$  as the Pauli transfer matrix of the ideal gate with no error. The main idea of decomposition comes from a very simple idea that a noisy gate operation comes from an ideal operation followed by a noise, i.e.,  $\mathcal{L} = \mathcal{N}\mathcal{L}^{(0)}$ . Hence, the inverse of the noise is given by

$$\mathcal{N}^{-1} = \mathcal{L}^{(0)}\hat{\mathcal{L}}^{-1} = \sum_i q_{\mathcal{L},i} \hat{\mathcal{B}}_i. \quad (69)$$

Notice that the matrices in the above equation are obtained from the first GST step. The remaining task is to determine quasiprobabilities  $q_{\mathcal{L},i}$  for each qubit and gate involved by solving the above equation. We note that instead of quasiprobabilistic decomposition of quantum gates, one could in principle use randomized compiling technique proposed in Ref. (Wallman and Emerson, 2016).

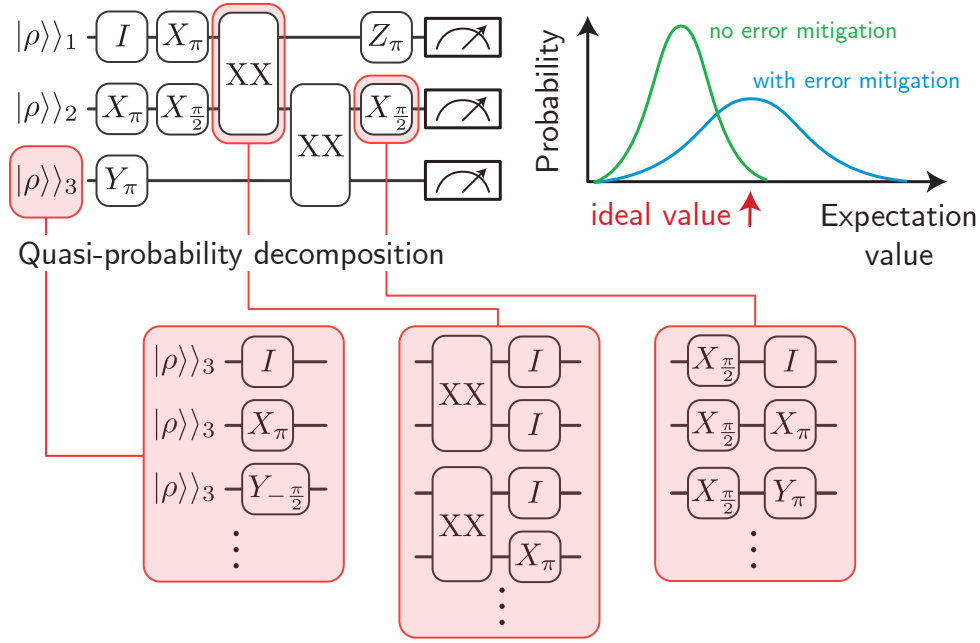


Figure 5 Quantum computing of the expected value of an observable using gate set tomography-based PEC. Quasiprobability decomposition of initial state preparation, exemplary single- and two-qubit processes are computed. Implementing the resulting decomposition is done using the Monte Carlo approach. With QEM, the probability distribution of expected value of the physical observable is now centered around an ideal value with larger variance as compared to the one without QEM. Inspired by (Zhang *et al.*, 2020a).

*Monte Carlo method* is used to compute  $\langle A^{(0)} \rangle$  by randomly choosing the  $l$ th circuit with the probability  $p_l = |q_l|/C$ , where  $C = \sum_l |q_l|$ . Lastly, the computed result is given by the expected value of effective measurement outcomes  $\langle A^{(0)} \rangle = CE[\mu_{\text{eff}}]$ , where the effective outcome is  $\mu_{\text{eff}} = \text{sgn}(q_l)\mu^{(l)}$  if the  $l$ th circuit is chosen and  $\mu^{(l)}$  is the outcome from the  $l$ th circuit. As a consequence, the mean value of the PEC outcome centers near around the ideal one with larger variance due to  $C$  (see Fig. 5). GST-based PEC experiments have recently been done in trapped-ion systems (Zhang *et al.*, 2020a) and superconducting circuits (Song *et al.*, 2019).

### 3. Other QEM strategies

We have seen that the quantum error mitigation techniques discussed so far do not require any ancilla or extra qubits with the caveat that one needs to perform more measurements. At the same time, one is only interested in information about the expectation value. Along this line of thought, there exist several proposals, which we will outline below. However, some of the methods might require ancilla qubits.

*Subspace expansion method* (Barron *et al.*, 2020; Colless *et al.*, 2018; McArdle *et al.*, 2019c; McClean *et al.*, 2020c,

2017; Sagastizabal *et al.*, 2019a) are designed to mitigate errors in the VQE routine, where we often tend to find an approximate ground state  $|\psi_a\rangle$  of a system Hamiltonian  $H$ . However, such state may differ from the true ground state  $|\psi_g\rangle$  due to noisy processes. In general, we do not know which error occurred to the state. The subspace expansion method works by resolving the action of  $H$  on the linear combination of quantum states ansatz Eq. (58). The subspace is spanned by a set of operators  $\hat{O}_i$ , i.e.,  $\{|\hat{O}_i\psi_a\rangle\}$ . Now, one proceeds to evaluate  $H_{ij} = \langle\psi_a|\hat{O}_i H \hat{O}_j|\psi_a\rangle$ , and  $S_{ij} = \langle\psi_a|\hat{O}_i \hat{O}_j|\psi_a\rangle$ . The latter is needed since the subspace states are in general not orthogonal to each other. By solving the generalized eigenvalue problem  $HC = SCE$ , with eigenvectors  $C$  and diagonal matrix of eigenvalues  $E$ , we can obtain the Hamiltonian spectra including the excited states (see Sec. V.A.4). This method requires an appropriate choice of subspace operators to mitigate errors due to external noise. In general, without knowing the noise models of quantum device, it would require an exponential number of expansion operators to obtain the optimal ground-state.

*Stabilizer based approach* (Bonet-Monroig *et al.*, 2018; McArdle *et al.*, 2019c; Sagastizabal *et al.*, 2019a) relies on the information associated with conserved quantities such as spin and particle number conserving ansatz. If



any change in such quantities is detected, one can pinpoint an error in the circuit, which is akin to stabilizer measurement in quantum error correction schemes. We can implement the stabilizer measurements by adding ancilla qubits to the qubit registers or taking additional measurements and post-processing.

*Individual error reduction* method was proposed in Ref. (Otten and Gray, 2019). As we have seen earlier, Markovian noise can be modelled using the Lindblad master equation, Eq. (61), where we have  $\frac{d\hat{\rho}}{dt} = \hat{\mathcal{L}}(\hat{\rho}) = \sum_i \mathcal{L}_i(\hat{\rho})$ , where each  $\mathcal{L}_i$  denotes a noise channel present. Here, we have absorbed the unitary component into  $\hat{\mathcal{L}}$ . It was shown that

$$\tilde{\rho}(T) = \hat{\rho}(T) - \sum_{j=1}^m \frac{1}{g_j} (\hat{\rho}(T) - \hat{\rho}_j(T)), \quad (70)$$

$$= \hat{\rho}^{(0)}(T) + \mathcal{O}(\tau^2). \quad (71)$$

Notational explanations are as follow.  $\hat{\rho}(T)$  is the density matrix after applying quantum gates with the presence of all associated noise channels at the final evolution time  $T$ . In contrast,  $\hat{\rho}_j(T)$  is the state under the influence of all the noise channels but one less  $\mathcal{L}_j$  according to the ratio  $g_j$ . Notice that if  $g_j = 1$ , we have fully removal of the entire channel  $\mathcal{L}_j$ .  $\hat{\rho}^{(0)}(T)$  is the ideal output state without any error, while  $\tau$  is the evolution time for each noise process after the gate application. We note that the first-order error  $\mathcal{O}(\tau)$  is removed. As usual, what we want to obtain is  $\langle \hat{A} \rangle = \text{Tr}[\hat{\rho}^{(0)}(T)\hat{A}]$  for a physical observable  $\hat{A}$ . We can arrive at it by using Eq. (70). Though its result is neat and beautiful, this method assumes a perfect removal of individual noise channel. Hence, it is relatively unrealistic on current quantum hardware as compared to other strategies.

*Dynamical decoupling* (DD) (Santos and Viola, 2005; Viola and Knill, 2005; Viola et al., 1999) is a very well-known and widely used quantum control technique in the literature. In short, it is designed to suppress decoherence via fancy pulses to the system so that it cancels the system-bath interaction to a given order in time-dependent perturbation theory (Lidar, 2014). Recently, DD experiments were performed on the 16-qubit IBMQX5, 5-qubit IBMQX4, and the 19-qubit Rigetti Acorn chips and the gain in substantial gate fidelity relative to unprotected, free evolution of individual transmon qubits was demonstrated in Ref. (Pokharel et al., 2018).

*Pulse shaping technique* is a strategy for passive cancellation of system-bath interaction, in contrast to DD as a method for active cancellation of system-bath interaction. Traditionally, this method stands on the shoulder

of a mean to obtain high-fidelity quantum gate in nonlinear qubits such as transmons, commonly known as derivative removal of adiabatic gate (DRAG) scheme (De, 2015; Gambetta et al., 2011; Motzoi et al., 2009). With the availability of IBM Qiskit Pulse (McKay et al., 2018) that allows users to control backend pulse shapes and sequences of a quantum processor on the fly, a recent study (Gokhale et al., 2020) achieved both  $1.6\times$  lower error rates and  $2\times$  faster execution time, relative to standard gate-based compilation. In light of these recent developments, we envisage a possibility to realize/encode holonomic quantum gates (Zanardi and Rasetti, 1999; Zhang et al., 2015) which are robust against parameter fluctuations and attain even better gate fidelity.

*Lanczos-inspired approach* (Suchsland et al., 2020) estimates the expectation value of a physical observable  $\text{Tr}[\hat{\rho}^{(0)}\hat{A}]$  by constructing a basis of the order- $m$  Krylov subspace  $\mathcal{K}^{(m)}$ . For  $m = 2$ , it is shown that the estimate is given by  $\text{Tr}[\hat{\rho}^e(a_0, a_1)\hat{A}]$ , where

$$\hat{\rho}^e(a_0, a_1) = \frac{(a_0 - a_1\hat{A})\hat{\rho}(a_0 - a_1\hat{A})}{\text{Tr}[\hat{\rho}(a_0 - a_1\hat{A})^2]}, \quad (72)$$

with  $\forall a_0, a_1 \in \mathbb{R}$ . As seen above, due to the Krylov expansion, this technique can reduce the impact of different sources of noise with cost of an increase in the number of measurements to be performed, without additional experimental overhead.

*Depolarizing error mitigation* was proposed in Ref. (Vovrosh et al., 2021). Depolarizing errors are one of the main contributions to errors on current NISQ devices. Assuming that only these types of errors are occurring, a simple ansatz captures the effect of the depolarizing channel

$$\rho = (1 - p_{\text{tot}})\rho_{\text{exact}} + p_{\text{tot}} \frac{I^{\otimes N}}{2^N}, \quad (73)$$

with the total error probability  $p_{\text{tot}}$ . One can show that the error-free value of an observable can be reconstructed by knowing the purity of the prepared quantum state  $\text{Tr}(\rho^2)$ , which can be efficiently measured by randomized measurements (Elben et al., 2018). With this device-independent approach, quantitative correct simulations of quantum problems have been performed on IBM quantum processors.

## B. Barren plateaus

Recently, it was shown that the expectation value of the gradient of the objective function corresponding to random initialized PQC (RPQC) decays exponentially to

zero as a function of number of qubits (McClellan *et al.*, 2018). In particular, the authors considered an  $L$ -layered RPQC of the form of Eq. (26). The mathematical basis of this result hinges on the fact that the PQC in Eq. (26) become unitary 2-design as the circuit depth increases polynomially in the circuit width i.e., number of qubits. The notion of unitary 2-design has been used extensively in the recent proofs of barren plateau in RPQCs, which necessitates a small discussion about their mathematical structure.

The primary objective of unitary  $t$ -designs is to substitute the integration over the space of unitaries with a finite sum. The aforementioned approach provides an accessible way to find the average of functions over unitaries and then prove intriguing theorems about them. The set of unitaries  $U(d)$  forms a topologically compact and connected group. Furthermore, they correspond to the set of norm-preserving matrices in  $\mathbb{C}^d$ . We can define a unique “translation-invariant” measure, i.e. *Haar measure* on  $U(d)$ , which can be employed to calculate expectation values of functions of unitaries, i.e.,

$$\langle f \rangle = \int_{U(d)} f(U) dU. \quad (74)$$

We present the aforementioned statement in a relatively formal manner.

**Definition 4** Let  $\mathbb{U}$  be a group of  $n \times n$  unitaries. A probability measure  $\mu$  on  $\mathbb{U}$  is called *Haar measure* on  $\mathbb{U}$  if for any subset  $\mathbb{S} \subseteq \mathbb{U}$  and for any fixed  $K \in \mathbb{U}$ , we have

$$\mu(K\mathbb{S}) = \mu(\mathbb{S}K) = \mu(\mathbb{S}) \quad (75)$$

where  $K\mathbb{S} \equiv \{KU : U \in \mathbb{S}\}$  and  $\mathbb{S}K \equiv \{UK : U \in \mathbb{S}\}$ . The property in Eq. (75) is called *translation invariance*.

Because of translation invariance, if  $U_1$  is a Haar-distributed random unitary matrix, then so are  $UU_1$  and  $U_1U$  for a fixed unitary matrix  $U$ . To present the concept of unitary designs, we need to first discuss the notions of *homogeneous polynomials* ( $\text{Hom}(a, b)$ ) and *weight functions*, defined in the context of unitary matrices. Any polynomial of degree  $a$  in the entries of  $U \in U(d)$  and  $b$  in  $U^\dagger$  is called a homogeneous polynomial  $\text{Hom}(a, b)$ . Some of the examples of homogeneous polynomials are  $U^\dagger V^\dagger U V \in \text{Hom}(2, 2)$  and  $V U^\dagger V U \in \text{Hom}(3, 1)$ . A weight function on a set  $S$ ,  $w : S \rightarrow (0, 1]$ , for all  $U \in S$  satisfies the following two properties: (i)  $w(U) \geq 0$  and (ii)  $\sum_{U \in S} w(U) = 1$ . Having discussed the concepts of homogeneous polynomials and weight function, we proceed to define the notion of unitary  $t$ -designs.

**Definition 5** *Unitary  $t$ -design:* Given a finite set  $S \subset U(d)$  and a weight function  $w : S \rightarrow (0, 1]$ , the tuple  $(S, w)$  is called a *unitary  $t$ -design* if for all  $f \in \text{Hom}(t, t)$

$$\sum_{U \in S} w(U) f(U) = \int_{U(d)} f(U) dU. \quad (76)$$

We will refer to unitary  $t$ -design as  $t$ -design henceforth. Note that to verify if a tuple  $(S, w)$  forms a  $t$ -design, we need to check condition 76 for every function  $f$  in  $\text{Hom}(t, t)$ . For many cases, such an approach might be impractical and hence it requires more tractable condition. We discuss a more tractable condition in the form of the following Lemma.

**Lemma 6** *Given a finite set  $S \subset U(d)$  and a weight function  $w : S \rightarrow (0, 1]$ , the tuple  $(S, w)$  forms a  $t$ -design if and only if  $\sum_{U \in S} w(U) U^{\otimes t} \otimes (U^\dagger)^{\otimes t} = \int_{U(d)} U^{\otimes t} \otimes (U^\dagger)^{\otimes t} dU$ .*

Using the notion of 2-design, the appearance of barren plateaus in the training landscape has been established for various kind of ansätze. Barren Plateaus can be thought of as a consequence of the exponentially large dimension of the Hilbert space when the number of qubits increases and the fact that the variational circuit unitary, when the parameters are initialized at random, is a 2-design. Consequently, the strategies proposed to tackle this problem focus on reducing the space dimension of this unitary or breaking the randomness properties related to the 2-designs. Another way to think of the origin of the barren plateau issue could be the problem-agnostic nature of the ansatz, faced with exponentially large parameter space. Thus, one could attempt to devise ansätze as well as the optimization methodology in a problem-aware manner by using physically-inspired or problem-specific ansätze presented in Sec. II.B or those proposed in Sec. III.E.

Besides the exponential parameter space induced barren plateaus, other physical phenomena can also generate them. In particular, the noise and decoherence present in the quantum computing experiments also generates this problem in VQAs (Wang *et al.*, 2020b). Entanglement-induced barren plateaus have also been reported recently (Marrero *et al.*, 2020).

While certain ansätze can be assumed or proven to form (approximate) 2-designs, such proofs are challenging for the general ones. To numerically verify the presence of barren plateaus, past studies often considered computing the gradients and variances of a local observable using a particular ansatz over increasing system sizes (McClellan *et al.*, 2018; Skolik *et al.*, 2020).

Another attempt to avoid a barren plateau is to initialize the variational circuit with a particular state choice. Intuitively, the algorithm will start in a particular region of the Hilbert space allowing the optimization subroutine to potentially find the minima in a closer region. Any algorithm that starts with a particular initial state will contribute to avoiding the generation of Barren Plateaus. This include all physically inspired strategies mentioned in Sec. II.B. The use of clever encodings for the algorithm parameters can also be understood as a initialization strategy (Cervera-Lierta *et al.*, 2020). For further details, refer to Sec. V.B. Classical algorithms such as

neural networks can also be used to learn the proper circuit encodings (Verdon *et al.*, 2019a).

A good choice for the initial state is often not enough to reduce the size of the Hilbert space. Although expressible circuit ansätze are usually a requirement for the success of a VQA (see Sec. IV.C for more details), it can expand the parameter space that the optimizer has to explore. Several works propose circuit structures that reduce that space by introducing correlations between the variational parameters of the circuit (Volkoff and Coles, 2020), block-wise initialization of those parameters (Grant *et al.*, 2019) or exploring particular ansatz structures (Sharma *et al.*, 2020).

In most cases, the use of clever initialization and circuit ansatz strategies do not necessarily guarantee the absence of Barren Plateaus. Moreover, some of the mentioned works require a circuit design that is not necessarily hardware efficient. Other ideas focus on the classical parts of the VQA instead of the quantum circuit designs. One example is using local instead of global cost functions for the optimization. It has been shown (Cerezo *et al.*, 2020c) that Barren Plateaus also emerge in shallow depth circuits, and that the use of local cost functions reduces the exponential decay tendency to a polynomial one. The optimization strategy may also reduce the effect of the vanishing gradients, for instance by training the circuit layer by layer (Lyu *et al.*, 2020; Skolik *et al.*, 2020) or by measuring low depth gradients (Harrow and Napp, 2019).

Barren plateaus are a roadblock in the trainability and hence any PQC ansatz which suffers from barren plateau will fail to properly train the parameters in its search for near-optimal (or optimal) performance. As shown in Ref. (Arrasmith *et al.*, 2020a), even gradient-free approaches not necessarily always alleviate the severity of the barren plateaus. However, one can circumvent this issue by using hybrid quantum states of the form of equation Eq. (58) or hybrid density matrices as introduced in Ref. (Haug and Bharti, 2020). The idea is to write the overall ansatz as a classical combination of quantum states i.e.,  $|\phi(\alpha(t), \theta)\rangle = \sum_{i=0}^{m-1} \alpha_i(t) |\psi_i(\theta_i)\rangle$ . Tuning the  $\theta_i$ s can often lead to barren plateaus. One can avoid such issues by fixing  $\theta_i$ s by harnessing the structure of the problem to find the basis states of the ansatz i.e.,  $\{|\psi_i(\theta_i)\rangle\}$  (see (Bharti and Haug, 2020a,b; Haug and Bharti, 2020; Huang *et al.*, 2019)). The  $\alpha_i$ s are classical parameters and can be tuned without changing the basis states.

### C. Expressibility of variational ansätze

A cornerstone in the success of variational algorithm is choosing the proper ansatz for the problem. In addition to trainability, i.e. how well the ansatz can be optimized, another major quality is expressibility. This concerns whether a given PQC is able to generate a rich class of

quantum states. The expressibility as a figure-of-merit has been studied in (Nakaji and Yamamoto, 2020a; Sim *et al.*, 2019) by introducing the concepts of *expressibility* and *entangling capability*.

*Expressibility.* A given PQC  $U(\theta)$  generates the states  $|\psi(\theta)\rangle = U(\theta)|0\rangle$ . Now, one randomly samples from the PQC and generates a distribution of states. Expressibility is defined as the deviation of this distribution from the Haar measure (see Definition 4), which samples uniformly from the full Hilbert space

$$A_U^{(t)} = \left\| \int_{\text{Haar}} (|\psi\rangle\langle\psi|)^{\otimes t} d\psi - \int_{\theta} (|\psi_{\theta}\rangle\langle\psi_{\theta}|)^{\otimes t} d\psi_{\theta} \right\| \quad (77)$$

where  $\int_{\text{Haar}}$  denotes the integration over a state  $|\psi\rangle$  distributed according to the Haar measure. An ansatz circuit  $U$  with small  $A_U^{(t)}$  is more expressible, with  $A_U^{(t)} = 0$  corresponding to being maximally expressive, as it generates quantum states with a distribution closer to the Haar measure. The PQC samples uniformly from the full Hilbert space and thus is able to approximate any possible state. This is especially important in the case where one wants to train the PQC to represent a particular quantum state while having little prior information about the state. A highly expressible PQC is more likely to be able to represent the target state. With this measure, it is possible to compare the expressive power of different types of PQCs.

*Entangling capability.* This method uses the power of a PQC to create entangled states as a way to quantify the expressiveness of an ansatz. Various methods to quantify entanglement exist. In (Sim *et al.*, 2019) the Meyer-Wallach  $Q$  measure has been proposed to estimate the number and types of entangled states a particular PQC can generate. For a system of  $n$  qubits define a linear mapping  $\iota_j(b)$  that acts on the computational basis

$$\iota_j(b) |b_1 \dots b_n\rangle = \delta_{b_j} |b_1 \dots \hat{b}_j \dots b_n\rangle, \quad (78)$$

where  $b_j \in \{0, 1\}$  and the symbol  $\hat{\phantom{x}}$  denotes absence of the  $j$ -th qubit. The entanglement measure  $Q$  is then defined as

$$Q(|\psi\rangle) \equiv \frac{4}{n} \sum_{j=1}^n D(\iota_j(0)|\psi\rangle, \iota_j(1)|\psi\rangle), \quad (79)$$

with the generalized distance  $D$  defined as

$$D(|u\rangle, |v\rangle) = \frac{1}{2} \sum_{i,j} |u_i v_j - u_j v_i|^2, \quad (80)$$

with  $|u\rangle = \sum u_i |e_i\rangle$  and  $|v\rangle = \sum v_i |e_i\rangle$ , where  $\{|e_i\rangle\}$  denotes the computational basis set. The entangling capability of a PQC is then defined as the average  $Q$  of states

randomly sampled from the circuit.

$$\text{Ent} = \frac{1}{|S|} \sum_{\theta_i \in S} Q(|\psi_{\theta_i}\rangle), \quad (81)$$

where  $S = \{\theta_i\}_i$  is the set of sampled circuit parameters. When the PQC can only generate product states  $Q$  is minimal with  $Q = 0$ , whereas for highly entangled states it will reach the maximum of  $Q = 1$ .

In (Nakaji and Yamamoto, 2020a; Sim *et al.*, 2019), a wide class of circuits have been investigated with above expressibility measures. It has been found that certain types of ansätze are more expressible, e.g. layered PQCs consisting of CNOT are more expressive than CZ. It has more recently been shown that there is a trade-off between an ansatz being expressive and trainable. Making an ansatz more expressible most likely will result in reducing the gradient of the objective function (Holmes *et al.*, 2021). Given this trade-off, the authors suggest several strategies for reducing expressibility (and improving trainability), including correlating parameters and restricting rotation directions and/or angles of parameterized gates. For Alternating Layered Ansatz (see Sec. II.B.2), a favorable trade-off was found as this type of ansatz is both relatively expressive as well as does not exhibit barren plateaus in a certain regime (Nakaji and Yamamoto, 2020a).

#### D. Reachability

Reachability discusses the question whether a given PQC  $|\psi(\theta)\rangle$  with parameter  $\theta$  is capable of representing a quantum state that minimizes some objective function. This can be quantified by the reachability deficit over finding the minimum of an objective function  $\hat{O}$  (Akshay *et al.*, 2020a)

$$f_R = \min_{\psi \in \mathcal{H}} \langle \psi | O | \psi \rangle - \min_{\theta} \langle \psi(\theta) | O | \psi(\theta) \rangle \quad (82)$$

where the first term on the right side is the minimum over all states  $|\psi\rangle$  of the Hilbertspace, whereas the second term is the minimum over all states that can be represented by the PQC. The reachability deficit is equal or greater zero  $f_R \geq 0$ , with  $f_R = 0$  when the PQC can generate a state  $|\psi(\theta^*)\rangle$  with parameter  $\theta^*$  that minimizes the objective function.

Reachability has been studied in-depth for QAOA. Although QAOA has been shown to exhibit quantum computational universality (Lloyd, 2018; Morales *et al.*, 2019), which implies that any unitary operator is reachable under the QAOA ansatz, this statement does not hold true for finite fixed depths  $p$ . In fact, it was recently found by (Akshay *et al.*, 2020a) that QAOA exhibits reachability deficits for the MAX-2-SAT and MAX-3-SAT problems, where the optimal value of the objective function cannot be found using a fixed circuit depth  $p$

beyond a critical clause density (defined as the ratio between the number of clauses and the number of variables in the problem) (Akshay *et al.*, 2020a). In other words, for problems with a certain clause density, there is a critical depth  $p^*$  for which the optimal solution can only be found (up to a threshold) if  $p \geq p^*$ , and  $p^*$  grows with the clause density. This finding sets strong limitations on the performance of variational quantum algorithms and is related to the expressibility problem where the set of reachable states of the fixed-depth QAOA for  $p < p^*$  does not contain the states which are close to the optimal solution.

Similar reachability deficits have also been found in the variational Grover search problem. Moreover, by re-analyzing the experimental data from Google's Sycamore quantum processor on the application of QAOA to various graph optimization problems (Arute *et al.*, 2020b), Akshay *et al.* also discovered reachability deficits in this case, where the graph density (defined as the ratio between the number of graph edges to the number of graph nodes) replaces the clause density as the order parameter (Akshay *et al.*, 2020b).

Note that the reachability deficits are distinct from the barren plateau problem, where the gradients of the objective function concentrate to zero for many choices of initial variational parameters, thus slowing down the optimization process. On the other hand, the reachability deficit for  $p < p^*$  is independent of the initial parameters.

#### E. Theoretical guarantees of the QAOA algorithm

The QAOA has several key analytical results which have contributed to its considerable interest in recent years.

The quantum advantage of QAOA algorithm has been studied in (Farhi and Harrow, 2016), where they showed that the efficient sampling of the output distribution of QAOA, even for the lowest depth case of  $p = 1$ , implies the collapse of the polynomial hierarchy (see Sec. I.A for more details about complexity theory). Following the conjecture from complexity theory that the polynomial hierarchy does not collapse, this result propels QAOA as a possible candidate for establishing some quantum advantage.

For the case where the problem Hamiltonian  $H_P$  takes the form

$$H_P = \sum_i \omega_A \hat{\sigma}_z^{2i} + \omega_B \hat{\sigma}_z^{2i+1} + \gamma_{AB} \hat{\sigma}_z^{2i} \hat{\sigma}_z^{2i+1} + \gamma_{BA} \hat{\sigma}_z^{2i+1} \hat{\sigma}_z^{2i+2}, \quad (83)$$

where  $\omega_{A(B)}$  are the coefficients for the even (odd) sites and  $\gamma_{AB(BA)}$  are the interaction strength between first (second) neighbour spins. Taking  $H_M$  as defined in Eq. (22) for a 1D lattice, Ref. (Lloyd, 2018) showed that QAOA can be used to implement universal quantum computation. This result was proven and generalized in



a later work (Morales *et al.*, 2019) to include a larger class of problem and mixing Hamiltonians that can provide computational universality for QAOA. By connecting VQA with optimal control theory, the Pontryagin’s minimum principle of optimal control is used to show that the bang-bang protocol (in which the evolution switches abruptly between two Hamiltonians) is optimal for a fixed total time  $T$  (Yang *et al.*, 2017). Since QAOA can be regarded as a bang-bang ansatz by switching between unitary evolution under  $H_P$  and  $H_M$  respectively, this suggests the optimality of QAOA as a VQA. However, recent works have challenged this claim. By generalizing the argument in Ref. (Yang *et al.*, 2017), it has been shown that the optimal protocol actually possesses the ‘bang-anneal-bang’ structure (Brady *et al.*, 2020). Such protocols begin and end with a bang, with regions of smoothly varying control function (akin to quantum annealing) in between. It was also shown that when the total time  $T$  is large, bang-bang QAOA suffers from the proliferation of local minima in the control parameters, rendering it difficult to find optimal (or near-optimal) QAOA parameters.

## F. Circuit compilation

As it will be discussed in Sec. VII, a quantum computer is composed of its hardware (quantum) and software (classical). The software translates a quantum algorithm to a set of instructions that implement the desired quantum operations and read out the qubit states. This process can be understood as *quantum compilation* (Chong *et al.*, 2017), but the term is not limited to this particular application. When mapping a quantum circuit to a specific device architecture, one needs to consider the available quantum gates, the qubits connectivity that allows two-qubit gates implementation, and experimental limitations such as decoherence time, which imposes a certain circuit depth in terms of the number of gates. For these reasons, it has become indispensable to develop tools that allow for circuit simplifications and efficient mappings to the real hardware. These tools are also known as quantum compilers since they translate the theoretical circuit to the realistic simulation or device. In the following lines, we describe some of these tools. Many of them are suited both for NISQ and fault-tolerant quantum computation.

### 1. Native and universal gate sets

The available gates that can be implemented experimentally are sometimes referred to as the *native gate set*. With a universal gate set  $\mathcal{G} \in SU(d)$  (also called *instruction gate set*), any unitary operation can be constructed efficiently. More formally, the *Solovay-Kitaev theorem*

(Dawson and Nielsen, 2006) states that giving this universal set  $\mathcal{G}$ , any unitary operation  $U \in SU(d)$  can be approximated with  $\epsilon$  accuracy with a finite sequence  $S$  of gates from  $\mathcal{G}$ . This sequence scales logarithmically as  $\mathcal{O}(\log^c(1/\epsilon))$ , where  $c$  is a constant that depends on the theorem proof. For  $d = 2^n$  this theorem guarantees that qubit quantum circuits can be decomposed efficiently using a finite gate sequence. Although this is one of the most important theorems in quantum computation, it is an existence theorem, i.e. it does not provide the efficient decomposition that it predicts.

Clifford group is an important object in quantum information science because of its applications in quantum error correction, randomized benchmarking and investigations for quantum advantage. The generalized Pauli operators in prime dimension  $p$  are given by

$$T_{(a,b)} = \begin{cases} \omega^{-\frac{ab}{2}} Z^a \mathcal{X}^b & (a,b) \in \mathbb{Z}_p \times \mathbb{Z}_p, p \neq 2 \\ i^{ab} Z^a \mathcal{X}^b & (a,b) \in \mathbb{Z}_2 \times \mathbb{Z}_2, p = 2 \end{cases} \quad (84)$$

where  $\omega = \exp\left(\frac{2\pi i}{p}\right)$  and  $\mathbb{Z}_p$  denotes integer modulo  $p$ . The  $Z$  and  $\mathcal{X}$  operators are defined via their action on computational basis states  $\{|k\rangle\}_k$ , with  $\mathcal{X}|k\rangle = |k+1 \bmod p\rangle$ , and  $Z|k\rangle = \omega^k|k\rangle$ . The unitaries which map the set of generalized Pauli operators to themselves up to a phase are called *Clifford unitaries*. Let us denote the set of  $p$  dimensional Clifford unitaries by  $\mathcal{C}_p$ . Mathematically speaking,

$$U \in \mathcal{C}_p \iff \exists \phi : UT_{(a_1,b_1)}U^\dagger = \exp(i\phi)T_{(a_2,b_2)} \quad (85)$$

where  $T_{(a_1,b_1)}$  and  $T_{(a_2,b_2)}$  are generalized Pauli operators. The set of Clifford unitaries  $\mathcal{C}_p$  forms a group, called *Clifford group*.

There are technically infinite universal gate sets, but the Clifford group’s gates compose the most commonly used (in particular, the  $H$ ,  $S$  and  $CNOT$  gates) and the  $T$  gate. The Clifford group alone can be simulated efficiently utilizing the *Gottesman-Knill theorem* (Aaronson and Gottesman, 2004), proving that no quantum advantage can be found without the use of the  $T$  gate. For this reason, many algorithms try to simplify and reduce quantum circuits to the minimal number of  $T$  gates, giving an estimation of the classical efficiency of that particular circuit (Amy *et al.*, 2013; Amy and Mosca, 2019; Gosset *et al.*, 2013; Heyfron and Campbell, 2018; Kissinger and van de Wetering, 2019).

Besides these minimal reduction algorithms, other basic decompositions are useful. Even if only a native gate set is available experimentally, other basic gates can be constructed and used in the algorithms. As an example,  $S$  and  $T$  gates are particular cases of the single-qubit rotational gate around the  $z$  axis, i.e.  $R_z(\theta) = e^{-i\theta\sigma_z/2}$ , and  $H$  gate can be obtained from  $R_y(\theta) = e^{-i\theta\sigma_y/2}$  and  $R_x(\theta) = e^{-i\theta\sigma_x/2}$  gates as  $H = R_y(-\pi/2)R_x(\pi)$ . Any single-qubit gate can be decomposed into the following

gate sequence:

$$U(\theta, \phi, \lambda) = R_z(\phi)R_y(\theta)R_z(\lambda). \quad (86)$$

This motivates using single-qubit rotational gates and at least one entangling gate (e.g. the *CNOT* or *CZ* gate) as native gate sets. Any two-qubit gate can be obtained from this minimal set by using circuit identities (Barenco *et al.*, 1995; Blaauboer and De Visser, 2008; de Guise *et al.*, 2018; Watts *et al.*, 2013). The particular choice of the entangling gate can be motivated from the experimental platform. Depending on the technology used to construct the quantum device, a natural 2-qubit gate implementation can be more suited. Some examples are the use of *CZ* gates in tunable superconducting circuits (Kjaergaard *et al.*, 2020; Krantz *et al.*, 2019), cross-resonance gates in fixed frequency superconducting qubits (Krantz *et al.*, 2019), or the *XX* gates in trapped ions (Häffner *et al.*, 2008; Kjaergaard *et al.*, 2020). More expressive gate sets with continuous gate parameters or long-range interactions can be achieved by further control over the hardware parameters in time (Bastidas *et al.*, 2020; Foxen *et al.*, 2020; Krinner *et al.*, 2020; Lacroix *et al.*, 2020). Interactions involving multiple qubits are possible by periodically driving the system’s parameters in time (Bastidas *et al.*, 2020).

## 2. Circuit decompositions

Once the native gate set is fixed, the next step consists of decomposing the theoretical unitary circuit into this basic set. A raw translation of each single and two-qubit gates into the native set might imply a large circuit depth, reducing the effectiveness of that decomposition. Moreover, to find the decomposition of gates involving more than one qubit might prove challenging in general. Besides common circuit identities mentioned before (such as controlled-qubit gates decompositions), one may need mathematical tools to understand and derive general circuit reductions to particular smaller pieces.

One of these mathematical tools is the so-called *ZX-Calculus*. It is a graphical language that maps quantum circuits to particular graph representations and derives a set of rules to manipulate these graphs. Its application range goes from measurement-based quantum computation to quantum error correction. For a complete review about ZX-calculus and its variety of applications, see Ref. (van de Wetering, 2020). For the purpose of this review, we are interested in the quantum circuit simplification applications (de Beaudrap *et al.*, 2019; Cowtan *et al.*, 2019b; Duncan *et al.*, 2020; Hanks *et al.*, 2020; Kissinger and van de Wetering, 2020b). There exist an open-source library for ZX-calculus manipulation called *PyZX* (Kissinger and van de Wetering, 2020a).

Other approaches use well-known artificial intelligence algorithms to find optimal circuit decompositions, for in-

stance, the use of reinforcement learning (Pirhooshayan and Terlaky, 2020; Zhang *et al.*, 2020b). Evolutionary algorithms such as *genetic algorithms* have been widely studied (Bang and Yoo, 2014; Lamata *et al.*, 2018a; Las Heras *et al.*, 2016; Li *et al.*, 2017b; Massey *et al.*, 2004; Massey *et al.*, 2006; Potoček *et al.*, 2018; Spagnolo *et al.*, 2017; Williams and Gray, 1999). In these approaches, multiple random circuits composed by the native gate set are generated and evolved later on. The evolution strategy includes the definition of possible *mutations* such as introducing a new gate in a particular qubit, the swap between circuit gates or the deletion of a particular gate. Then, a multi-objective loss function is used to estimate the success of each circuit family until a given convergence, when the circuit with the best performance is selected. These works have to be added to those focusing finding the optimal PQC for a given VQA, as discussed in Sec. II.B. A VQA for circuit compilation using a genetic algorithm as optimization subroutine is presented in Ref. (Khatri *et al.*, 2019), called Quantum Assisted Quantum Compiler.

## 3. The qubit mapping problem

After decomposing the quantum circuit into the native gates and simplifying it, a hardware-specific task remains: mapping the resultant circuit to the particular qubits connectivity or topology. Due to experimental limitations, in general, not all qubits are connected, which means that two-qubit operations are not always possible. A naive approach to circumvent this limitation consist of swapping each qubit state with its neighbour (by using *SWAP* gates) until we find a connected pair, perform the desired two-qubit gate and swap again the state of the qubits involved, returning to the original pair. This translates into a significant growth of the circuit depth for circuit topologies with far connected qubits.

Some NISQ algorithms presented in this review may include the qubits’ connectivity by means of the loss function or the available rules used to decompose the unitaries. However, this is not always enough. Moreover, quantum compilation is a hardware-specific transformation; it might be more useful to apply this step independently of the quantum circuit and depending on the chip architecture. Unfortunately, the qubit mapping problem is NP-complete (Botea *et al.*, 2018). Several heuristic approaches based on dynamic programming and depth partitioning have been explored (Cowtan *et al.*, 2019a; Li *et al.*, 2019; Siraichi *et al.*, 2018; Zulehner *et al.*, 2018; Zulehner and Wille, 2019). Exact methodologies based on reasoning engines such as Boolean satisfiability solvers have also been proposed (Wille *et al.*, 2019). The so-called LHZ architecture is an approach that solves the connectivity issue at the cost of increasing the number of qubits. With this architecture the problem is mapped

from  $n$  to  $n^2$  qubits, but this allows for parallel execution of  $CNOT$  gates for an all-to-all coupled network (Lechner, 2020). The same framework can be applied to a quantum annealing system as well (Lechner et al., 2015).

## V. APPLICATIONS

### A. Many-body physics and chemistry

Understanding the static and dynamic properties of quantum mechanical systems is a core challenge at the heart of many fields such as chemistry and physics. Numerical methods on classical computers to solve these quantum mechanical problems struggle when the number of constituent particles is scaled up due to exponential increase of resources needed. Owing to their quantum mechanical nature, quantum computers offer a way to simulate even large-scale many-body systems. Here, we now review various NISQ algorithms that have been proposed to tackle quantum chemistry and many-body physics related problems. We start by introducing concepts on mapping physical problems onto the quantum computer. Then, we introduce algorithms for common challenges, such as finding the static as well as dynamic properties of quantum systems in various settings. All NISQ algorithms discussed in this section are listed in Table I.

#### 1. Qubit encodings

In general, any physical system can be written in terms of a Hamiltonian which is the sum of its kinetic and potential energy. In quantum theory, each physical system is associated with a language of operators and an algebra establishing such language. Depending on the system constituents, there are three types of particles (operators) in Nature: fermions, bosons and anyons. The first two are elementary particles obeying Fermi-Dirac (FD) statistics and Bose-Einstein (BE) statistics, respectively. The latter being quasiparticles obeying continuous or anyonic statistics, and existing only in two-dimensional confinement. Quantum computers (QC) operate in a language of qubits (a distinguishable set of spin-1/2 particles). Hence, the quantum simulation of a physical system refers to performing a one-to-one mapping from the system operator language to the QC language, preserving the underlying statistics.

In the standard model of quantum computation, a two-level system or spin-1/2 particle is denoted by its spin orientation  $|\uparrow\rangle = |0\rangle = (1, 0)^T$  and  $|\downarrow\rangle = |1\rangle = (0, 1)^T$ . An  $N$ -qubit system is then constructed from the standard Pauli matrices  $\hat{\sigma}_x^i, \hat{\sigma}_y^i, \hat{\sigma}_z^i$ , where the superscript  $i$  refers to the  $i^{\text{th}}$  local qubit site. These operators satisfy the commutation relations of an  $\bigoplus_{i=1}^N su(2)_i$  algebra

$[\hat{\sigma}_\mu^l, \hat{\sigma}_\nu^m] = 2i\delta_{lm}\epsilon_{\mu\nu\lambda}\hat{\sigma}_\lambda^l$ , where  $\epsilon_{\mu\nu\lambda}$  is the totally antisymmetric Levi-Civita symbol with  $\mu, \nu, \lambda \in \{x, y, z\}$ .

**Fermions.** In the second quantized notation,  $N$  fermions are denoted by fermionic operators  $\hat{f}_i^\dagger(\hat{f}_i)$ , the creation (annihilation) operators of a fermion in the  $i^{\text{th}}$  mode/site ( $i = 1, \dots, N$ ). The fermionic operators obey Pauli's exclusion principle and the antisymmetric nature of the fermion wave function. Hence, the fermionic algebra is defined by the anti-commutators  $\{\hat{f}_i, \hat{f}_j\} = 0, \{\hat{f}_i^\dagger, \hat{f}_j\} = \delta_{ij}$ . There are a number of well-known mappings that allow the description of a fermionic system by the standard model of QC. They are the Jordan-Wigner transformation (Jordan and Wigner, 1928), Bravyi-Kitaev transformation (Bravyi and Kitaev, 2002) and Ball-Verstraete-Cirac transformation (Ball, 2005; Verstraete and Cirac, 2005). More advanced mappings using the interaction graph of the Hamiltonian (Setia and Whitfield, 2018) or customized quasi-local encodings (Chien and Whitfield, 2020) have been introduced as well. For the concreteness, the Jordan-Wigner transformation is briefly outlined here:

$$\hat{f}_j \rightarrow \left( \prod_{l=1}^{j-1} -\hat{\sigma}_z^l \right) \hat{\sigma}_-^j; \hat{f}_j^\dagger \rightarrow \left( \prod_{l=1}^{j-1} -\hat{\sigma}_z^l \right) \hat{\sigma}_+^j, \quad (87)$$

with  $\hat{\sigma}_\pm^j = (\hat{\sigma}_x^j \pm i\hat{\sigma}_y^j)/2$ . In this new transformation, one can verify that  $\hat{f}_j^\dagger, \hat{f}_j$  satisfy the above anticommutation relations, while  $\hat{\sigma}_\mu^j$  satisfy the commutation relations showed above. The reader is referred to the literature (Aspuru-Guzik et al., 2005; Seeley et al., 2012; Somma et al., 2003; Tranter et al., 2018, 2015) and the respective original references for details and comparisons regarding the other transformations.

**Bosons.** Bosonic operators satisfy the commutation relations  $[\hat{b}_i, \hat{b}_j] = 0, [\hat{b}_i, \hat{b}_j^\dagger] = \delta_{ij}$  in an infinite-dimensional Hilbert space. At first, it seems it is impossible to simulate bosonic systems due to the nature of infinite dimensions. However, sometimes we are interested in studying some finite modes of excitations above the ground state. Hence, the use of the entire infinite dimensional Hilbert space is unnecessary. In a finite dimensional basis, the bosons  $\hat{b}_i^\dagger, \hat{b}_i$  obey the following commutation relations (Batista and Ortiz, 2004)

$$[\hat{b}_i, \hat{b}_j] = 0; [\hat{b}_i, \hat{b}_j^\dagger] = \delta_{ij} \left[ 1 - \frac{N_b + 1}{N_b!} (\hat{b}_i^\dagger)^{N_b} (\hat{b}_i)^{N_b} \right], \quad (88)$$

with  $\hat{b}_i^\dagger \hat{b}_i |n_i\rangle = n_i |n_i\rangle$  with  $n_i = 0, \dots, N_b$ , where  $N_b$  is the maximum truncated excitation number, corresponding to the  $i^{\text{th}}$  bosonic site/mode. A direct consequence is one can then write down the creation and annihilation

operators as

$$\hat{b}_i^\dagger = \sum_{n=0}^{N_b-1} \sqrt{n+1} |n+1\rangle \langle n|, \quad (89)$$

and  $\hat{b}_i$  is complex conjugate of  $\hat{b}_i^\dagger$ . There are infinite means to translate such truncated operators into the QC language: Pauli words. Commonly used one is known as standard binary or compact encoding (McArdle *et al.*, 2019b; Sawaya and Huh, 2019; Sawaya *et al.*, 2020; Somma *et al.*, 2003; Veis *et al.*, 2016), where  $\{\alpha, \beta \in \mathbb{W}\}$  in  $|\alpha\rangle\langle\beta|$  are now written in terms of binary strings. Using the following identities:  $|0\rangle\langle 1| \equiv \hat{\sigma}_-$ ;  $|1\rangle\langle 0| \equiv \hat{\sigma}_+$ ;  $|0\rangle\langle 0| \equiv (\hat{1} + \hat{\sigma}_z)/2$ ;  $|1\rangle\langle 1| \equiv (\hat{1} - \hat{\sigma}_z)/2$ , Pauli words translation can be accomplished. Recently, detailed studies on various encodings: binary, Gray, Unary, block Unary, have been studied and Gray code in particular is found to be resource efficient (number of qubits and two-qubit entangling gates) in simulating some specific bosonic and spin Hamiltonians (Sawaya *et al.*, 2020).

*Anyons.* As seen above, we can now proceed to simulate more general particle statistics, in particular, hard-core anyons. With “hard-core”, we refer to the Pauli’s exclusion principle where only zero or one particle can occupy a single mode. The anyonic operators  $\hat{a}_i, \hat{a}_i^\dagger$  obey the following commutation relations:

$$\begin{aligned} [\hat{a}_i, \hat{a}_j]_\theta &= [\hat{a}_i^\dagger, \hat{a}_j^\dagger]_\theta = 0; [\hat{a}_i, \hat{a}_j^\dagger]_{-\theta} = \delta_{ij}(1 - (e^{-i\theta+1})\hat{n}_j); \\ [\hat{n}_i, \hat{a}_j^\dagger] &= \delta_{ij}\hat{a}_j^\dagger, \end{aligned} \quad (90)$$

where  $\hat{n}_j = \hat{a}_j^\dagger \hat{a}_j$ ,  $[\hat{A}, \hat{B}]_\theta = \hat{A}\hat{B} - e^{i\theta}\hat{B}\hat{A}$ , with  $(i \leq j)$  and  $0 \leq \theta < 2\pi$ . Specifically,  $\theta = \pi \bmod(2\pi)$  gives rise to canonical fermions, and  $\theta = 0 \bmod(2\pi)$  would recover hard-core bosons. By simply applying the following isomorphic mapping between algebras (Somma *et al.*, 2003):

$$\begin{aligned} \hat{a}_j^\dagger &= \prod_{i<j} \left( \frac{e^{-i\theta} + 1}{2} + \frac{e^{-i\theta} - 1}{2} \hat{\sigma}_z^i \right) \hat{\sigma}_+^j, \\ \hat{a}_j &= \prod_{i<j} \left( \frac{e^{i\theta} + 1}{2} + \frac{e^{i\theta} - 1}{2} \hat{\sigma}_z^i \right) \hat{\sigma}_-^j, \\ \hat{n}_j &= \frac{1 + \hat{\sigma}_z^j}{2}, \end{aligned} \quad (91)$$

we would obtain Pauli words for the QC. The above mapping would also ensure the anyonic algebra, Eq. (90).

## 2. Constructing electronic Hamiltonians

The electronic structure problem is one of the most prominent task within variational quantum algorithms (see for example the reviews (Cao *et al.*, 2019; McArdle *et al.*, 2020)) and was the pioneering task for the variational quantum eigensolver (McClean *et al.*, 2016;

Peruzzo *et al.*, 2014). In this section, we will illustrate how the original continuous many-electron problem can be discretized to a second-quantized formulation that can itself be encoded into qubits by the techniques introduced in the beginning of Sec. V.A. This encoded qubit systems define then the central problem of the VQAs further described in the proceeding sections Sec. V.A.3.

The electronic structure problem aims to approximate eigenfunctions of an electronic Hamiltonians

$$H_e = \mathcal{T}_e + \mathcal{V}_{ee} + \mathcal{V}_{\text{ext}}, \quad (92)$$

describing a system of  $N_e$  electrons through their accumulated kinetic energies  $\mathcal{T}_e = -\frac{1}{2} \sum_{k=1}^{N_e} \Delta_{\mathbf{r}_k}$ , the electronic Coulomb repulsion  $\mathcal{V}_{ee} = \sum_{k \neq l} V_{ee}(\mathbf{r}_k - \mathbf{r}_l) = \sum_{k \neq l} \frac{1}{|\mathbf{r}_k - \mathbf{r}_l|}$ , and an external potential  $\mathcal{V}_{\text{ext}} = \sum_{k=1}^{N_e} V_{\text{ext}}(\mathbf{r}_k)$  that is usually given by the accumulated Coulomb potential of nuclear point charges. If the external potential is not explicitly spin dependent, the electronic Hamiltonian only acts on the spatial coordinates  $\mathbf{r}_k \in \mathbb{R}^3$  of the electrons and in order to ensure proper electronic wave functions, the fermionic anti-symmetry is ensured via restrictions in the Hilbert-space. We refer to (Herbst, 2018; Kottmann, 2018; Rohwedder, 2010) and the textbook (Yserentant, 2010) for the direct construction and discretization of this continuous Hilbert spaces. A more compact, but formally equivalent, definition is offered through second quantization by introducing the abstract anti-commuting field operators  $\hat{\psi}^\dagger(x)$  and  $\hat{\psi}(x)$  that create and annihilate electrons at spin-coordinate  $x_k \in \mathbb{R}^3 \times \{\pm \frac{1}{2}\}$  (Jordan and Klein, 1927; Jørgensen, 2012; Surján, 2012). The electronic Hamiltonian can then be written as

$$\begin{aligned} H_e &= \int d\mathbf{x} \hat{\psi}^\dagger(\mathbf{x}) (T(\mathbf{x}) + V_{\text{ext}}(\mathbf{x})) \hat{\psi}(\mathbf{x}) \\ &+ \int d\mathbf{x} d\mathbf{y} \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{y}) V_{ee}(\mathbf{x} - \mathbf{y}) \hat{\psi}(\mathbf{y}) \hat{\psi}(\mathbf{x}) \end{aligned} \quad (93)$$

where the potential operators still only act on the spatial part of the spin components. Although direct approaches on real-space grids are possible (Kivlichan *et al.*, 2017; Kottmann, 2018; Kottmann and Bischoff, 2017; Kunita and Hirata, 2020) the majority of variational quantum algorithms employs a fixed set of three dimensional functions (so called orbitals) to capture the spatial part of the electronic Hilbert space. The orbitals are usually determined by solving a mean-field problem (Hartree-Fock) within a set of globally defined atomic orbitals. Alternatives to the standard representation are for example direct determination of system adapted orbitals (Kottmann *et al.*, 2021), compactification of basis sets through intrinsic atomic orbitals (Barison *et al.*, 2020) and optimized virtual orbitals represented by plane-waves (Bylaska *et al.*, 2020). For the formal description of the discretized second-quantized electronic Hamiltonian, the origin of the orbitals is not important as long as they form an orthonormal set of  $H^1(\mathbb{R}^3)$  functions. Using such a



set of spatial orbitals we can formally expand the field operators in the corresponding spin-orbitals

$$\hat{\psi}^\dagger(x) = \sum_k \phi_k^*(x) f_k^\dagger \quad (94)$$

$$\hat{\psi}(x) = \sum_k \phi_k(x) f_k \quad (95)$$

where  $f_k^\dagger$  and  $f_k$  are fermionic creation and annihilation operators obeying the anticommutation relations shown in the previous subsection. Using this expansion in Eq. (93) leads to the common discretized second-quantized Hamiltonian

$$H_e = \sum_{kl} h_{kl} f_k^\dagger f_l + \sum_{klmn} g_{klmn} f_k^\dagger f_l^\dagger f_n f_m \quad (96)$$

with the molecular integrals (Fermann and Valeev, 2020)

$$h_{kl} = \int \phi_k^*(x) (T(x) + V_{\text{ext}}(x)) \phi_l(x) dx \quad (97)$$

$$g_{klmn} = \int \phi_k^*(x) \phi_l^*(y) V_{\text{ee}}(x-y) \phi_m^*(x) \phi_n^*(y) dx dy.$$

Note that the indices of the two body integrals are denoted in the standard Dirac notation  $g_{klmn} \equiv \langle kl | V_{\text{ee}} | mn \rangle$  but other notations, such as Mulliken  $(km|lm) = \langle kl | V_{\text{ee}} | mn \rangle$  are sometimes used. Generally speaking, an arbitrary set of spatial orbitals, that can in principle be any set of orthonormal  $H^1(\mathbb{R}^3)$  functions, defines a discretized second-quantized Hamiltonian as in Eq. (96) over the corresponding molecular integrals Eq. (97). This discretized Hamiltonian can then be encoded into a qubit Hamiltonian by corresponding fermion to qubit mappings discussed in Sec. V.A.

### 3. Variational quantum eigensolver

Estimating the ground state and its energy of Hamiltonians is an important problem in physics, which has numerous applications ranging from solid-state physics to combinatorial optimization (see Sec. V.C). While this problem is in general QMA-hard and even quantum computers are not expected to be able to efficiently solve it in general (Kempe et al., 2006), there is hope that approximate solutions of the ground state could be found faster and for larger system sizes compared to what is possible with classical computers.

To this end, VQE (McClean et al., 2016; Peruzzo et al., 2014) has been proposed, to find the ground state of a Hamiltonian  $H$  in a manner that is suited for NISQ devices (Wecker et al., 2015). Following the concept introduced in Sec. II.A and Sec. II.B, a parameterized circuit  $U(\theta)$  is minimized with respect to the objective function, which in general is the expectation value of the energy of the Hamiltonian

$$\langle H_\theta \rangle \equiv \langle 0 | U^\dagger(\theta) H U(\theta) | 0 \rangle. \quad (98)$$

The approximated ground state is given by the quantum state  $|\psi_{\min}\rangle = U(\theta_{\min})|0\rangle$  which minimizes the energy  $\min_{\theta} \langle H_\theta \rangle \geq E_g$  upper bounded by the true ground state energy  $E_g$  as guaranteed by the Rayleigh-Ritz variational principle (Gould, 2012). VQE has been intensively studied in both theory and experiments, and various adaptations and extensions have been proposed, which we discuss in the following paragraphs.

*Self-verification.* Whether the variational quantum simulator has converged to an actual eigenstate of the Hamiltonian can be checked directly on the quantum processor by verifying that the variance of the energy  $\text{var} = \langle (H - \langle H \rangle)^2 \rangle$  is zero. This has been demonstrated for solving a many-body Hamiltonian on 8 qubits on a ion-trap (Kokail et al., 2019) (see also Sec. V.F.6).

*Accelerated VQE.* A key computational effort in VQE lies in estimating the cost function, which is achieved by repeatedly running the circuit and taking measurements of the Pauli strings (see Sec. II.C). For a given desired additive error bounded by  $\epsilon$ , it takes  $O(1/\epsilon^2)$  number of samples. This can be improved by using the Quantum Phase estimation algorithm to estimate the expectation value, which takes only  $O(\log(1/\epsilon))$  samples, however at the cost of additional computation which may be hard in the NISQ era. To leverage a trade-off between the advantages and disadvantages of both methods, an accelerated versions of VQE that interpolates between regular measurements and quantum phase estimation has been proposed (Wang et al., 2019).

*Measurement-based VQE.* In (Ferguson et al., 2020), the authors present two strategies to implement the VQE algorithm on a measurement-based quantum computer, an alternative quantum computing paradigm that uses entanglement as a resource and achieves the the desired computation by performing particular sets of local measurements (see (Briegel et al., 2009) for a review). They propose to a way to generate the needed variational state families using measurements on a highly entangled state and provide an equivalence between the measurement- and gate-based schemes.

*Reusing qubits in VQE.* A recent proposal suggested a VQE method that relies on fewer qubits by re-using qubits (Liu et al., 2019). The core idea is to represent a virtual  $N$  qubit state by  $R + V < N$  physical qubits, where  $R$  qubits have to be reusable qubits, e.g. they can be measured and re-initialized during the circuit runtime. These intermittent measurements are possible on current ion trap hardware (Pino et al., 2020). The  $R + V$  qubits

are entangled by a PQC, then  $R$  qubits are measured and the outcome is recorded. The  $R$  qubits are re-initialized to the  $|0\rangle^{\otimes R}$  state, and again entangled with the  $V$  other qubits by another PQC. This procedure is repeated until in total  $N$  qubits have been measured. The concept and expressiveness of this type of ansatz is the same as Tensor networks methods such as MPS, which have been highly useful for the classical calculation of many-body problems, and open up a way to perform quantum computing of many qubits on devices with limited number of qubits.

*Adiabatically assisted VQE.* The ground state of more challenging Hamiltonians can be difficult to find for standard VQE due to convergence to local minima instead of the global minima of the energy. To alleviate this, quantum annealing (see Sec. III.A) can be used to adiabatically assist the optimisation procedure, as proposed in the adiabatically assisted VQE (Garcia-Saez and Latorre, 2018). This approach uses an objective function  $O(s) = \langle 0|U^\dagger(\theta)H(s)U(\theta)|0\rangle$ , where  $H(s) = (1-s)H_0 + sH_1$ . Here,  $H_0$  is a Hamiltonian with easily preparable ground-state and the goal is to find the ground state of a Hamiltonian  $H_1$ . In this algorithm, VQE is run for multiple discrete steps  $s_n$ . One starts with  $s_0 = 0$  and finds the minimal parameters  $\theta_0^*$  of the objective function  $O(s_0)$ . Then,  $\theta_0^*$  is used as initial guess for VQE for the next increasing step  $s_1 = s_0 + \Delta s$  with objective function  $O(s_1)$ . This procedure is repeated until  $s = 1$  is reached. This approach eases the optimization task, as the initial Hamiltonian  $H_0$  is a simple Hamiltonian with a ground state that can be easily found via optimization. For small steps  $\Delta s$ , the ground state of the Hamiltonian  $H(s)$  and  $H(s + \Delta s)$  will not differ too much, making the optimization task at every step less challenging compared to directly solving for  $H(1)$ . Previous works (McClean et al., 2016) also suggest to use adiabatically prepared states as initial states of a VQE algorithm (see Sec. II.B).

#### 4. Variational quantum eigensolver for excited states

The methods of VQE have been extended to solve for the excited states of a given Hamiltonian. Finding excited states or the spectrum of Hamiltonian is an important problem in quantum chemistry and many-body physics. Various proposals have been put forward.

*Folded spectrum method.* A straightforward way of calculating excited states is the folded spectrum method proposed by Ref. (Peruzzo et al., 2014). To find an excited state of an Hamiltonian  $H$  with approximate energy  $\lambda$ , above defined VQE method is here applied to the objective function  $C(\theta) = \langle (H - \lambda)^2 \rangle_{\mathcal{U}(\theta)}$ . VQE will target

the eigenstate with an energy that is closest to  $\lambda$ . This method requires an approximate knowledge of the energy of the excited state that one wants to find, as well as estimating  $\langle H^2 \rangle$ , which may require a excessively large number of measurements to be performed.

An extension of this method can also be used to find states that are constrained to a specific value of the conserved quantity of the problem, such as total particle number, magnetization (Ryabinkin et al., 2018a). Here, one defines the objective function  $C(\theta) = \langle H \rangle_{\mathcal{U}(\theta)} + \sum_i \mu_i (\langle S_i \rangle_{\mathcal{U}(\theta)} - s_i)^2$ , where  $S_i$  is the operator corresponding to the conserved quantity, and  $s_i$  is the target value of the conserved quantity. Note, that this does not restrict the target space to be an eigenstate of  $S_i$ .

*Orthogonally constrained VQE.* Excited states can be found by constraining the VQE objective function such that it penalizes the ground state (Higgott et al., 2019). First, one finds an approximation to the ground state of Hamiltonian  $H$  via VQE with  $\theta_0 = \arg \min_{\theta} \langle H \rangle_{\mathcal{U}(\theta)}$  and approximated ground state  $|\psi(\theta_0)\rangle = \mathcal{U}(\theta_0)|0\rangle$ . Now, one uses this information to formally project out the approximate ground state, in order to find the next highest excited state. One defines the Hamiltonian  $H_1 = H + a|\psi(\theta_0)\rangle\langle\psi(\theta_0)|$ , with some sufficiently large positive parameter  $a$ . The ground state of  $H_1$  then corresponds to the first excited state of  $H$  and can be found with a VQE. This procedure can be repeated to find higher excited states up to any order by sequentially accumulating the projector terms of all states found. The Hamiltonian for the  $k$ -th excited state is then given by  $H_k = H + \sum_{i=1}^{k-1} a_i |\psi(\theta_i)\rangle\langle\psi(\theta_i)|$ . Combined with the unitary coupled cluster ansatz, the orthogonally constrained VQE can find excited states of small molecules (Higgott et al., 2019; Lee et al., 2018). It was further extended for adaptive circuit construction (Kottmann et al., 2020b) and imaginary time evolution (Jones et al., 2019b).

The projector term requires calculating the overlap  $|\langle\psi(\theta)|\psi(\theta_0)\rangle|$ , which can be achieved for example by the SWAP test, by applying the inverse of the circuit that generated the ground state  $|\langle 0|U^\dagger(\theta)U(\theta_0)|0\rangle|^2$ , or randomized measurements (Elben et al., 2020). An alternative approach that relies on a discriminator circuit that is trained in parallel to distinguish between the excited state to be learned and previously found lower-lying states has been proposed (Tilly et al., 2020) and demonstrated on a small model system. Scalable proposals still remain an open research question. Since the projector term does not require the overlap itself, but the absolute square of it, it can be computed with the help of Eq. (6) by computing the fidelity of the current trial state with the previously found states (Kottmann et al., 2021; Lee et al., 2018).

*Subspace expansion.* The subspace expansion method was introduced in Sec. IV.A.3 for error mitigation. This method can be used at the same time to find excited states, which has been proposed in (McClean *et al.*, 2017) and demonstrated for a small molecule in (Colless *et al.*, 2018). After finding the ground state of a Hamiltonian  $H$  with VQE, one follows the steps that were detailed in Sec. IV.A.3. One expands the prepared quantum state with different appropriate operators that match the low-energy excitations of  $H$  and generates a set of states that span the low-energy subspace. Then, overlaps between the states are measured, which are then used to solve a generalized eigenvalue problem on a classical computer. The eigenvalues and eigenstates give the excited states of the Hamiltonian. For quantum chemistry problems, the subspace expansion method was also proposed for including dynamical correlations to ground states over external corrections (Takeshita *et al.*, 2020), in the spirit of classical quantum chemistry methods, like for example CAS-CI (Roos *et al.*, 1980).

As alternative approach, the expansion in the subspace can also be accomplished by real-time evolving a reference state, and picking states at different evolution times as basis for expansion (Stair *et al.*, 2020). This is motivated by the fact that the time evolution can be seen as an approximate Krylov expansion of the quantum state. Then, one proceeds to solve the generalized eigenvalue problem to find eigenstates and eigenvalues of the Hamiltonian.

*Subspace-search VQE.* The core idea of subspace-search VQE (SSVQE) is to minimize the energy of a PQC  $U(\theta)$  over a set of orthogonal quantum states (Nakanishi *et al.*, 2019). The goal is to find the  $k$ -th eigenstates with lowest eigenenergy of a Hamiltonian  $H$ . In the weighted SSVQE the cost function is

$$L(\theta) = \sum_{j=1}^k w_j \langle \varphi_j | U^\dagger(\theta) H U(\theta) | \varphi_j \rangle, \quad (99)$$

where  $\{|\varphi_j\rangle\}_{j=0}^k$  is a set of  $k$  easily preparable mutually orthogonal quantum states (with  $\langle \varphi_i | \varphi_j \rangle = \delta_{i,j}$ ) and  $\{w_j\}_j$  are positive real numbers with  $w_i > w_j$  for  $i < j$ . Minimizing  $\theta^* = \arg \min_{\theta} L(\theta)$  to its global minimum gives us the ground state and excited states  $|\psi_j\rangle = U(\theta)|\varphi_j\rangle$ , where  $j = 1$  is the ground state and  $j > 1$  the excited states sorted in ascending order. This algorithm gives all  $k$  eigenstates in a single optimization routine. Note however that the more states are to be optimized, the more complex the optimization landscape and the effort to minimize it can become. An alternative formulation of the algorithm to find specifically the  $k$ -th lowest eigenstate is the unweighted SSVQE. Here, one

minimizes

$$L_1(\theta) = \sum_{j=1}^k \langle \varphi_j | U^\dagger(\theta) H U(\theta) | \varphi_j \rangle. \quad (100)$$

However, due to the absence of weights, the found states  $|\psi'_j\rangle = U(\theta^*)|\varphi_j\rangle$  for minimal  $\theta^*$  are not proper eigenstates of  $H$ , but are superposition states that span the subspace of the  $k$  lowest eigenenergies. As final step to find the  $k$ -th eigenstate, one fixes  $\theta = \theta^*$  to its minimized value, and then maximizes  $\phi^* = \max_{\phi} L_2(\phi)$ , with  $L_2(\phi) = \sum_{j=1}^k \langle \varphi_j | V^\dagger(\phi) U^\dagger(\theta^*) H U(\theta^*) V(\phi) | \varphi_j \rangle$  and  $V(\phi)$  being a unitary that acts only on the Hilbertspace of the  $k$  lowest eigenstates. Then, for the maximized  $\phi^*$ , the  $k$ -th lowest eigenstate is given by  $|\psi_k\rangle = U(\theta^*)V(\phi^*)|\varphi_k\rangle$ .

*Multistate contracted VQE.* The multistate contracted VQE combines the non-weighted SSVQE with the subspace expansion to find the ground state and excited states (Parrish *et al.*, 2019a). First, one runs the non-weighted SSVQE routine to find the unitary  $U(\theta^*)$  to find  $k$  states that span the subspace of the  $k$  smallest eigenvalues  $|\psi'_j\rangle = U(\theta^*)|\varphi_j\rangle$ . Then, to find the correct eigenstates, one runs the subspace expansion and measures the overlap matrix  $H_{ij} = \langle \psi'_i | H | \psi'_j \rangle$ , and diagonalizes it to find estimates of the  $k$  lowest eigenenergies and eigenstates.

*Fourier transform of evolution.* Recent experiments have determined the spectra of molecular and many-body Hamiltonians using superconducting processors (Aleiner *et al.*, 2020; Quantum *et al.*, 2020; Roushan *et al.*, 2017). A particular method to determine the eigenenergies of Hamiltonians via Fourier transforming the dynamics of observables has been applied here (Aleiner *et al.*, 2020; Roushan *et al.*, 2017). The idea is to prepare a Fock state that has overlap with the eigenstates whose eigenvalues one wants to calculate. The Fock state is then evolved in time with the Hamiltonian and specific observables are measured over a range of time. The Fourier transform of the time evolution of the observables can be used to deduce the eigenenergies of the Hamiltonian.

*Witness-assisted variational eigenspectra solver (WAVES).* WAVES core idea is to use a single reference qubit as an eigenstate witness to variationally find the ground state and excited states (Santagati *et al.*, 2018). A variational ansatz applied to a reference state is chosen. Then, the time evolution operator  $U(t) = \exp(-iHt)$  is evolved on the ansatz state as a control unitary  $CU(t)$ , with the control being the single qubit in a superposition state. Then, full tomography is performed on the single qubit to read out its von-Neumann entropy. If the variational

state is an eigenstate of the Hamiltonian  $H$ , then the entropy is zero. Further, the energy of the state can be estimated from the state of the qubit as well. The ansatz is variationally updated using the information from the qubit in an iterative fashion until the ground state is found. Excited states can be found by applying an appropriate excitation operator on the found ground state, and then variationally minimizing the von-Neumann entropy of the qubit. As last step, the authors suggest to use the iterative phase estimation algorithm to further improve the accuracy of the excited state as well determine its eigenvalue. This method requires to implement a controlled time evolution operator, similar to non-variational proposals (Jensen *et al.*, 2020), which are considered to be challenging for larger systems on NISQ devices.

## 5. Hamiltonian simulation

A major application for quantum computers is the simulation of the dynamics of Hamiltonians for problems such as many-body physics and chemistry. One standard approach for quantum simulation of Hamiltonians is based on the Trotter-Suzuki expansion from Eq. (17), where the evolving unitary is split up into small discrete timesteps of efficiently implementable unitaries, which can be run on the quantum computer. Naturally, the depth of the quantum circuit increases polynomially with the desired time to be evolved and target accuracy, which may not be feasible on NISQ devices without access to error correction. The relevant algorithms are reviewed in the following. We remark that some necessary tools to simulate many-body interaction Hamiltonian (Bravyi *et al.*, 2008; Menke *et al.*, 2019) have also been proposed.

*Variational quantum simulator.* Variational type algorithms have been proposed to solve dynamical problems in the NISQ era (Li and Benjamin, 2017). The core idea is to iteratively update an efficiently implementable variational quantum state  $|\psi(\theta)\rangle$  with a new set of parameters  $\theta \rightarrow \theta'$  such that it minimize the error between the actual time evolution  $\exp(-iH\delta t)|\psi(\theta)\rangle$  for a timestep  $\delta t$  and the updated variational state  $|\psi(\theta')\rangle$ . The rules to update the parameters  $\theta$  to solve the Schrödinger equation  $i d/dt |\psi(t)\rangle = H |\psi(t)\rangle$  can be found by the variational McLachlan's principle  $\delta \|(d/dt + iH)|\psi(\theta)\rangle\| = 0$  with  $\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle}$  and demanding that  $\theta$  remains real-valued. One finds a set of linear equations of motion  $A\dot{\theta} = C$  with

$$\begin{aligned} A_{i,j} &= \text{Re}(\partial_{\theta_i} \langle \psi(\theta) | \partial_{\theta_j} |\psi(\theta)\rangle) \\ C_i &= \text{Im}(\partial_{\theta_i} \langle \psi(\theta) | H |\psi(\theta)\rangle). \end{aligned}$$

At a given step of the iteration, one needs to measure the elements of  $A$  and  $C$  using the Hadamard test or

methods from (Mitarai and Fujii, 2019) (see Sec. II.C), and then update  $\theta$  with the solution of the linear equation of motion by a small timestep  $\delta t$ . The solver can be combined with adaptive strategies to reduce the complexity of the Ansatz circuit (Yao *et al.*, 2020c; Zhang *et al.*, 2020c). VQS has been applied on the IBM quantum processor to simulate energy transfer in molecules (Lee *et al.*, 2021) as well as to simulate a time-dependent Hamiltonian (Lau *et al.*, 2021). A straightforward extension of the variational quantum simulator can be applied to solve the Schrödinger equation in imaginary time (McArdle *et al.*, 2019a), for time-dependent problems (Yuan *et al.*, 2019) or for general linear differential equations (Endo *et al.*, 2020c; Kubo *et al.*, 2020). The variational quantum simulator can be extended to open quantum systems (Endo *et al.*, 2020c; Yuan *et al.*, 2019), which we discuss in Sec. V.A.7.

*Subspace variational quantum simulator.* The subspace variational quantum simulator (SVQS) (Heya *et al.*, 2019) builds upon the idea of the SSVQE (Nakanishi *et al.*, 2019) introduced earlier in Sec. V.A.4. The core idea is to rotate the initial state to be evolved onto the low-energy subspace found by the weighted SSVQE, then evolve it in time within the subspace, and then apply the reverse mapping. First, run the weighted SSVQE by preparing  $k$  initial states  $\{|\varphi_j\rangle = \sigma_j^x |0\rangle\}_{j=0}^k$  which are orthogonal with each other ( $\langle \varphi_i | \varphi_j \rangle = \delta_{i,j}$ ) and lie in the computational subspace, as well as a PQC  $U(\theta)$ . Now as in the weighted SSVQE minimize Eq. (99).

Then, prepare an initial state  $|\psi_{\text{in}}\rangle$  to be evolved, which is encoded into the computational subspace by applying the Hermitian conjugate of the obtained circuit  $U^\dagger(\theta)$ . Here, the evolution of the state in time is performed by applying single-qubit rotations on each qubit  $\mathcal{T}(t) = \bigotimes_j R_Z(-E_j t)$ , where  $\{E_j\}_j$  are the eigenenergies of the eigenstates  $\{|E_j\rangle\}_j$  obtained by SSVQE earlier. Finally, the state  $\mathcal{T}(t)U^\dagger(\theta)|\psi_{\text{in}}\rangle$  in the computational subspace is reverse mapped by applying  $U(\theta)$ , giving the evolved state

$$|\psi(t)\rangle = U(\theta)\mathcal{T}(t)U^\dagger(\theta)|\psi_{\text{in}}\rangle. \quad (101)$$

This method has the key advantage that since the evolution is directly implemented as simple rotations in the computational subspace, the circuit depth is independent of the evolution time to be simulated. However, the initial SSVQE optimization can be difficult, especially when one considers many eigenstates  $k$ .

*Variational fast forwarding.* Similar to the idea of the SVQS, variational fast forwarding (VFF) relies on the idea of evolving a quantum state in time  $\exp(-iHt)$  within a diagonal subspace, such that an enhanced evolution time can be achieved (Cirstoiu *et al.*, 2020). First,



a circuit that implements a small timestep of the desired evolution is implemented as  $V(\delta t) = \exp(-iH\delta t)$ . Then, an approximate diagonal factorization of  $V(\delta t)$  is trained for a particularly structured variational circuit

$$U(\boldsymbol{\theta}, \boldsymbol{\gamma}, \delta t) = W(\boldsymbol{\theta})D(\boldsymbol{\gamma}, \delta t)W^\dagger(\boldsymbol{\theta}). \quad (102)$$

Here,  $D(\boldsymbol{\gamma}, \delta t)$  is composed of commuting unitaries and chosen to parameterize the eigenvalues of unitary  $V(\delta t)$ , whereas  $W(\boldsymbol{\theta})$  represents its eigenvectors. Then, the evolution to an arbitrary time  $T = N\delta t$ , where  $N$  is some integer, is found by fast forwarding with  $U(\boldsymbol{\theta}, \boldsymbol{\gamma}, N\delta t) = W(\boldsymbol{\theta})D^N(\boldsymbol{\gamma}, \delta t)W^\dagger(\boldsymbol{\theta})$ . For the training of the variational Ansatz, the fidelity between  $V(\delta t)$  and  $U(\boldsymbol{\theta}, \boldsymbol{\gamma}, \delta t)$  is maximized by a quantum-classical feedback loop with a cost function that uses the local Hilbert-Schmidt test (Khatri *et al.*, 2019). As alternative approach, it was proposed to diagonalize the Hamiltonian  $H$  instead of the unitary  $V(\delta t)$ , and fast forward via  $U(\boldsymbol{\theta}, \boldsymbol{\gamma}, T) = W(\boldsymbol{\theta})\exp(-iD(\boldsymbol{\gamma})T)W^\dagger(\boldsymbol{\theta})$  (Commeau *et al.*, 2020).

*Quantum Assisted Simulator.* The VQS algorithm employs a classical-quantum feedback loop to update the parameters of the PQC. Until the classical processor has calculated its output, the classical-quantum feedback loop thwarts any use of the quantum device, slowing the algorithm on current cloud computing framework. The VQS algorithm, as well as its VQE based variant, i.e. SVQS share similarities and most of the concerns faced by VQE, such as the barren plateau issue (see Sec. IV.B) as well as lacking a systematic approach to select the ansatz. Further, the VQS algorithm requires controlled-unitaries, which make it difficult to realize for current-term devices.

To tackle the issues faced by VQS, the quantum assisted simulator (QAS) was suggested recently (Bharti and Haug, 2020b). The QAS algorithm does not need any classical-quantum feedback loop, can be parallelized, evades the barren plateau problem by construction, supplies a systematic approach to constructing the Ansatz and does not require any complicated unitaries.

The QAS algorithm shares its approach with IQAE (see Sec. III.E). The ansatz is given as linear combination of states  $|\phi(\boldsymbol{\alpha}(t))\rangle = \sum_{|\psi_i\rangle \in \mathbb{C}S_K} \alpha_i(t)|\psi_i\rangle$  (see Eq. (58)), with classical coefficients  $\boldsymbol{\alpha}(t)$  for ansatz state  $|\psi_i\rangle$ , which can be systematically constructed (see Definition 3). The Hamiltonian  $H$  is given as linear combination of unitaries (see Eq. (59)). The QAS algorithm employs Dirac-Frenkel principle to obtain the following classical evolution equation for  $\boldsymbol{\alpha}(t)$

$$\mathcal{E} \frac{\partial \boldsymbol{\alpha}(t)}{\partial t} = -i\mathcal{D}\boldsymbol{\alpha}(t). \quad (103)$$

Here,  $\mathcal{E}_{i,j} = \langle \psi_i | \psi_j \rangle$  and  $\mathcal{D}_{i,j} = \sum_k \beta_k \langle \psi_i | U_k | \psi_j \rangle$  are overlap matrices that can be efficiently measured on a quantum computer, i.e. for  $H$  given as combination of Pauli

strings, the overlaps are measurement of Pauli strings. Recently, QAS was run on the IBM quantum computer and showed superior performance compared to Trotter and VQS for a time-dependent Hamiltonian (Lau *et al.*, 2021).

## 6. Quantum information scrambling and thermalization

Quantum information scrambling is a quantum phenomena occurring when initially local states become increasingly non-local with the time-evolution of the system. It can be analyzed by computing the so-called out-of-time-ordered correlation function (OTOC) and have strong implications in thermalization in closed quantum systems dynamics. Recent experiments have been carried out to study this phenomena in a few qubits trapped-ion devices and simulators (Joshi *et al.*, 2020; Landsman *et al.*, 2019). The algorithms proposed are based on the well-known teleportation algorithm and use single and two-qubit gates to reproduce the the scrambling process.

In the context of VQAs, a variation of the VQE algorithm has been proposed to obtain the thermal evolution of quantum systems (Verdon *et al.*, 2019b). The authors present the Quantum Hamiltonian-Based Models (QHBM), an extension of the VQA's PQC to mixed states instead of pure states. Within this approach, the QHBM are classically trained to learn a mixed state distribution as a function of the optimization parameters. A direct application of such a model is the Variational Quantum Thermalizer (VQT), an algorithm which goal is to prepare a fixed-temperature thermal state of a given Hamiltonian.

The limitations of using variational QML algorithms to learn a scrambling unitary have also been studied in (Holmes *et al.*, 2020), where it is found trainability issues related with barren plateaus (see Sec. IV.B).

## 7. Simulating open quantum systems

In the following, we deal with the physics of open quantum systems (Huh *et al.*, 2014) which are well-described by the Lindblad master equation:  $d\hat{\rho}(t)/dt = -i[\hat{K}(t), \hat{\rho}(t)] + \hat{\mathcal{L}}[\hat{\rho}(t)]$  (see Sec. IV.A.1). By sampling from a mixture of pure state trajectories evolved by a non-Hermitian Hamiltonian and random quantum jumps, one recovers the Lindblad dynamics.

*Trotter simulation of open systems.* NISQ quantum hardware can be used to directly simulate the dynamics of small-scale open systems by using ancillas combined with measurements in the spirit of the quantum jump method (Hu *et al.*, 2020; Koppenhöfer *et al.*, 2020). Here, the unitary part of the dynamics is implemented via a

Suzuki-Trotter decomposition (see Sec. II.B). The non-unitary part of the dynamics that encodes the interaction with the external degrees of freedom is simulated by entangling the circuit with ancillas and subsequently measuring them. For every time step of the dynamics a new set of ancilla qubits has to be provided. Current quantum computers based on superconducting circuits do not allow to measure and re-use qubits, thus requiring a linear increase in the number qubits with every timestep. Further, in general the circuit depth scales polynomially with simulation time.

*Generalized variational quantum simulator.* In Ref. (Endo *et al.*, 2020c) the VQS algorithm is extended to simulate the method of quantum jumps in a variational setting. They implemented the algorithm for 2D Ising Hamiltonians for 6 qubits and observed a dissipation induced phase transition. In another work (Yuan *et al.*, 2019), VQS is extended to mixed states and simulate the Lindblad dynamics fully without the need of stochastic sampling. The idea is to write the density matrix as  $\rho = \rho(\theta(t))$  and simulate the evolution of  $\rho$  via evolution of the parameters  $\theta(t)$ . Equation Eq. (61) can be re-expressed as  $\frac{d}{dt}\rho = \sum_i g_i S_i \rho T_i^\dagger$ , where  $S_i$  and  $T_i$  are unitaries and  $g_i$  are coefficients. Using Dirac and Frenkel equation, the evolution of parameters is given by

$$\sum_j M_{i,j} \dot{\theta}_j = V_i, \quad (104)$$

$$M_{i,j} = \text{Tr} \left[ (\partial_i \rho(\theta(t)))^\dagger \partial_j \rho(\theta(t)) \right] \quad (105)$$

$$V_i = \text{Tr} \left[ (\partial_i \rho(\theta(t)))^\dagger \sum_j g_j S_j \rho T_j^\dagger \right]. \quad (106)$$

This method can also be extended to deep quantum neural network type Ansatzes (Liu *et al.*, 2020b).

These algorithms, however, suffer from the canonical drawbacks of the VQS algorithm, such as the requirement of feedback loop, trainability issues and necessity of controlled unitaries.

*Generalized quantum assisted simulators.* Recently, the generalized quantum assisted simulator (GQAS) (Haug and Bharti, 2020) was proposed as extension of the quantum assisted simulator to tackle above issues (see Sec. V.A.5. Instead of using a density matrix, the GQAS algorithm introduced the concept of “hybrid density matrix”

$$\hat{\rho} = \sum_{k,l} \beta_{k,l} |\psi_k\rangle \langle \psi_l| \quad (107)$$

for  $\beta_{k,l} \in \mathbb{C}$  and  $|\psi_l\rangle$  are chosen from the set of cumulative  $K$  moment states (see Definition 3). A classical device stores the coefficients  $\beta$  and the quantum states

correspond to some quantum register. A hybrid density matrix is a valid density matrix, if  $\text{Tr}(\hat{\rho}) = 1$  and  $\hat{\rho} \geq 0$ . Note that the normalization condition is fulfilled when  $\text{Tr}(\hat{\rho}) = \text{Tr}(\beta \mathcal{E}) = 1$ , where  $\mathcal{E}_{k,l} = \langle \psi_k | \psi_l \rangle$ . Using Dirac-Frenkel principle, the simulation of open system dynamics for the hybrid density matrix is given by

$$\mathcal{E} \frac{d}{dt} \beta(t) \mathcal{E} = -\iota(\mathcal{D}\beta(t)\mathcal{E} - \mathcal{E}\beta(t)\mathcal{D}) + \sum_{n=1}^K \gamma_n (\mathcal{R}_n \beta(t) \mathcal{R}_n^\dagger - \frac{1}{2} \mathcal{F}_n \beta(t) \mathcal{E} - \frac{1}{2} \mathcal{E} \beta(t) \mathcal{F}_n), \quad (108)$$

where  $\mathcal{D}_{k,l} = \langle \psi_k | H | \psi_l \rangle$ ,  $\mathcal{R}_{k,l}^n = \langle \psi_k | L_n | \psi_l \rangle$  and  $\mathcal{F}_{k,l}^n = \langle \psi_k | L_n^\dagger L_n | \psi_l \rangle$ . For a given choice of ansatz, the quantum computers only have to compute the overlap matrices as measurements of Pauli strings. Then, the classical computer uses this information to simulate the dynamics. There is no quantum-classical feedback loop, which on the currently available quantum computers can speed up computations substantially.

## 8. Nonequilibrium steady state

Unlike the previous Sec. V.A.7, we concern the physics of open quantum system that is out-of-equilibrium in nature, which is common in designing devices for molecular-scale electronics (Xiang *et al.*, 2016) as well as quantum thermodynamics (Vinjanampathy and Anders, 2016). By “out-of-equilibrium”, we mean that a quantum system and bath/s are constantly driven by external forces such as voltage differences, during which the composite particles of the system and bath are also interacting each other.

Notice that the method used in the previous Sec. V.A.7 would also lead to extremely high dimensional matrices in the Lindblad like master equation approach  $d\hat{\rho}/dt = \hat{\mathcal{L}}\hat{\rho}$  (see Sec. IV.A.1, Eq. (61)), and it deems impossible to capture all the degrees-of-freedom involved. However, one may relax some of the constraints involved in the problem setup, say time-independent dissipation and non-interaction among particles with small system size. The steady state density matrix of a quantum system  $\hat{\rho}_{\text{SS}}$  at the limit  $t \rightarrow \infty$  is then given by solving

$$\hat{\mathcal{L}}|\hat{\rho}_{\text{SS}}\rangle = 0, \quad (109)$$

or equivalently  $\hat{\mathcal{L}}^\dagger \hat{\mathcal{L}}|\hat{\rho}_{\text{SS}}\rangle = 0$ . A recent study (Yoshioka *et al.*, 2020) has shown that with ancilla qubits, the above non-Hermitian superoperator  $\hat{\mathcal{L}}$  can be simulated. The main idea is to map the density matrix of  $N$  qubits onto a vector of twice the number of qubits  $2N$

$$\hat{\rho} = \sum_{ij} \rho_{ij} |i\rangle \langle j| \rightarrow |\hat{\rho}\rangle = \sum_{ij} \frac{\rho_{ij}}{C} |i\rangle_P |j\rangle_A, \quad (110)$$

where  $C = \sqrt{\sum_{ij} |\rho_{ij}|^2}$ . By using a digital quantum computer and the variational approach one iteratively minimize the expectation value of a parameterized density matrix  $|\rho_\theta\rangle = U(\theta)$  with  $\min_\theta \langle 0|^{\otimes 2N} U^\dagger(\theta) \hat{\mathcal{L}}^\dagger \hat{\mathcal{L}} U(\theta) |0\rangle^{\otimes 2N}$ . A drawback of this approach is that measuring expectation values from the parameterized density matrix directly is difficult and thus requires an additional transformation.

Beyond the Lindblad master equation, to capture and describe truly “out-of-equilibrium” processes, the nonequilibrium Green’s function (NEGF) formalism (Dalla Torre *et al.*, 2013; Sieberer *et al.*, 2016; Stefanucci and Van Leeuwen, 2013) is commonly used. These existing Green’s function techniques are very complicated to be solved. Many assumptions need to be made in order to have some closed form and do some calculations. In particular, it requires that the interaction among particles are weak such that one does not need to find higher-order Feynman diagrams in finding the self-energy functional.

Since some of the existing quantum algorithms provide promising speedup over classical ones, one may wonder to use quantum algorithms to solve the NEGF, with a strategy of leaving classically hard computational tasks to the quantum processor and feeding its output back to classical computer, which could be done in a variational fashion. There exists a number of proposals (Endo *et al.*, 2020b; Jaderberg *et al.*, 2020; Kreula *et al.*, 2016) in the literature that undertake such hybrid quantum-classical approach. However, these methods assume no interaction among composite particles. In generic open quantum system in which many-body effects cannot be neglected, one would like to go beyond those assumptions. It is yet to see any quantum advantage of those near-term quantum algorithms over existing methods (Fitzpatrick *et al.*, 2017; Härtle *et al.*, 2008; Li *et al.*, 2016) for solving nonequilibrium steady state solution of an extremely complex physical setup such as vibrationally-coupled electron transport with multiple electronic levels (Härtle *et al.*, 2008).

## 9. Gibbs state preparation

Finding the ground state of quantum Hamiltonians is known to be QMA-hard. Under reasonable assumptions, preparing Gibbs state corresponding to arbitrarily small temperatures is as challenging as the Hamiltonian ground state problem. Gibbs state preparation has applications in many areas including quantum annealing, quantum SDP solvers, Boltzmann training and simulation of equilibrium physics. For a Hamiltonian  $H$ , the Gibbs state at temperature  $T$  (with  $k_B = 1$ ) is given by

$$\rho(T) = \frac{\exp\left(-\frac{H}{T}\right)}{\text{Tr}\left(\exp\left(-\frac{H}{T}\right)\right)}. \quad (111)$$

Some of the approaches to prepare Gibbs state are mentioned in the following

- (1) Starting with  $d$ -dimensional maximally mixed state  $\frac{I_d}{d}$ , under imaginary time evolution for time  $\tau$ , one gets Gibbs state corresponding to temperature  $T = \frac{1}{2\tau}$  (Verstraete *et al.*, 2004).
- (2) One can start with maximally entangled state  $|\xi\rangle_d = \frac{1}{\sqrt{d}} \sum_j |j, j\rangle_{AB}$  of a system combined of two equally sized subsystems A and B, and evolve it under imaginary time evolution using Hamiltonian  $H \otimes I$ . After tracing out system B, the state of system A at time  $\tau$  is given by Gibbs state corresponding to temperature  $T = \frac{1}{2\tau}$ .
- (3) The Gibbs state of a system is the density matrix which corresponds to minimum of its free energy. Thus, one can variationally tune the parameters of a parametrized density matrix such that it leads to minimization of free energy.

Recently, a few NISQ algorithms for Gibbs state preparation has been proposed, which apply the aforementioned ideas. In (Endo *et al.*, 2020c), authors used VQS based imaginary time evolution to prepare Gibbs state following the second approach. The first approach does not work in VQS based imaginary time evolution. In another work (Chowdhury *et al.*, 2020), the third approach was used to prepare Gibbs states. The aforementioned works require complicated controlled unitaries and classical-quantum feedback loop. In (Haug and Bharti, 2020), QAS based imaginary time evolution (see Sec. V.A.5) was suggested to prepare the Gibbs state with either first and second approach. The QAS approach does not require any classical quantum feedback loop or complicated controlled unitaries.

## 10. Many-body ground state preparation

The preparation of non-trivial many-body quantum states is crucial for many applications in quantum information processing and quantum metrology. QAOA has been used as a resource-efficient scheme for many-body quantum state preparation. In this context, the state  $|\psi\rangle$  for a system with linear dimension  $L$  (e.g.  $L$  can refer to the number of spins in a 1D spin chain) is non-trivial if there is no local unitary circuit  $U$  with depth  $\mathcal{O}(1)$  which can generate  $|\psi\rangle$  from a product state  $|\phi\rangle$ :  $|\psi\rangle = U|\phi\rangle$  (Ho and Hsieh, 2019). The Greenberger-Horne-Zeilinger (GHZ) state, which is an essential resource in several quantum metrology proposals (Dür *et al.*, 2014; Tóth and Apellaniz, 2014), is an example of a non-trivial quantum state due to its highly-entangled nature, and is the ground state of the 1D Ising Hamiltonian with periodic boundary conditions, i.e.  $H_P = -\sum_{i=1}^L \hat{\sigma}_z^i \hat{\sigma}_z^{i+1}$

Using QAOA, it has been shown that the GHZ state can be prepared efficiently with perfect fidelity using  $p = L/2$ , where  $p$  is the QAOA depth (Ho and Hsieh, 2019).

The authors conjectured that the ground state of the 1D transverse-field Ising model

with  $L$  even and periodic boundary conditions, can be prepared perfectly at any point in the phase diagram using QAOA with  $p = L/2$ . The ground state of the antiferromagnetic Heisenberg model with open boundary conditions

$$H_P = \sum_{i=1}^{L-1} \hat{\sigma}^i \cdot \hat{\sigma}^{i+1} \quad (112)$$

where  $\hat{\sigma}^i \equiv (\hat{\sigma}_x^i, \hat{\sigma}_y^i, \hat{\sigma}_z^i)$  has also been prepared with near perfect fidelity using QAOA. Using a long-range 1D Ising Hamiltonian

$$H_P = - \sum_{i < j} J_{ij} \hat{\sigma}_z^i \hat{\sigma}_z^j \quad (113)$$

where  $J_{ij} = J_0/|i-j|^\alpha$ , QAOA can achieve the ultrafast preparation of a GHZ state with a circuit depth of  $\mathcal{O}(1)$  (for  $\alpha = 0$ ) (Ho *et al.*, 2019). This result was generalized by (Wauters *et al.*, 2020a), which showed that QAOA can prepare the ground states of the fully-connected ferromagnetic  $q$ -spin model (note that  $q$  is used here instead of the conventional  $p$  in order to avoid confusion with the QAOA depth  $p$ )

$$H = -\frac{1}{N^{q-1}} \left( \sum_{i=1}^N \hat{\sigma}_z^i \right)^q - h \left( \sum_{i=1}^N \hat{\sigma}_x^i \right) \quad (114)$$

with resources scaling polynomially with the number of spins  $N$ . Since the system can encounter a first-order phase transition where the spectral gap becomes very small, QAOA greatly outperforms quantum annealing in this instance since an exponentially long annealing time is needed.

## 11. Quantum autoencoder

The quantum autoencoder (Romero *et al.*, 2017) (QAE) is a variational algorithm for the compression of data on a quantum computer. It finds a new data state representation which requires fewer qubits than the data was originally defined upon. This new encoding is said to be a representation in the *latent space*. The process of transforming the data into the latent space is referred to as *encoding*, and the converse, transformation of states in the latent space back onto the original, is known as *decoding*.

Training a QAE requires the minimization of an objective defined over several related quantum states. For a set of  $n$ -qubit states  $\{|\psi\rangle_i\}$ , the goal of the QAE is to find a unitary circuit  $\mathcal{E}(\theta)$  which accomplishes the following transformation

$$E : \mathcal{H}^n \rightarrow \mathcal{H}^k \otimes \mathcal{H}^{n-k} |E|\psi\rangle_i = |\phi\rangle_i \otimes |0\rangle^{\otimes(n-k)}, \quad (115)$$

where  $k$  is the dimension of the latent space. Thus, the application of a perfectly trained autoencoder to any state of the relevant set yields a product state that consists of the transformed state on  $k$  qubits with a  $(n-k)$ -qubit “trash” state. In principle, the trash state could be any state, but the all-zero state is chosen for simplicity.

The loss function of the QAE may be defined in several ways. It is a fidelity loss function (see Sec. II.A), in which minimization is performed by increasing the overlap between a (partial) measurement of the state resulting from the application of the encoder and a known state. The most practical definition for training the autoencoder, called “trash training”, uses as its objective the overlap between the “trash” qubits and the  $|0\rangle^{\otimes(n-k)}$  state. Formulated in the density matrix picture, the objective of minimization is

$$O = -\text{Tr}(I^{\otimes k} \otimes |0\rangle\langle 0|^{\otimes(n-k)} \rho_i), \quad (116)$$

where  $\rho_i = \sum_i p_i |\psi_i\rangle\langle\psi_i|$  with, in general, all states in the set equally weighted.

The QAE can be trained by training only the encoding circuit, due to the unitarity of the encoder; the decoding operation is achieved by the complex conjugate of the encoder circuit. Improvements in the encoding results translates to improvements in the decoder, a boon not possessed by classical autoencoders.

After the successful training of a QAE, the encoder and decoder circuits may be used for data transformation in further algorithms, action upon the data in the latent space, in which the data is represented more densely, may prove powerful in further applications.

A data re-uploading strategy to construct a QAE encoder is presented in (Bravo-Prieto, 2020), where the so-called enhanced feature quantum autoencoder (EF-QAE) is trained to compress the ground state of the 1-D Ising model as a function of the external field and samples of handwritten digits. The QAE has also been deployed experimentally in the compression of qutrits on a photonic device (Pepper *et al.*, 2019). Small states have been experimentally compressed losslessly on photonic devices by (Huang *et al.*, 2020a). (Bondarenko and Feldmann, 2020) designed a QAE capable of denoising entangled quantum states, such as GHZ or W states, subject to spin flip errors and random unitary noise.

## 12. Quantum computer-aided design

Two recent proposals focus the computing power of NISQ devices back on the processors themselves: Techniques were developed to simulate quantum hardware on a quantum computer (Kottmann *et al.*, 2020c; Kyaw *et al.*, 2020b). They establish the paradigm of “quantum computer-aided design”, indicating that classically intractable simulations of quantum hardware properties



can be performed on a quantum computer, thereby improving the prediction of device performance and reducing experimental testing cycles.

In the first approach, optical path modes are mapped to sets of qubits, and quantum optical elements are mapped to digital quantum circuits that act on the qubits (Kottmann *et al.*, 2020c). Photonic setups can then be flexibly simulated. The framework is used to simulate both a Boson sampling experiment and the optimization of a setup to prepare a high-dimensional multipartite entangled state.

The second proposal introduces quantum simulation techniques for superconducting circuit hardware (Kyaw *et al.*, 2020b). A circuit module consisting of coupled transmon qubits is designed. The corresponding superconducting circuit Hamiltonian, which is written in a basis of multi-level operators, is efficiently mapped to a set of data qubits (Sawaya *et al.*, 2020). Simulations of a multi-level extension to the VQE algorithm (Higgott *et al.*, 2019) are used to determine the spectrum of the superconducting circuit. The resulting states and eigenenergies are directly related to experimentally relevant device characteristics and can be used to seed simulations of time dynamics.

Device and setup design is a key challenge in improving and scaling quantum systems. Therefore, digital quantum simulation of quantum processors will be a relevant application for NISQ quantum computers as classical resources become too small to capture the relevant Hilbert space of the hardware.

## B. Machine learning

The goal of machine learning is to facilitate a computer to act without being explicitly programmed to. As per Tom Mitchell (Mitchell, 1997), given some class of tasks  $\mathcal{T}$  and performance metric  $\mathcal{P}$ , a computer program is said to learn from experience  $\mathcal{E}$  if

$$\mathcal{P}(\mathcal{T}) \propto \mathcal{E}, \quad (117)$$

i.e. its performance measured by  $\mathcal{P}$  for task  $\mathcal{T}$  increase with  $\mathcal{E}$ .

Depending on the kind of experience  $\mathcal{E}$  permitted to have during the learning process, the machine learning algorithms are classified into three categories:

- (1) *Supervised learning.* Given a function  $y = f(x)$ , the goal is to learn  $f$  so that it returns the label  $y$  for the unlabelled data  $x$ . A canonical example would be pictures of cats and monkeys, with the task to recognize the correct animal. Given training examples from the joint distribution  $P(Y, X)$ , the task of supervised learning is to infer the probability of a label  $y$  given example data,  $x$ , i.e.,  $P(Y = y|X = x)$ .

- (2) *Unsupervised learning.* The data is provided without any label. The task is to recognize an underlying pattern in this data. Given access to several examples  $x \in X$  learning is the algorithm goal is to learn the probability distribution  $P(X)$  or some important properties of the aforementioned distribution.
- (3) *Reinforcement learning.* In this case, neither data nor label is provided. The machine has to generate data and improve the aforementioned data generation process via optimizing a given reward function. This is similar to how a human child learns to walk. If it falls, that acts as a negative reward.

Machine learning has uncovered applications in physics such as Monte Carlo Simulation (Huang and Wang, 2017; Liu *et al.*, 2017), many-body physics (Carleo and Troyer, 2017), phase transition (Wang, 2016), quantum foundations (Bharti *et al.*, 2019a), and state tomography (Torlai *et al.*, 2018). For a meticulous review on machine learning for physics, refer to (Benedetti *et al.*, 2019b; Bharti *et al.*, 2020a; Carleo *et al.*, 2019; Dunjko and Briegel, 2018).

Most of the success in machine learning come from the use of artificial neural networks, structures capable of learning sophisticated distributions and that encompass multiple features that can be fine-tuned depending on the problem to tackle. In that direction, there are several proposals to define a model for quantum neural networks with different kind of activation functions (Schuld *et al.*, 2014; Torrontegui and García-Ripoll, 2019; Wan *et al.*, 2017). However, in some of these cases, these models are not suited for the NISQ era.

The merger of quantum theory and machine learning has recently led birth to a new discipline, known as quantum machine learning (QML). Both algorithms that deal classically with data from a quantum origin and quantum algorithms that process quantum and classical data are usually known as QML applications. However, in this review, we will focus only on those algorithms that process data quantum-mechanically, in particular, those that use quantum algorithms that can be run in NISQ computers. For QML review that mainly focus on fault-tolerant quantum algorithms check (Biamonte *et al.*, 2017). For a survey of quantum computational learning theory, refer to (Arunachalam and de Wolf, 2017). An analysis of QML from a classical ML perspective can be found at (Ciliberto *et al.*, 2018; Dunjko and Briegel, 2018), and for near-term devices in (Benedetti *et al.*, 2019b; Perdomo-Ortiz *et al.*, 2018).

It might be surprising that a linear theory as quantum physics can generate the non-linearities that a machine learning model needs. However, the linearity of quantum mechanics comes from the dynamical part (quantum states evolution) and one can encounter multiple sources of non-linearities arising from measurement,

post-selection or coupling the system with environment. Quantum operations in the Hilbert space can also encode non-linear behaviour, as it will be shown with Kernel methods.

In the following subsections, we will present the quantum mechanical analogs of the three machine learning categories defined above. The algorithms discussed will be listed in Table II.

### 1. Supervised learning

The two prominent methods to perform a supervised learning classification task using a NISQ computer are *quantum Kernel estimation* (Havlíček *et al.*, 2019; Kusumoto *et al.*, 2019; Schuld and Killoran, 2019) and *Variational Quantum Classifier* (VQC) (Farhi and Neven, 2018; Mitarai *et al.*, 2018).

Classical Kernel methods include well-known machine learning algorithms such as Support Vector Machines (SVM) (Cortes and Vapnik, 1995), Principal Component Analysis (PCA) or Gaussian Processes, among others. The rich theoretical structure of Kernel methods can be expanded to the quantum world by defining and working in the Hilbert space with the quantum equivalent of feature vectors (Schuld and Killoran, 2019). To that aim, one needs to modify and adapt the well-known theorems to work in a quantum feature space. For more details about classical Kernel methods we refer to (Hofmann *et al.*, 2008). A review on Kernel methods in the context of Quantum Machine Learning can be found in (Mengoni and Di Pierro, 2019). In the following lines, we will directly describe the quantum versions of them. The basics of supervise learning with quantum computers are presented in (Schuld, 2018).

Given an input set  $\mathcal{X}$  and quantum Hilbert space  $\mathcal{H}$ , data  $\mathbf{x} \in \mathcal{X}$  is encoded into a quantum state (*quantum feature vector*)  $|\Phi(\mathbf{x})\rangle$  by means of the *quantum feature map*, i.e.  $\Phi: \mathcal{X} \rightarrow \mathcal{H}$ . The inner product of two quantum feature vectors defines a kernel

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) \equiv \langle \Phi(\mathbf{x}_i) | \Phi(\mathbf{x}_j) \rangle_{\mathcal{H}}, \quad (118)$$

for  $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}$ . In comparison with classical kernels, the inner product is defined in a Hilbert space by replacing the standard definition  $\langle \cdot, \cdot \rangle$  by the Dirac brackets  $\langle \cdot | \cdot \rangle$ . For a map  $\Phi$ , the *reproducing kernel Hilbert space* takes the form

$$\mathcal{R}_{\Phi} = \{f: \mathcal{X} \rightarrow \mathbb{C} \mid f(\mathbf{x}) = \langle w | \Phi(\mathbf{x}) \rangle_{\mathcal{H}}, \forall \mathbf{x} \in \mathcal{X}, |w\rangle \in \mathcal{H}\}. \quad (119)$$

The orthogonality of  $|w\rangle$  w.r.t.  $|\Phi(\mathbf{x})\rangle$  defines a decision boundary, i.e. depending on the sign of the inner product,  $\mathbf{x}$  lies in one side of the hyperplane. The function  $f$  is thus a linear function in  $\mathcal{H}$ . The *representer theorem* (Schölkopf *et al.*, 2001) states that this function can

be approximated by the linear function  $f^*$  by using the kernel defined above, i.e.

$$f^*(\mathbf{x}) = \sum_{i=1}^{\mathcal{D}} \alpha_i \kappa(\mathbf{x}, \mathbf{x}_i) \quad (120)$$

for an input dataset  $\mathcal{D}$ . Using Eq. (120), one can solve a convex optimization problem to get the coefficients  $\alpha_i$ . The analysis so far entails the connection between linear models in reproducing kernel Hilbert space with kernelized models in the input space.

One can use a quantum computer to calculate the inner product of feature mapped quantum states to obtain the kernel  $\kappa$ . This kernel can be fed to a classical device, which can use Eq. (120) to obtain the coefficients  $\alpha_i$ , for instance, by maximizing a cost function of the form (Havlíček *et al.*, 2019)

$$C(\boldsymbol{\alpha}) = \sum_{i=1}^{\mathcal{D}} \alpha_i - \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j \kappa(\mathbf{x}_i, \mathbf{x}_j), \quad (121)$$

where  $y_i$  are the labels of the training points and constrained to  $\sum_{i=1}^{\mathcal{D}} \alpha_i y_i = 0$ .

Also, a Gaussian Boson Sampling device (see Sec. III.B) can be used for computing kernel functions (Schuld *et al.*, 2020b).

Another approach is to use a variational circuit  $\mathcal{U}(\boldsymbol{\theta})$  and directly perform the classification task in the reproducing kernel Hilbert space, without using Eq. (120). This approach is sometimes referred as a variational quantum classification. Data is also embedded into the state  $|\Phi(\mathbf{x})\rangle$  and then processed with a PQC  $\mathcal{U}(\boldsymbol{\theta})$ . The resultant state becomes

$$|\Psi(\mathbf{x}, \boldsymbol{\theta})\rangle = \mathcal{U}(\boldsymbol{\theta})|\Phi(\mathbf{x})\rangle, \quad (122)$$

which parameters are estimated by training it to match the target states  $|y_i\rangle$  that represent the  $y_i$  labels of the training points, i.e. by minimizing the infidelity

$$C(\boldsymbol{\theta}) = \sum_{i=1}^{\mathcal{D}} (1 - |\langle y_i | \Psi(\mathbf{x}_i, \boldsymbol{\theta}) \rangle|^2). \quad (123)$$

Both methods require a way to encode the data into a quantum state. There are several strategies to define the quantum feature map. It is a key step in the success of the classification task, as the needed non-linearities must come from it. Furthermore, to eventually obtain any quantum advantage, one should search from the set of classically intractable feature maps. One of the first proposed approaches was the amplitude encoding (Schuld *et al.*, 2016) also required in other quantum algorithms (Harrow *et al.*, 2009). This approach encodes the classical data points into the amplitudes of a quantum state, i.e.  $|\Phi(\mathbf{x})\rangle = \sum_i x_i |e_i\rangle$ , where  $|e_i\rangle$  are the basis states. However, this raw encoding requires *i)* knowing which gates can be used to perform this operation for general

data points and *ii*) having an efficient way to extract and process these amplitudes. Although the same point can eventually be overcome by using similar approaches as the ones used to define a PQC, the second one requires tools as QRAM (Giovannetti *et al.*, 2008), experimentally challenging and not suited for the NISQ era.

In general, the encoding strategies used in state-of-the-art algorithms consist on introducing the classical data points into the parameters of the quantum circuit gates. As briefly mentioned in Sec. II.B, one designs a state preparation circuit  $E$  that encodes the data points,

$$|\Phi(\mathbf{x})\rangle = E(\mathbf{x}, \phi)|0\rangle. \quad (124)$$

The use of  $\phi$  parameters is optional and they can be subject to the optimization subroutine too.

Typically, the encoding gate is designed using the same structure of a layer-wise PQC from Eq. (26). Data points are introduced in layers of single-qubit rotational gates  $R$ , as defined in Eq. (16), followed by an entangling gate unitary  $W$ , e.g.

$$E(\mathbf{x}) = \prod_{k=1}^{L_E} \left( \bigotimes_{i=1}^n R_k(\mathbf{x}_i) \right) W_k, \quad (125)$$

with  $L_E$  being the total number of encoding layers. Then, the whole VQC is composed by this encoding circuit and the processing one to be optimized, i.e.  $\mathcal{U}_{VQC}(\boldsymbol{\theta}, \mathbf{x}) = E(\mathbf{x})U(\boldsymbol{\theta})$ .

Alternatively, some works propose to remove the distinction between the encoding  $E$  and processing  $U$  circuits and introduce the data values along the circuit (Lloyd *et al.*, 2020; Pérez-Salinas *et al.*, 2020a; Schuld *et al.*, 2020c; Vidal and Theis, 2019). This strategy, sometimes called *input-redundancy* or *data-reuploading*, introduce the data in all circuit layers, e.g.

$$\mathcal{U}_{VQC}(\boldsymbol{\theta}, \mathbf{x}) = \prod_{k=1}^L \left( \bigotimes_{i=1}^n R_k(\mathbf{x}_i, \boldsymbol{\theta}) \right) W_k, \quad (126)$$

where  $L$  is now the total number of circuit layers. This strategy has proved the universality when applied to one qubit (Pérez-Salinas *et al.*, 2020a) and can reconstruct the coefficients of the Fourier series (Schuld *et al.*, 2020c; Vidal and Theis, 2019).

The inclusion of encoding strategies and, in particular, the data re-uploading, can help well-known VQA such as the VQE. In general, one of the final goals of a VQE can be the identification of interesting points on a potential energy surface generated by a parametrized Hamiltonian. Often, one is interested in the ground state energy as a function of some Hamiltonian parameter  $\lambda$ , e.g. the interatomic distance, but other properties, like the energy gap between ground state and first excited state can be interesting as well. (Kyaw *et al.*, 2020b) To do so, one often needs to scan discretely over  $\lambda$  for some particular interval and run a VQE to obtain the ground state

energy on each of these points. This becomes an extra computational cost, especially if we are interested only in a particular region of this ground state profile, e.g. to extract the  $\lambda_{min}$  whose ground state has minimal energy. In that direction, some proposals suggest to encode the parameters of the Hamiltonian into the PQC and learn the energy profiles (Mitarai *et al.*, 2019). In particular, the *Meta-VQE* algorithm (Cervera-Lierta *et al.*, 2020) proposes to encode the  $\lambda$  into the PQC gates together with the optimization parameters. Then optimize an objective function that corresponds with the sum of expectation values for some  $M$  training  $\lambda$  parameters, i.e.  $\langle \hat{O} \rangle = \sum_{i=1}^M \langle H \rangle_{\mathcal{U}(\boldsymbol{\theta}, \lambda_i)}$ . Once the circuit has been optimized, one can run it again with the new  $\lambda_i$  to directly extract an estimation of the ground state, without having to optimize the full circuit again. An extension of this approach is the *optimized Meta-VQE* (*opt-meta-VQE*), which consist of using the optimized parameters from the Meta-VQE as starting points of a standard VQE. This tries to avoid vanishing gradients issues (discussed in Sec. IV.B) by starting in a particular region of the parameter space instead of random initialization.

VQC also need an extra piece, the definition of the target state  $|y_i\rangle$ . The goal of the quantum circuit is to divide and push the quantum states that encode the data points into two or more regions of the Hilbert space. To that aim, the parameters of the circuit are trained to match every encoded state into a particular representative of one of these regions. Therefore, the more separated these regions are, the lesser misclassified points are expected. As discussed in Sec. II.C, qubits measurement implies a certain computational cost. For that reason, many proposals suggest to use the state of only one qubit to train the whole circuit (Farhi and Neven, 2018; Schuld *et al.*, 2020a). The cost function estimation reduces to measuring the probability distribution of one qubit. Other works use a more sophisticated definition of these target states by selecting the most orthogonal states of the qubits space (Lloyd *et al.*, 2020; Pérez-Salinas *et al.*, 2020a). This strategy is inspired from optimal state discrimination (Helstrom, 1969).

Using the nonlinear character of quantum mechanical processes as “reservoir”, the notion of *quantum reservoir computing* has been suggested. The reservoir is a highly nonlinear system whose parameters are arbitrary but fixed. One can perform reservoir computing by employing a basic training algorithm such as linear regression at the readout stage. Since the reservoir parameters are fixed, only training of the readout stage parameters is required. The aforementioned idea helps utilize the high nonlinearities of the reservoir without the high computational cost of training. The concept of employing quantum systems as quantum reservoirs was first introduced in (Fujii and Nakajima, 2017; Nakajima *et al.*, 2019). Quantum reservoir computing has been proposed for many experimental platforms such as Gaus-

sian states in the optical set-up (Nokkala *et al.*, 2020), two-dimensional fermionic lattices (Ghosh *et al.*, 2019) and nuclear spins (Negoro *et al.*, 2018). Quantum gate based implementation of quantum reservoir computing for NISQ devices has also been discussed (Chen *et al.*, 2020). NISQ devices have also been used for regression (Mitarai *et al.*, 2018).

Quantum annealing has been also applied to supervised learning to predict biological data (Li *et al.*, 2018). Here, the quantum annealer is used to train the parameters of the classification model, which is done by mapping the problem of finding the optimal parameters to a minimization of a QUBO.

## 2. Unsupervised learning

The use of quantum devices to speed up different unsupervised learning tasks has been investigated thoroughly, leading to different algorithms for generative modelling (Benedetti *et al.*, 2019a, 2017), clustering (Otterbach *et al.*, 2017), among others (Lloyd *et al.*, 2013). An analysis of quantum speedup in unsupervised learning for Fault-Tolerance algorithms is presented in (Aïmeur *et al.*, 2013). The task of learning probabilistic generative models in particular has been of interest to the quantum machine learning community, because of the potential advantage quantum computers may exhibit over their classical counterparts in the near future.

Generative Modelling involves learning the underlying probability distribution from a finite set of samples from a data set, and generating new samples from the distribution. There have been several proposals for using parameterized quantum circuits as models for generative learning (Amin *et al.*, 2018; Benedetti *et al.*, 2019a, 2018), including quantum Boltzmann machines, quantum circuit Born machines, quantum assisted Helmholtz machines, quantum generative adversarial networks, amongst others (Benedetti *et al.*, 2019b). We discuss some of these proposals in detail hereafter.

*Quantum Boltzmann Machines.* The quantum Boltzmann machine (Amin *et al.*, 2018) (QBM) extends the classical Boltzmann machine (Ackley *et al.*, 1985), a neural architecture capable of several tasks including generative modeling of data. Such models take their name from their physical inspiration, namely, the Boltzmann distribution over the Ising model in the classical case, and the Boltzmann distribution over the transverse-field Ising model, for the quantum case. Such a network consists of a mixture of *visible* and *hidden* vertices, connected by weighted edges. The visible vertices function as both input and outputs to the network, whilst the hidden vertices add extra degrees of freedom to the network.

The QBM can be modeled with the Hamiltonian

$$H = - \sum_a^N (b_a \hat{\sigma}_z^a + \Gamma_a \hat{\sigma}_x^a) - \sum_{a,b} \omega_{ab} \hat{\sigma}_z^a \hat{\sigma}_z^b, \quad (127)$$

where  $b_a$ ,  $\Gamma_a$ , and  $\omega_{ab}$  are the parameters to be fine-tuned to generate the training data. Defining the density matrix  $\rho = \frac{e^{-H}}{Z}$  with  $Z$  the usual partition function,  $Z = \text{Tr}(e^{-H})$ , the marginal probability that the visible variables are in some state  $\mathbf{v}$  is given by  $P_v = \text{Tr}(\Lambda_v \rho)$ , with  $\Lambda_v = \left( \otimes_v \frac{1 + v_v \hat{\sigma}_z^v}{2} \right) \otimes I_h$ , a projector onto the subspace spanned by the visible variables tensor the identity acting on the hidden variables. The objective of training the QBM, then, is to get the family of probability distributions  $P_v$  to match the family inherent to the data,  $P_v^{\text{data}}$ , for arbitrary  $\mathbf{v}$ . This is achieved by minimizing the negative log-likelihood measure shown below

$$\mathcal{L} = - \sum_v P_v^{\text{data}} \log \frac{\text{Tr} \Lambda_v e^{-H}}{\text{Tr} e^{-H}}. \quad (128)$$

The gradients of  $\mathcal{L}$  with respect to the Hamiltonian parameters are difficult to calculate by sampling the Boltzmann machine, both classically and in the quantum variant. Methodologies of approximating these gradients are necessary to advance the deployment of QBM's.

The QBM may be trained both to be a generator, or a discriminator, with respect to the distribution it is trained to mimic. Consider the joint distribution of input and output variables  $x$  and  $y$  respectively. In the discriminative case, the objective is to minimize negative log-likelihood with respect to  $P_{y|x}$ . For generative learning, the goal is to learn the joint distribution  $P_{x,y}$  directly.

The implementation of the QBM designed by (Amin *et al.*, 2018) found that a ten qubit QBM with only visible vertices is able to learn a mixture of randomly generated Bernoulli distributions more effectively than a classical Boltzmann machine, and performed better in generative applications. (Kieferová and Wiebe, 2017) found that a QBM outperformed classical Boltzmann machines in generative training to reproduce small Haar-random states. Extensions to the QBM, such as the Variational Quantum Boltzmann Machine (VQBM) (Zoufal *et al.*, 2020), have improved upon trainability. In addition to its generative capacities, QBMs have shown potential in reinforcement learning (Crawford *et al.*, 2019a), in which they have been shown to achieve better fidelity to data distributions than do restricted Boltzmann machines or deep Boltzmann machines (classical boltzman machines with layers of hidden vertices) of similar sizes. To suit NISQ devices, it has been shown that QBM can be trained robustly using QAOA as a subroutine (Verdon *et al.*, 2017).

*Quantum Circuit Born Machines.* Parametrized quantum circuits can function as generative models to sample from



probability distribution. The Quantum Circuit Born Machine (QCBM) outputs bitstrings  $\mathbf{x}$  sampled from measurements in the computational basis of a quantum circuit  $U(\boldsymbol{\theta})$ , with the probability of each bit string given by the Born rule  $p_{\boldsymbol{\theta}}(\mathbf{x}) \sim |\langle \mathbf{x} | U(\boldsymbol{\theta}) | 0 \rangle|^2$ . The goal is that the distribution of the QCBM matches the one from a given target distribution.

QCBMs can prepare entangled quantum states by training them to match the probability distribution corresponding to the desired quantum state (Benedetti *et al.*, 2019a). In (Liu and Wang, 2018), training of QCBMs using the gradients of the parameterized quantum circuit was proposed using the maximum mean discrepancy loss, which calculates the difference of the sampled output from the quantum circuit and the desired distribution in a Kernel feature space.

QCBMs are well suited to be run on current NISQ hardware and can serve as benchmarks (Hamilton *et al.*, 2019; Leyton-Ortega *et al.*, 2019) and have been applied to tasks such as generating images (Rudolph *et al.*, 2020) or financial data (Alcazar *et al.*, 2020; Coyle *et al.*, 2020a). It has been shown that QCBMs can potentially outperform classical computers as they are able to sample from probability distributions that are difficult for classical computers (Coyle *et al.*, 2020b; Du *et al.*, 2020).

**Quantum Generative Adversarial Networks.** Generative adversarial learning (Goodfellow *et al.*, 2014) has been one of the most recent breakthrough in machine learning, and have become very powerful tool in the machine learning community, for image and video generation, and materials discovery. The GAN consists of two networks, a generator,  $F_G(z; \theta_g)$  and a discriminator,  $F_D(x; \theta_d)$  - with parameters  $\theta_g$  and  $\theta_d$  respectively, playing an adversarial game, which can be summarized as follows:

$$\min_{\theta_g} \max_{\theta_d} (E_{x \sim p_{data}(x)} [\log(F_D(x))] + E_{z \sim p_z(z)} [\log(1 - F_D(F_G(z)))]) \quad (129)$$

where  $p_z(z)$  is a fixed prior distribution,  $p_{data}(x)$  is the target distribution,  $x$  is the data sampled from  $p_{data}(x)$ , and  $z$  is the noise sampled from  $p_z(z)$ . The training of GAN is carried iteratively, until the generator produces a distribution indistinguishable from the target distribution.

A quantum version of generative adversarial networks (GANs) was proposed theoretically in Refs. (Dallaire-Demers and Killoran, 2018; Lloyd and Weedbrook, 2018) and further developed for near term quantum devices in Refs. (Romero and Aspuru-Guzik, 2019; Situ *et al.*, 2020; Zeng *et al.*, 2019), where parameterized quantum circuits are used for adversarial learning instead of classical neural networks.

The different adaptations of quantum GANs can be divided into different categories, based on the data and

networks used being classical and quantum (Romero and Aspuru-Guzik, 2019). There have been different studies with hybrid models of GANs using both classical and quantum data, and it has been shown that the training of these networks are robust to moderate level of noise (Anand *et al.*, 2020b).

The training of quantum GANs has been demonstrated experimentally on various quantum processing units, for a variety of tasks including, quantum state estimation (Hu *et al.*, 2019), image generation (Huang *et al.*, 2020b,d), generating continuous distributions (Anand *et al.*, 2020b), learning distribution (Zoufal *et al.*, 2019), among others (Nakaji and Yamamoto, 2020b).

### 3. Reinforcement learning

The general framework of reinforcement learning (RL) involves an agent interacting with an environment attempting to maximize an underlying reward function.

The mathematics of RL can be captured using Markov decision process (MDP) (Sutton and Barto, 2018). An MDP is a 4-tuple  $(S, A, R, P)$ , where  $S$  is the set of all possible valid states;  $A$  is the set of all possible actions;  $R$  is the reward function, i.e. a map  $R: S \times A \times S \rightarrow \mathbb{R}$ ; and  $P$  is the transition probability, i.e. a map  $P: S \times A \rightarrow [0, 1]$ . Specifically, the transition probability  $P(\tilde{s}|s, a)$  represents the probability of transition to state  $\tilde{s}$  given the present state is  $s$  and the action  $a$  has been taken. The term ‘‘Markov’’ in MDP means that transitions are memory-less and depend only on the current state and action. The agents in reinforcement learning learn via trial and error. For a successful training, a proper balance between exploration of unknown strategies and exploitation of prior experience is required.

The training happens via agent-environment interaction. At the beginning of time step  $t$ , the environment state is  $s_t$ . From the set  $A$ , the agent selects an action  $a_t$ . The transition probability dictates the next state of the environment  $s_{t+1}$  and the agent gets reward  $r_{t+1}$  based on the reward function  $R$ . The agent-environment interaction yields a series of states and actions of the form  $\tau = (s_1, a_1, s_2, a_2, \dots, s_H, a_H)$ . The aforementioned series is called a trajectory and the number of interactions ( $H$ ) in an episode is called horizon. Suppose the probability of a trajectory is  $P(\tau)$  and the corresponding cumulative reward is  $R_{tot}(\tau)$ . Then, the expected reward is  $\sum_{\tau} P(\tau) R_{tot}(\tau)$ .

By harnessing quantum mechanical phenomena such as superposition and entanglement, one can expect to achieve speedups in the reinforcement learning tasks (Dong *et al.*, 2008; Dunjko *et al.*, 2016, 2017; Paparo *et al.*, 2014). The aforementioned intuition has led to recent works towards quantum reinforcement learning (Cornelissen, 2018; Dunjko *et al.*, 2017)

We discuss the essence of quantum reinforcement learning

ing by providing a brief synopsis of quantum agent environment (AE) paradigm. For details, refer to (Dunjko *et al.*, 2017). In the AE paradigm, agent and environment are modelled via sequences of unitary maps  $\{\mathcal{E}_A^j\}_j$  and  $\{\mathcal{E}_E^j\}_j$ , respectively. The agent and environment have access to memory registers belonging to Hilbert spaces  $\mathcal{H}_A$  and  $\mathcal{H}_E$ . The communication register between the agent and the environment belongs to Hilbert space  $\mathcal{H}_C$ . The agent maps  $\{\mathcal{E}_A^j\}_j$  act on  $\mathcal{H}_A \otimes \mathcal{H}_C$  and the environment maps  $\{\mathcal{E}_E^j\}_j$  act on  $\mathcal{H}_E \otimes \mathcal{H}_C$ . The agent and environment interact with each other by applying their maps sequentially. The set of actions and states correspond to orthonormal set of vectors  $\{|a\rangle|a \in A\}$  and  $\{|s\rangle|s \in S\}$  respectively. The Hilbert space corresponding to the communication register is given by  $\mathcal{H}_C = \text{span}(|y\rangle|y \in S \cup A)$ . The classical AE paradigm corresponds to the case where the agent and environment maps are classical.

Quantum reinforcement learning has been studied for algorithm such as SARSA, and Q Learning (Jerbi *et al.*, 2019), which are some of the elementary reinforcement learning algorithms (Sutton and Barto, 2018).

In the set-up of variational quantum circuits, reinforcement learning has been explored for small input sizes (Chen *et al.*, 2020). This work revealed a possibility of quadratic advantage in parameter space complexity. Using better encoding schemes, (Lockwood and Si, 2020b) showed the case of reinforcement learning with variational quantum circuits for larger input sizes. In a follow-up work, (Lockwood and Si, 2020a) demonstrated the possibility of dealing with the relatively complicated example of playing Atari games.

Reinforcement learning with quantum annealers has also been investigated by (Crawford *et al.*, 2016). In their framework, they explore reinforcement learning with quantum Boltzmann machines. A detailed study of basic reinforcement learning protocols with superconducting circuits is provided in (Lamata, 2017). Some exciting proposals of reinforcement learning with trapped ions and superconducting circuits have also been proposed recently (Cárdenas-López *et al.*, 2018). For quantum eigensolvers, reinforcement learning study has been carried out recently (Albarrán-Arriagada *et al.*, 2020). Reinforcement learning with optical set-up has been discussed in (Yu *et al.*, 2019).

### C. Combinatorial optimization

Given a finite set of objects, say  $S$ , combinatorial optimization deals with finding an optimal object from the set  $S$ . It is a sub-discipline of mathematical optimization theory, with applications in diverse fields such as artificial intelligence, logistics, supply chain and theoretical computer science. Some typical examples of combinatorial optimization problems are the traveling sales-

man problem (Lenstra and Kan, 1975), job-shop scheduling (Manne, 1960), max-cut (Festa *et al.*, 2002) and Boolean satisfiability (Tovey, 1984).

To understand combinatorial optimization, let us consider the canonical problem of Boolean satisfiability. Boolean variables admit two truth values, “TRUE” and “FALSE”. Boolean variables can be combined together using operators AND or conjunction (denoted by  $\wedge$ ), NOT or negation (denoted by  $\neg$ ), and OR or disjunction (denoted by  $\vee$ ). These combinations are called Boolean expressions.

A Boolean expression is said to be satisfiable if it can be TRUE for appropriate assignment of logical values to its constituent Boolean variables. Given a Boolean expression  $E$ , the Boolean satisfiability problem (SAT) is to check if  $E$  is satisfiable. The famous Cook-Levin theorem showed that SAT is NP-complete (Arora and Barak, 2009). The study of SAT finds applications in theoretical computer science, cryptography and artificial intelligence.

Every combinatorial optimization problem can be expressed as  $m$  clauses over  $n$  Boolean variables. A Boolean variable is known as positive literal, while its negation is known as a negative literal. Thus, a literal is either a variable or its negation. A disjunction of literals is known as clause or constraint. For every constraint  $C_\alpha$  for  $\alpha \in \{1, 2, \dots, m\}$  and every string  $z \in \{0, 1\}^n$ , let

$$C_\alpha(z) = \begin{cases} 1 & \text{if } z \text{ satisfies } C_\alpha(z) \\ 0 & \text{if } z \text{ does not satisfy} \end{cases}$$

The goal of a combinatorial optimization problem, framed as such, is to find a string which maximizes the following objective function,

$$C(z) = \sum_{\alpha=1}^m C_\alpha(z), \quad (130)$$

which counts the number of satisfied constraints.

Approximate optimization algorithms such as QAOA seeks to find a solution  $z$  (usually a bit-string) with a desired approximation ratio  $C(z)/C_{\max} \geq r^*$ , where  $C_{\max}$  is the maximum value of  $C(z)$ . Using  $C(z)$  and computational basis vectors  $|e_i\rangle \in \mathbb{C}^{2^n}$  for  $i = 1, \dots, 2^n$ , one can construct the problem Hamiltonian as the one in Eq. (21). The list of the NISQ algorithms for combinatorial optimization discussed in the following lines are listed in Table III.

#### 1. Max-Cut

Max-Cut is an important combinatorial optimization problem with applications in diverse fields such as theoretical physics and circuit design. In theoretical physics, the Max-Cut problem is equivalent to finding the ground

state and ground state energy of a spin glass Hamiltonian. Given a graph  $G = (V, E)$  for vertex set  $V$  and edge set  $E$ , a cut is a partition of the elements of  $V$  into two disjoint subsets. Given a weight function  $w : E \rightarrow \mathbb{R}^+$  such that the edge  $(i, j) \in E$  has weight  $E_{ij}$ , the Max Cut Problem is to find a cut  $K \cup \bar{K} = V$  that maximizes

$$\sum_{i \in K, j \in \bar{K}, (i, j) \in E} w_{ij}. \quad (131)$$

For every vertex  $v_i \in V$ , let us associate a variable  $x_i$  which takes values  $\pm 1$ . Given an arbitrary cut  $K \cup \bar{K} = V$ , let us define  $x_i = 1$  if  $v_i \in K$  and  $-1$  otherwise. The Max Cut problem is equivalent to the following quadratic program,

$$\max \sum_{(v_i, v_j) \in E} w_{ij} \frac{(1 - x_i x_j)}{2}, \quad (132)$$

subject to  $x_i \in \{-1, +1\} \forall v_i \in V$ .

Considering  $n$  vertices as  $n$  qubits in the computational basis, we can classify qubits by assigning quantum states  $|0\rangle$  or  $|1\rangle$ . For the classical objective function in the optimization program in 132, we can use the following Hamiltonian as the problem Hamiltonian,

$$H_P = \sum_{(i, j) \in E} \frac{1}{2} (I - \hat{\sigma}_z^i \otimes \hat{\sigma}_z^j) \equiv \sum_{(i, j) \in E} C_{ij}, \quad (133)$$

It has been shown that it is NP-hard to achieve an approximation ratio of  $r^* \geq 16/17 \approx 0.9412$  for Max-Cut on all graphs (Håstad, 2001). For the QAOA with  $p = 1$ , it has been proven that for a general graph,

$$\begin{aligned} \langle C_{ij} \rangle &= \frac{1}{2} + \frac{1}{4} (\sin 4\beta \sin \gamma) (\cos^{d_i} \gamma + \cos^{d_j} \gamma) \\ &\quad - \frac{1}{4} (\sin^2 \beta \cos^{d_i + d_j - 2\lambda_{ij}} \gamma) (1 - \cos^{\lambda_{ij}} 2\gamma) \end{aligned} \quad (134)$$

where  $d_i + 1$  and  $d_j + 1$  denote the degrees of vertices  $i$  and  $j$  respectively, and  $\lambda_{ij}$  is the number of triangles containing the edge  $(i, j)$  in the graph (Wang et al., 2018b). Here,  $\gamma$  and  $\beta$  refer to the QAOA parameters in Eq. (23).

In the case of unweighted 3-regular (u3R) graphs, the above result gives the approximation ratio of 0.692, which is consistent with the pioneering result by Farhi, Goldstone and Gutman (Farhi et al., 2014). In comparison, the best classical algorithms to date give the approximation ratio of  $r^* \approx 0.8786$  for general graphs (Goemans-Williamson algorithm (Goemans and Williamson, 1995)) and  $r^* \approx 0.9326$  for u3R graphs (Halperin et al., 2004), using semidefinite programming. While QAOA for  $p = 1$  does not outperform its classical counterparts for the Max-Cut problem, QAOA has been found to surpass the Goemans-Williamson bound for larger values of  $p$  (Crooks, 2018). QAOA has also been applied to the clustering problem (from unsupervised learning) by mapping it to Max-Cut (Otterbach et al., 2017).

Recently, the performance of QAOA for a class of Max-Cut problems was investigated theoretically and it was found to be equal or inferior to a simple classical local algorithm (Hastings, 2019). Furthermore, it was proven that for a wide class of Max-Cut graphs constant depth  $p$  QAOA is inferior to the Goemans-Williamson algorithm (Bravyi et al., 2019). This casts some severe doubts on how QAOA can show quantum advantage for combinatorial problems. In response to these results, a non-local version of QAOA called recursive QAOA (RQAOA) was proposed (Bravyi et al., 2019). The core idea is as follows. Run QAOA as a subroutine on a specific problem with  $N$  qubits and measure the expectation values of the correlations between the all qubit pairs  $i, j$  with  $M_{ij} = \langle \sigma_z^i \sigma_z^j \rangle$ . Then, pick out the qubit pair with maximal absolute value of correlation  $n, m = \max_{(i, j)} |M_{ij}|$ . For  $M_{nm} > 0$ , the selected qubit pair  $n, m$  are positively correlated and very likely to be in the same state, whereas for  $M_{nm} < 0$  they are anti-correlated and likely to be in opposite state. Now, this correlation is fixed as a constraint on the problem by fixing the state of the qubit  $\sigma_z^m = \text{sign}(M_{nm}) \sigma_z^n$ . With this constraint, one of the two qubits can be removed as its state is completely determined by the other, reducing the qubit number by one. The QAOA subroutine is run again on the reduced problem of size  $N - 1$  qubits together with the above steps of measuring correlation and fixing a qubit pairs. The RQAOA algorithm is run recursively until the size of the problem is reduced to a small number of qubits such that it can be solved easily classically.

When RQAOA is run with the QAOA subroutine of depth  $p = 1$ , it can efficiently simulated on a classical computer, which can serve as an important benchmark with classical algorithms (Bravyi et al., 2019). Numerical experiments with higher  $p$  suggest similar or better performance on combinatorial problems compared to other classical algorithms (Bravyi et al., 2020b; Egger et al., 2020b).

Finally, QAOA with depth  $p = 1$  has been investigated in comparison with quantum annealing (Streif and Leib, 2020). QAOA is connected to quantum annealing in the sense that in the limit of infinite depth  $p$ , QAOA is equivalent to quantum annealing (refer to Sec. II.B.1 for QAOA, as well as Sec. III.A for quantum annealing). However, QAOA can outperform quantum annealing on specific problems even at depth  $p = 1$ . In fact, QAOA can solve specific problems perfectly for  $p = 1$ , arriving at the correct solution with unit probability, whereas quantum annealing struggles here to find the solution (Streif and Leib, 2020). This shows that QAOA is strictly more powerful than quantum annealing.

## 2. Other combinatorial optimization problems

While the usage of QAOA on Max-Cut has been studied extensively, QAOA has also seen applications in other important combinatorial optimization problems, such as: Max- $k$  Vertex Cover, which seeks to find the set of  $k$  vertices on a graph that maximizes the number of edges incident on the vertices (Cook *et al.*, 2019); Exact cover problem (given a set  $X$  and several subsets  $S_i$ , find the combination of subsets which contains all elements just once) with applications to the tail-assignment problem (Bengtsson *et al.*, 2020; Vikstål *et al.*, 2020); lattice protein folding (Fingerhuth *et al.*, 2018); knapsack problem as applied to battery revenue optimization (de la Grandrève and Hullo, 2019); multi-coloring graph problems (Oh *et al.*, 2019); maximum independent set problems (Choi *et al.*, 2020; Saleem, 2020) with applications to scheduling; and the vehicle routing problem (Utkarsh *et al.*, 2020). An adiabatically assisted approach was suggested in (Garcia-Saez and Latorre, 2018) to tackle combinatorial optimization problems. Gaussian Boson Sampling (see Sec. III.B) has been used to assist in a wide variety of combinatorial optimization problems (Arrazola *et al.*, 2018; Bromley *et al.*, 2020), most prominently to solve Max-Clique (Arrazola and Bromley, 2018; Banchi *et al.*, 2020b). This has applications in predicting molecular docking configurations (Banchi *et al.*, 2020a), computing vibrational spectra of molecules (Huh *et al.*, 2015), and electron-transfer reactions (Jahangiri *et al.*, 2020a).

Quantum Annealing, which has been the inspiration of QAOA, is a prominent platform that has been applied to various combinatorial optimization problems and its applications, such as protein folding (Perdomo-Ortiz *et al.*, 2012). These have been reviewed in (Bouland *et al.*, 2020; Hauke *et al.*, 2020).

## D. Numerical solvers

We proceed to discuss NISQ algorithms used to solve numerical problems such as factoring, singular value decomposition, linear equations and non-linear differential equations, all of them listed in Table IV.

### 1. Variational quantum factoring

The factoring problem accepts a composite positive integer  $N$  as input and returns its prime factors as output. There is no known efficient classical algorithm for prime factorization and the hardness of factoring is used to provide the security in the RSA public-key cryptosystems. The famous Shor’s factoring algorithm is a polynomial time quantum algorithm for the factoring problem (Shor, 1999) (implying prime factorization is in BQP) and hence has been extensively investigated by quantum computing

researchers (for details refer to (Anshuetz *et al.*, 2019) and references therein). The resource estimates for implementing the Shor’s algorithm is, however, way beyond the capabilities of the NISQ era. A detailed analysis has shown that factoring a 2048-bit RSA number would necessitate a quantum processor with  $10^5$  logical qubits and circuit depth on the order of  $10^9$  to run for roughly 10 days (Jones *et al.*, 2012; Van Meter *et al.*, 2010). On a photonic architecture, using 1.9 billion photonic modules, factoring a 1024-bit RSA number is expected to take around 2.3 years (Devitt *et al.*, 2013). To tackle the factoring problem in the near-term quantum devices, it is imperative to develop NISQ-era compatible alternatives to Shor’s factoring algorithm.

The factoring problem can be mapped to the ground state problem of an Ising Hamiltonian (Burgess, 2002; Dattani and Bryans, 2014). To understand the afore-said mapping, let us consider the factoring of  $m = p \times q$ . Suppose the binary representations of  $m, p$  and  $q$  are  $m = \sum_{k=0}^{n_m-1} 2^k m_k$ ,  $p = \sum_{k=0}^{n_p-1} 2^k p_k$  and  $q = \sum_{k=0}^{n_q-1} 2^k q_k$ . Here,  $m_k \in \{0, 1\}$  is the  $k$ th bit of  $m$  and the total number of bits for  $m$  has been denoted by  $n_m$ . Similar notation has been employed for  $p$  and  $q$ . Since  $m = p \times q$ , it induces  $n_c = n_p + n_q - 1$  constraints on the individual bits of  $m, p$  and  $q$ ,

$$\sum_{j=0}^i q_j p_{i-j} + \sum_{j=0}^i z_{j,i} - m_i - \sum_{j=1}^{n_c} 2^j z_{i,i+j} = 0, \quad (135)$$

for  $i \in [0, n_c)$  and the carry bit from position  $i$  to position  $j$  has been represented by  $z_{i,j}$ . The constraint  $i$  in Eq. (135) induces clause  $C_i \equiv \sum_{j=0}^i q_j p_{i-j} + \sum_{j=0}^i z_{j,i} - m_i - \sum_{j=1}^{n_c} 2^j z_{i,i+j}$  over  $\mathbb{Z}$  such that factoring can be modelled as assignment of binary variables  $\{m_i\}$ ,  $\{p_i\}$  and  $\{q_i\}$  which solves  $\sum_{i=0}^{n_c-1} C_i^2 = 0$ .

One can map the binary variables to quantum observables to quantize the clause  $C_i$  to  $\hat{C}_i$  using the mapping  $b_k \rightarrow \frac{1}{2}(1 - \sigma_{b,k}^z)$  and obtain the Hamiltonian  $H_P = \sum_{i=0}^{n_c-1} \hat{C}_i^2$ , which we refer as factoring Hamiltonian. Note that the factoring Hamiltonian is a 4-local Ising Hamiltonian.

By using the aforementioned ideas, one can use NISQ algorithms for the ground state problem to tackle the factoring problem (see Sec. V.A.3 and Sec. V.C). In Ref. (Anshuetz *et al.*, 2019), authors employ QAOA to find the ground state of the factoring Hamiltonian and refer to their Algorithm as variational quantum factoring (VQF) algorithm. Numerical simulations were provided for numbers as high as 291311. For a recent experimental realization and detailed analysis of VQF, refer to (Karamlou *et al.*, 2020).



## 2. Singular value decomposition

Given a matrix  $M \in \mathbb{C}^{m \times n}$ , the Singular Value Decomposition (SVD) provides a factorization of the form  $M = U \Sigma V^\dagger$ , where  $U \in \mathbb{C}^{m \times m}$  is a unitary matrix,  $\Sigma \in \mathbb{R}_+^{m \times n}$  is a rectangular diagonal matrix with non-negative real diagonal entries and  $V \in \mathbb{C}^{n \times n}$  is a unitary matrix. The diagonal entries of  $\Sigma$  are called the singular values of matrix  $M$ . The columns of the unitary matrices  $U$  and  $V$  are called left-singular and right-singular vectors of  $M$ . Using Dirac notation, one can write

$$M = \sum_{j=1}^r d_j |u_j\rangle\langle v_j|. \quad (136)$$

where  $d_j, |u_j\rangle, |v_j\rangle$  are singular values, left-singular vectors and right-singular vectors. The rank of matrix  $M$  is  $r$  and is equal to the number of non-zero singular values.

SVD finds applications in calculating pseudoinverse (Gregorcic, 2001), solving homogeneous linear equations (Klema and Laub, 1980), signal processing (Vandewalle and De Moor, 1991) and recommendation systems (Koren et al., 2009). Moreover, the notion of Schmidt decomposition which is used to study entanglement of bipartite quantum states, is related to SVD.

In the quantum information context, the SVD can be used to compute the Schmidt decomposition of bipartite quantum states. For a quantum state  $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , the Schmidt decomposition is given by

$$|\psi\rangle = \sum_i d_i |u_i\rangle |v_i\rangle, \quad (137)$$

where  $d_i$  are non-negative real numbers such that  $\sum_i d_i^2 = 1$ . Moreover,  $\{|u_i\rangle\}_i$  and  $\{|v_i\rangle\}_i$  correspond to orthonormal basis sets for  $\mathcal{H}_A$  and  $\mathcal{H}_B$  respectively. The number of non-zero  $d_i$ , say  $\chi$ , is called the Schmidt rank of the quantum state  $|\psi\rangle$  and is used to quantify the bipartite entanglement. To calculate the Schmidt decomposition, one can write the bipartite quantum state as a matrix  $|\psi\rangle = \sum_{i,j} A_{ij} |i\rangle |j\rangle$ , where  $|i\rangle$  and  $|j\rangle$  are the computational basis states of each qubit, and perform SVD of the matrix  $A$ .

Ref. (Bravo-Prieto et al., 2020) provide a NISQ algorithm to perform SVD of pure bipartite states. Starting with two unitary circuits, which act on different bipartitions of the system, the authors variationally determine the singular values and singular vectors by training the circuits on exact coincidence of outputs. The central ideas of their method is to variationally find circuits that provides the following transformation of the initial quantum state  $|\psi\rangle_{AB}$  with Schmidt rank  $\chi$ ,

$$U_A \otimes V_B |\psi\rangle_{AB} = \sum_{i=1}^{\chi} \lambda_i e^{i\gamma_i} |e_i\rangle_A |e_i\rangle_B, \quad (138)$$

where  $U_A |v_i\rangle_A = e^{i\alpha_i} |e_i\rangle_A$ ,  $V_B |v_i\rangle_B = e^{i\beta_i} |e_i\rangle_B$  such that  $\alpha_i = \beta_i + \gamma_i \in [0, 2\pi)$  and  $\{|e_k\rangle_{A,B}\}_k$  are the computational

basis states in  $\mathcal{H}_{A,B}$ . Using their algorithm, authors also suggest the possibility to implement SWAP gate between parties  $A$  and  $B$  without the requirement of any gate connecting the two subsystems.

Using variational principles for singular values and Ky Fan theorem (Fan, 1951), (Wang et al., 2020c) provide an alternative NISQ algorithm for SVD. The authors provide proof of principle application of their algorithm in image compression of handwritten digits. They also discuss the applications of their algorithm in recommendation systems and polar decomposition.

## 3. Linear system problem

Systems of linear equations play a crucial role in various areas of science, engineering and finance. Given a matrix  $A \in \mathbb{C}^{N \times M}$  and  $\mathbf{b} \in \mathbb{C}^N$ , the task of the linear system problem (LSP) consists of finding  $\mathbf{x} \in \mathbb{C}^M$  such that

$$A\mathbf{x} = \mathbf{b}. \quad (139)$$

Depending on the dimensions  $M$  and  $N$ , the LSP takes various forms. If  $M = N$  and  $A$  is invertible,  $\mathbf{x} = A^{-1}\mathbf{b}$  is unique. If  $M \neq N$ , the LSP can be under-determined or over-determined. For the sake of simplicity, it is natural to assume the matrix  $A$  to be square i.e.  $M = N$ . If the matrix  $A$  has at most  $s$  non-zero elements per row or column, the LSP is called  $s$ -sparse.

The quantum version of the LSP, known as the quantum linear system problem (QLSP), assumes  $A$  to be  $N \times N$  Hermitian matrix and  $\mathbf{b}$  to be a unit vector, i.e. it can be represented as a quantum state  $|b\rangle = \sum_{i=1}^N b_i |e_i\rangle$ . The QLSP problem thus is formulated as

$$A|x\rangle = |b\rangle \rightarrow |x\rangle = A^\dagger |b\rangle. \quad (140)$$

The first quantum algorithm proposed for solving the QLSP was the famous Harrow-Hassidim-Lloyd (HHL) algorithm (Harrow et al., 2009). Apart from the size of the matrix  $A$ , i.e.  $N$ , and its sparsity  $s$ , two other dominant factors determining the running time of a LSP or QLSP algorithm are the condition number ( $\kappa$ ) of the matrix  $A$  and the additive error ( $\epsilon$ ) corresponding to the solution. The condition number of a matrix  $A$  is given by ratio of maximal and minimal singular values of  $A$ . The best classical algorithm for LSP is the conjugate gradient method with runtime complexity  $O(Ns\kappa \log(\frac{1}{\epsilon}))$ . On the other hand, the HHL algorithm for QLSP, as originally proposed, has runtime complexity  $O(\log(N)s^2 \frac{\kappa^2}{\epsilon})$ . Further works on the HHL algorithm has improved  $\kappa$  scaling to linear (Ambainis, 2012) and error dependence to  $\text{poly}(\log(\frac{1}{\epsilon}))$  (Childs et al., 2017). Implementation of HHL algorithm, however, requires the fault-tolerant architecture and hence its guarantees can not be leveraged on the NISQ architecture. The largest QLSP solved on

a gate based quantum computer corresponds to its implementation on an nuclear magnetic resonance (NMR) processor for  $N = 8$  (Wen *et al.*, 2019).

Recently, a few VQA based implementations of the QLSP were proposed (Bravo-Prieto *et al.*, 2019; Huang *et al.*, 2019; Xu *et al.*, 2019b). Given a QLSP with input  $A$  and  $|b\rangle$ , the idea is to find the ground state of the following Hamiltonian,

$$H(u) = A(u)P_{+,b}^\perp A(u), \quad (141)$$

where  $A(u)$  and  $P_{+,b}^\perp$  are defined as

$$A(u) \equiv (1 - u)\sigma_z \otimes I + u\sigma_x \otimes A, \quad (142)$$

$$P_{+,b}^\perp = I - |+, b\rangle\langle b, +|. \quad (143)$$

Both  $A$  and  $|b\rangle$  are assumed to be constructed efficiently with a quantum circuit, i.e.  $A = \sum_{k=1}^{K_A} \beta_k U_k$  and  $|b\rangle = U_b|0\rangle$ , with  $K_A = O(\text{poly}(\log N))$ . The phase in  $\beta_k$  can be absorbed in  $U_k$  and hence one can assume  $\beta_k > 0$ .

The Hamiltonian in Eq. (141) for  $u = 1$ , has a unique ground state,  $|+\rangle|x^*\rangle = |+\rangle \frac{A^{-1}|b\rangle}{\|A^{-1}|b\rangle\|_2}$ , with zero ground state energy. After removing the ancilla, the ground state can be seen to be proportional to  $A^{-1}|b\rangle$ . Thus, one can define the following loss function,

$$L_H(|x\rangle) = \langle +, x | H(1) | +, x \rangle. \quad (144)$$

Without the ancilla, the above loss function can be written as  $L_H(|x\rangle) = \langle x | A^2 | x \rangle - \langle x | A | b \rangle \langle b | A | x \rangle$ .

In (Huang *et al.*, 2019), authors analyze the optimization landscape for VQA based optimization for the loss function of Eq. (144) and show the presence of barren plateaus which persist independent of the architecture of the quantum circuit for generating  $|x(\theta)\rangle$ . Even techniques based on adiabatic morphing (Garcia-Saez and Latorre, 2018) fail to evade the effect of the barren plateaus. To circumvent the barren plateau problem, (Huang *et al.*, 2019) proposed a classical-quantum hybrid state (see also Sec. III.E and Eq. (58))  $x = \sum_{i=1}^r \alpha_i |\psi_i(\theta_i)\rangle$ , where  $\alpha_i \in \mathbb{C}$  and  $\theta_i \in \mathbb{R}^{k_i}$  for  $i \in \{1, 2, \dots, r\}$ . Note that  $\theta_i$  are the usual variational parameters and  $\alpha_i$  are the combination parameters. These parameters are stored on a classical device and the state  $x$  is not explicitly created on a quantum processor. Moreover,  $x$  may not be normalized. To solve the QLSP, one minimizes the following loss function,

$$L_R(x) = \|Ax - |b\rangle\|_2^2 = x^\dagger A^\dagger A x - 2\text{Re}\{\langle b | Ax \rangle\} + 1. \quad (145)$$

Since optimization with respect to  $\theta_i$  suffers from the barren plateau problem, one can fix and subsequently drops the variational parameter  $\theta_i$ .

The optimization landscape is convex in  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)$ . Starting from  $|\psi_1\rangle = |b\rangle$ , other quantum states can be generated using the Ansatz tree approach in (Huang *et al.*, 2019). It was proved that finding

the combination parameters of  $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_r\rangle$  to minimize  $L_R(\sum_{i=1}^r \alpha_i |\psi_i\rangle)$  is BQP complete. Moreover, using  $O(K_A^2 \frac{\tau^2}{\epsilon})$  measurements, one can find  $\epsilon$ -suboptimal solution. With this approach, linear systems as high as  $2^{300} \times 2^{300}$  can be solved by considering cases which are also classically tractable.

#### 4. Non-linear differential equations

Nonlinear differential equations (NLDE) are a system of differential equations (DE) that cannot be expressed as a linear system. The numerical approaches to tackle DE can be local or global. Local methods employ numerical differentiation techniques (Butcher, 1987) such as the Runge-Kutta or discretization of the space of variables. Global methods, on the other hand, represent the solution via a suitable basis set, and the goal remains to find optimal coefficients (Gottlieb and Orszag, 1977). In many cases, as the number of variables or nonlinearity in the differential equations increase, finding solutions becomes challenging. To achieve higher accuracy, local methods require a fine grid, which renders high computational cost. In the case of global methods, high accuracy necessitates a large number of elements in the basis set, leading to more extensive resource requirements. To tackle resource challenges, quantum algorithms are proposed.

Linear DE can be re-expressed as a system of linear equations using the finite difference method, and one can employ NISQ linear system algorithms to tackle the problem (see Sec. V.D.3). However NLDE defy this approach for large nonlinearities.

A canonical example of a NLDE appearing in quantum theory is the 1-D nonlinear Schrödinger equation  $\left[-\frac{1}{2} \frac{d^2}{dx^2} + V(x) + g|f(x)|^2\right]f(x) = Ef(x)$ . Here,  $E$  denotes energy,  $g$  quantifies nonlinearity, and  $V$  is the external potential. Recently, NISQ algorithms for NLDE have been proposed. Ref. (Lubasch *et al.*, 2020) use ancillary quantum registers and controlled-multiqubit operations to implement nonlinearities to simulate the nonlinear Schrodinger equation. Ref. (Haug and Bharti, 2020) propose the nonlinear quantum assisted simulator (NLQAS) to tackle NLDE without any controlled unitaries. Using NLQAS, they simulate this equation for 8 qubit system. NLDE have also been studied in (Gaitan, 2020) for fluid dynamics problems. Using differentiable quantum circuits, (Kyriienko *et al.*, 2020) have also proposed an interesting approach to solving NLDE via global methods.

#### E. Finance

The subject of finance deals with money and investments. The three typical subdivisions of this field are;

personal finance, corporate finance and public finance. Due to its practical relevance, comprehensive and rigorous investigations have been carried out to comprehend finance using techniques from mathematics. Some of the typical applications from finance which could potentially benefit from quantum technologies are portfolio optimization, modelling financial markets via differential equations and predicting market trends. In the following lines, we present one possible approach to classify mathematical problems in finance (Egger *et al.*, 2020a).

*Optimization based.* Many of the finance problems involve decision-making subject to certain constraints and consequently fall in the optimization framework. Some of the typical examples of optimization problems in finance are portfolio optimization, portfolio diversification and auctions. Concepts from convex optimization and combinatorial optimization turn out to be appropriate for the aforementioned class of problems. A considerable number of optimization centric problems can be converted to QUBO (see Eq. (53) in section Sec. III.A). Quantum algorithms for linear systems, convex optimization and QUBO have been employed with the hope for a possible advantage. In particular, the problems which can be reduced to QUBO have been investigated extensively using quantum annealing, VQE and QAOA (Boulant *et al.*, 2020; Hodson *et al.*, 2019).

*Simulation based.* The simulation problems in finance deal with simulating potential outcomes, typical examples being simulating the influence of volatility on risk or estimating asset values for pricing. Monte-Carlo simulations and algorithms for stochastic nonlinear differential equations are particularly beneficial for such tasks. Consequently, quantum Monte-Carlo algorithms and quantum algorithms for stochastic nonlinear differential equations have been investigated in the context of finance (Boulant *et al.*, 2020; Egger *et al.*, 2020a; Kubo *et al.*, 2020; Ramos-Calderer *et al.*, 2019).

*Machine learning based.* Some of the standard machine learning based problems in finance require predicting a future event based on historical data, pattern and anomaly detection, and classification of the end result in categories. Sample problems are fraud detection, algorithmic trading, risk assessment and credit scoring. One can potentially hope to apply quantum machine learning algorithms as the ones presented in Sec. V.B for such tasks.

The complete coverage of quantum finance is outside the scope of this review. We direct the reader to (Boulant *et al.*, 2020; Egger *et al.*, 2020a; Orus *et al.*, 2019) for a comprehensive treatment. We proceed to

discuss portfolio optimization and fraud detection; two canonical examples from finance. The algorithms that appear in this subsection are listed in Table V.

### 1. Portfolio optimization

The mathematical notion of portfolio optimization, was first proposed in (Markowitz, 1952). Given some capital and set of  $m$  assets, the goal is to find the best investment strategies under a set of constraints. Some of the typical constraints could be non-negativity of the investment or limit on total budget  $B$ . The objective function could be variance of the whole portfolio or overall risk. Let us represent the overall portfolio allocation by  $x \in \mathbb{R}^m$ , where  $x_k$  denotes the investment in the  $k$ -th asset. Suppose the return at time step  $t$  is  $c(t) \in \mathbb{R}^m$  with expected return  $\mu \in \mathbb{R}^m$  and covariance matrix  $\Sigma \in \mathbb{R}^{m \times m}$  given by

$$\mu(T) = \frac{1}{T} \sum_{t \in [T]} c(t), \quad (146)$$

$$\Sigma(T) = \frac{1}{T-1} \sum_{t \in [T]} (c(t) - \mu)(c(t) - \mu)^T. \quad (147)$$

With some additional equality constraints captured by  $A \in \mathbb{R}^{m \times m}$  and  $b \in \mathbb{R}^m$ , such as the total budget constraint  $\sum x_i = B$ , the portfolio optimization task corresponds to the following program,

$$\min_{x \in \mathbb{R}^m} x^T \Sigma x, \text{ s.t. } c^T x = \mu, \quad A^T x = b. \quad (148)$$

We can introduce Lagrange multipliers  $\eta$  and  $\theta$  for the equality constraints, and get the following Lagrangian corresponding to the aforementioned program,

$$\mathcal{L}(x, \eta, \theta) = \frac{1}{2} x^T \Sigma x + \eta (c^T x - \mu) + \theta (A^T x - b). \quad (149)$$

The solution to the portfolio optimization Eq. (148) reduces to solving the following linear system,

$$\begin{bmatrix} 0 & 0 & c^T \\ 0 & 0 & A^T \\ c & A & \Sigma \end{bmatrix} \begin{bmatrix} \eta \\ \theta \\ x \end{bmatrix} = \begin{bmatrix} \mu \\ b \\ 0 \end{bmatrix} \quad (150)$$

In (Rebentrost *et al.*, 2018) the HHL algorithm was proposed to implement the above approach with a quantum computer. One can use near-term linear system solvers for portfolio optimization tasks given assumptions about hardware capabilities and input model are satisfied (Huang *et al.*, 2019). A modified version of the program from Eq. (148) was recently tackled using VQE and QAOA (Egger *et al.*, 2020a). Specifically, the following optimization program was attempted,

$$\min_{x \in \{0,1\}^m} q x^T \Sigma x - \mu^T x, \text{ s.t. } 1^T x = B. \quad (151)$$

Here, the portfolio vector  $x \in \{0, 1\}^m$  is a vector of binary variables where  $x_i = 1$  means the  $i$ th asset was selected.  $x_i = 0$  means otherwise. Similar to the previous case,  $\mu$  and  $\Sigma$  denote expected return and covariance matrix respectively. The total budget has been denoted by  $B$  and  $q > 0$  captures the risk appetite of the decision maker. It is assumed that all assets have same price and total budget has to be spent. By adding the constraint as a penalty term  $(1^T x - B)^2$ , the authors convert optimization program Eq. (151) to a QUBO problem, which further reduces to an appropriate Hamiltonian ground state problem. They employ both VQE and QAOA hereafter. They also study portfolio diversification, by again converting the same to QUBO (Egger *et al.*, 2020a). Since problems which can be converted in QUBO can be a good fit for annealing based optimization, quantum annealing has been extensively applied for portfolio optimization (Cohen *et al.*, 2020; Marzec, 2016; Rosenberg *et al.*, 2016; Venturelli and Kondratyev, 2019). For details, refer to (Boulant *et al.*, 2020; Orus *et al.*, 2019).

## 2. Fraud detection

Machine learning can be utilised to detect financial frauds. A typical example could be somebody’s credit card getting blocked because higher than usual transaction was under process. Based on the historical data, often a fraudulent transaction can be caught. There are, however, false positives also. A typical example of false positive is somebody’s credit card transaction getting declined when trying it at a new store. Based on synthetic credit card transaction data in Ref. (Altman, 2019), variational quantum classification was performed in Ref. (Egger *et al.*, 2020a). The example problem in Ref. (Egger *et al.*, 2020a) contains dataset with 100 purchase transaction records with each transaction  $k$  marked fraudulent ( $y_k = -1$ ) or not fraudulent ( $y_k = 1$ ). The  $k$ -th transaction input vector  $x_k$  contains information about transaction amount, time, method and location. Variational quantum classification was used to predict the right label. A similar analysis using variational QBM was carried on in Ref. (Zoufal *et al.*, 2020).

## F. Other applications

In this subsection, we cover other applications for which NISQ algorithms can provide promising improvements. They are listed in Table VI.

### 1. Quantum foundations

One of the first experiments in digital quantum computers were the Bell nonlocality tests known as *Bell inequalities* (Brunner *et al.*, 2014). Those experiments

computed a type of Bell inequalities known as Mermin inequalities in up to five superconducting quantum qubits. The experiment consisted in preparing the GHZ state (Greenberger *et al.*, 1990), measure it in a particular basis state and obtain the expectation value of the Mermin operator (Alsina and Latorre, 2016).

In the context of VQA, the non-classicality in VQEs is examined using contextuality, which is a nonclassical feature of quantum theory (Amaral and Cunha, 2018). Using the notion of “strong contextuality”, (Kirby and Love, 2019) categorized VQE experiments into two categories: contextual and non-contextual. Such foundational works could be utilized to comprehend the possible sources of quantum advantage in NISQ algorithms. Using novel concepts from this field, contextual subspace VQE (CS-VQE) was recently proposed (Kirby *et al.*, 2020).

In another work, variational consistent history (VCH) algorithm was suggested to investigate foundational questions (Arrasmith *et al.*, 2019). The consistent history approach has been used to examine topics from quantum cosmology and quantum-classical transition. In the VCH algorithm, the quantum computer is used to compute the “decoherence functional”, which is challenging to calculate classically. The classical computer is employed to tune the history parameter so that the consistency is improved.

### 2. Quantum optimal control

Quantum optimal control is a topic of paramount importance in the pursuit to harness the potential of Near-Term quantum devices. For a given quantum control system and a cost function that measures the quality of control, it aims to find a control that can achieve optimal performance.

Some recent works have investigated quantum optimal control in the NISQ framework. Recent detailed perspective in this direction can be found in Ref. (Magann *et al.*, 2021). Ref. (Li *et al.*, 2017a) provides a hybrid quantum-classical approach to quantum optimal control. To remedy some of the difficulties of classical approaches to optimal control related to scaling of resources, (Dive *et al.*, 2018) proposed another NISQ framework. Experimental demonstration of quantum control for a 12-qubit system has also been carried on (Lu *et al.*, 2017). The aforementioned approaches, however, restrict their target states to be sparse matrices. For dense target states, (Policharla and Vinjanampathy, 2020) recently proposed a NISQ algorithm. Along with their algorithm, they also suggested a few algorithmic primitives to calculate overlap of quantum states and transition matrix elements.



### 3. Quantum metrology

Quantum metrology harnesses non-classical features of quantum theory for parameter estimation tasks. A canonical example could be estimating the parameter  $\phi$  of a unitary map under the action of Hamiltonian  $\hat{H}$ , given by  $\hat{\rho}(\phi) = e^{-i\hat{H}\phi}\hat{\rho}_0$  where the density matrix  $\hat{\rho}_0$  refers to the initial state of the system. The goal is to estimate  $\phi$  via measurements on  $\hat{\rho}(\phi)$ . The quantum Cramer-Rao bound provides a lower bound to the achievable precision,

$$(\Delta\phi)^2 \geq \frac{1}{nF_Q(\hat{\rho}(\phi))}. \quad (152)$$

Here,  $n$  represents number of samples,  $F_Q(\hat{\rho}(\phi))$  is quantum Fisher information and  $(\Delta\phi)^2$  is the variance in the estimation of  $\phi$ . In most of the experiments, the parameter of interest is either temperature or magnetic field.

Notice that the precision of the estimation procedure increases as the quantum Fisher information increases. Using it as a cost function, a few works have recently explored quantum metrology to prepare a better probe state in a VQA set-up (Beckey *et al.*, 2020; Kaubruegger *et al.*, 2019; Koczor *et al.*, 2020; Ma *et al.*, 2020). In addition, (Meyer *et al.*, 2020) provided toolbox for multi-parameter estimation.

### 4. Fidelity estimation

In Sec. II.A, we discussed how to use the fidelity as an objective function, a quantity which is useful to train some VQA algorithms. In addition, estimating the fidelity of a quantum state with respect to another state has a general interest in the context of quantum computing. For this reason, algorithms that can estimate this property may become useful in the NISQ era.

Given the density matrices of two quantum states  $\rho_1$  and  $\rho_2$ , their fidelity is given by

$$F(\rho_1, \rho_2) = \left( \text{Tr} \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2. \quad (153)$$

Due to the large dimensionality of the Hilbert spaces, computing fidelity can be challenging.

Recently, variational quantum fidelity estimation (VQFE) algorithm was proposed to tackle slightly modified version of the fidelity estimation task which works efficiently when  $\rho_1$  has low-rank. Ref. (Cerezo *et al.*, 2020a) provide lower and upper bounds on  $F(\rho_1, \rho_2)$  via VQFE. The algorithm calculate fidelity between  $\rho_1^n$ , which is a truncated version of  $\rho_1$  obtained by projecting  $\rho_1$  to subspace spanned by  $n$  largest eigenvalue eigenvectors of  $\rho_1$ . The bounds improve monotonically with  $n$  and is exact for  $n = \text{rank}(\rho_1)$ . The VQFE algorithm proceeds in three steps: *i*) a variational diagonalization of  $\rho_1$ ; *ii*) the matrix elements of  $\rho_2$  are computed in the eigenbasis of  $\rho_1$ ; and *iii*) using the matrix elements from *ii*), the fidelity is estimated.

### 5. Quantum error correction

The leading error correction schemes carry high resource overheads, which renders them impractical for near-term devices (Fowler *et al.*, 2012; Johnson *et al.*, 2017). Moreover, many of the schemes mandate knowledge of the underlying noise model (Fletcher *et al.*, 2008; Kosut and Lidar, 2009; Kosut *et al.*, 2008). For an encoding process  $\mathcal{E}$ , decoding process  $\mathcal{D}$  and noise model  $\mathcal{N}$ , the quality of a quantum error correction scheme can be characterized by how close  $\mathcal{D} \circ \mathcal{N} \circ \mathcal{E}$  is close to identity. The range of  $\mathcal{E}$  is called code space  $\mathcal{C}$ .

In (Johnson *et al.*, 2017), a variational error-correcting scheme i.e, quantum variational error corrector (QVECTOR) was proposed by defining an objective function over the code space  $\mathcal{C}$ . The authors employ two trainable parametric quantum circuits  $V(\mathbf{p})$  and  $W(\mathbf{q})$  for encoding and decoding respectively, with tunable parameter vectors  $\mathbf{p}$  and  $\mathbf{q}$ . For a given encoding-decoding pair, characterized by  $(\mathbf{p}, \mathbf{q})$ , the authors calculate a quantity called ‘‘average code fidelity’’ with respect to Haar distribution of states over the code space  $\mathcal{C}$ . The algorithm is model-free, i.e. no assumption of the noise model is involved. The goal of the QVECTOR algorithm is to maximize average code fidelity in a variational set-up.

In the context of variational quantum algorithms, error correction has also been explored in (Xu *et al.*, 2019b) where the target logical states are encoded as ground state of appropriate Hamiltonian. (Xu *et al.*, 2019b) employ imaginary time evolution to find the ground state. The authors implement there scheme for five and seven qubit codes. For a brief discussion on error correction and quantum fault-tolerance, refer to Sec. VIII.B.

### 6. Nuclear physics

The Standard Model of particle physics is the theory that describes the nature of the electromagnetic and nuclear interactions. Its current formulation consist of describing the forces as quantum fields, i.e. by using quantum field theory (QFT) formalism. Perturbative calculations of QFT provide with the dynamics of the physical processes at a given energy scale. However, in some cases as in quantum chromodynamics (QCD), perturbation theory can not be applied because the impossibility of observe a free quark or gluon (the fundamental particles affected by QCD interaction) due to confinement. For this reason, QCD calculations are obtained by means of numerical methods such as Monte Carlo simulations in a discretized version of QFT on a lattice structure (LQFT). The high computational cost of LQFT has motivated the study of using quantum computation or simulation to obtain the desired QCD predictions (Jo   *et al.*, 2019).

The Schwinger model describes the dynamics of the

quantum electromagnetic (QED) interaction in one spatial and temporal dimensions. It is used as a toy model to study QCD since it shows fermion confinement but it is simple enough to be solved analytically. The first experimental quantum simulations of this model were carried out in trapped ions (Hauke *et al.*, 2013) and later on a superconducting circuit quantum computer (Martinez *et al.*, 2016). A first proposal to use a quantum-classical algorithm to simulate this model was presented in (Klco *et al.*, 2018), where the quantum computer simulates the dynamics of the symmetry sectors suggested by a classical computation. In Ref. (Kokail *et al.*, 2019) a VQS is used in an analog setup to reduce the number of variational parameters and thus, reduce the computational cost of the algorithm. Their proposal is experimentally implemented in a trapped-ion analog simulator. A significant reduction of the computational cost of LQFT is proposed in (Avkhadiiev *et al.*, 2020) by using a VQA approach to compute the *optimized interpolating operators*, which are used to approximate the quantum wave functions, written in terms of the creation and annihilation operators, to reduce the Euclidean time evolution of the simulation. Adaptations of the UCC quantum chemistry ansatz, introduced in Sec. II.B.1, to study quantum-variationally QCD are presented in (Dumitrescu *et al.*, 2018; Liu and Xin, 2020), and for the study neutrino-nucleus scattering in (Roggero *et al.*, 2020).

## 7. Entanglement properties

Entanglement is a resource for numerous quantum information tasks. A bipartite quantum state  $\rho_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$  is called separable if it admits the form  $\rho_{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B$ , where  $p_i$  are non-negative and  $\sum_i p_i = 1$ . If a state is not separable, then it is called entangled. The problem of detecting whether a state is separable or entangled is known as the separability problem and has been shown to be NP-hard (Gurvits, 2003).

As mentioned in Sec. V.D.2, computing the Schmidt rank of  $\rho_{AB}$  gives a measure of the bipartite entanglement. Thus, those algorithms that tackle the SVD problem can also be used to extract entanglement properties (Bravo-Prieto *et al.*, 2020). In (Wang *et al.*, 2020a), authors propose a NISQ algorithm for the separability problem by providing a variational approach to employ *the positive map criterion*. This criterion establishes that the quantum state  $\rho_{AB}$  is separable if and only if for arbitrary quantum system  $R$  and arbitrary positive map  $\mathcal{N}_{B \rightarrow R}$  from  $B$  to  $R$ , we have  $\mathcal{N}_{B \rightarrow R}(\rho_{AB}) \geq 0$ . The authors start with a positive map and decompose it into a linear combination of NISQ implementable operations. These operations are executed on the target state, and the minimal eigenvalue of the final state is variationally estimated. The target state is deemed entangled if the optimized minimal eigenvalue is negative.

Exploring a similar strategy as the one presented in (Bravo-Prieto *et al.*, 2020), (Pérez-Salinas *et al.*, 2020b) propose a VQA to compute the tangle, a measure of tripartite-entanglement.

Variational algorithms have also been employed for extracting the entanglement spectrum of quantum systems in (Cerezo *et al.*, 2020b; LaRose *et al.*, 2019).

## VI. BENCHMARKING

One of the central questions at the intersection of software and hardware for NISQ devices is evaluating devices' performance and capabilities. This is where benchmarking concepts come in, to provide various metrics that attempt to measure different machines' capabilities and compare them across time and other devices. A benchmarking protocol can be characterized by its inherent assumptions, resource costs and the information gain. The goal is to build benchmarking protocols that make minimal and practical assumptions, have low resource costs, and have high information-gain.

Benchmarking protocols have been developed for NISQ as well as fault-tolerant devices. For a pedagogical summary, refer to (Eisert *et al.*, 2019). In this review, we focus on quantum benchmarking protocols for NISQ devices. Some of the leading NISQ benchmarking schemes are randomized benchmarking, quantum volume, cross-entropy benchmarking and application-based benchmarks.

### A. Randomized benchmarking

The most straightforward way of comparing devices is by simply counting qubits. To really compare different qubits, we must also have a sense of how many operations we can do with them before the noise arising from errors drowns out the signal. Randomized benchmarking (RB) is a convenient method for finding average error rates for quantum operations, in particular for single and two-qubit gates (Magesan *et al.*, 2011, 2012). RB is robust against state preparation and measurement (SPAM) errors and, unlike tomography, admits an efficient and practical implementation.

RB involves the following assumptions: *i)* for every gate, the incurred noise is independent of other Clifford gates; *ii)* the involved unitaries should constitute a 2-design (see Sec. IV.B) and should not be universal. In other words, no  $T$  gate is allowed; *iii)* during the experiment, there is no drifting in the noise processes; and *iv)* one can describe noise processes using completely positive trace-preserving (CPTP) maps.

A RB protocol starts by sampling a sequence of  $m$  Clifford gates (see Sec. IV.F.1). The sequence is applied to the initial state, followed by its inverse. Finally a

two-outcome POVM measurement is done to calculate the fidelity between initial state and the output state, followed by classical post-processing. The RB protocol discretizes time so that it is measured in the number of gates and it then averages over many sequences of each length  $m$ . More formally, a 4-step RB protocol consist of

- (1) Generate  $K_m$  sequences of  $m$  quantum operations  $C_{i_j}$  with  $i \in [1, m]$  and  $j \in [1, K_m]$ , where  $i$  indexes over the sequence of operations, and  $j$  over the statistical samples. These operations are randomly chosen from the Clifford group, and a  $m+1$ -th operation is chosen that cancels the first  $m$  operations such that the net operation is the identity. The operations can be chosen from the 2-, 4- or  $2^n$ -dimensional Clifford groups, depending on whether we are benchmarking single-, two- or  $n$ -qubit operations (McKay *et al.*, 2019). These operations come with some error, which is modeled with linear operators  $\Lambda_{i_j, j}$ , so that the full sequence of  $m$  operations is given by

$$S_{K_m} = \bigcirc_{j=1}^{m+1} (\Lambda_{i_j, j} \circ C_{i_j}) \quad (154)$$

Here,  $\circ$  denotes composition and  $\bigcirc$  represents composition of the terms defined with index  $j$ .

- (2) For each sequence we find the fidelity with the initial state by measuring  $\text{Tr}[E_\psi S_{K_m}(\rho(\psi))]$ , where  $\rho(\psi)$  is the initial state (with preparation errors) and  $E_\psi$  is a POVM measurement operator corresponding to the measurement including noise. Without noise, this would be the projector  $E_\psi = |\psi\rangle\langle\psi|$ .
- (3) Average over the  $K_m$  statistical samples to find the sequence fidelity  $F(m, \psi) = \text{Tr}[E_\psi S_m(\rho(\psi))]$  where  $S_m$  is the mean over the operations  $S_{K_m}$ .
- (4) Fit the data with the function

$$F_{fit}(m, \psi) = A_0 p^m + B_0, \quad (155)$$

where we have assumed the errors are independent of gate and time. This is not a fundamental assumption, but can be relaxed (Magesan *et al.*, 2011, 2012). The average gate error is here given by  $\epsilon_{RB} = 1 - p - (1 - p)/2^n$ , and the constants  $A_0$  and  $B_0$  absorb the SPAM errors.

The operations  $C_{i_j}$  are chosen from the Clifford group, because these are relatively easy to perform on quantum hardware, and because the final  $m+1$ -th operation that undoes the sequence can easily be pre-computed on a classical computer. Averaging over the Clifford group (or any other finite group) also has the property that even though the real noise-channel would be more complicated than the purely depolarizing one, the average over the group will still give rise to an exponential decay.

These gate errors extracted from randomized benchmarking can be used to compare the quality of quantum gates, and to estimate that an algorithm of depth  $\sim 1/\epsilon_{RB}$  gates can be run on the device before only statistical noise is output. The intuition behind the RB protocol is that a (purely) depolarizing channel will cause exponential decay of an excited state over time.

Simultaneous randomized benchmarking (SRB) has been proposed to acquire information about crosstalk and undesired coupling between the neighbouring qubits (Gambetta *et al.*, 2012). RB has also been extended for gatesets that do not form a Clifford group (Brown and Eastin, 2018; Carignan-Dugas *et al.*, 2015; Cross *et al.*, 2016; França and Hashagen, 2018; Gambetta *et al.*, 2012; Harper and Flammia, 2017; Hashagen *et al.*, 2018). In such cases, the expression for  $F_{fit}(m, \psi)$  does not follow equation Eq. (155) (Helsen *et al.*, 2019). Employing concepts from representation theory, an extension of RB has been proposed to extract the fidelity for a broad category of gatesets, including T-gate (Helsen *et al.*, 2019). A practically scalable protocol called cycle benchmarking was developed lately to characterize local and global errors for multi-qubit quantum computers (Erhard *et al.*, 2019).

## B. Quantum volume

To further refine the concept of the computational power of a quantum computer from just qubit count and gate-errors, the IBM Quantum team introduced the “quantum volume” (Cross *et al.*, 2019; Moll *et al.*, 2018). It is one of the widely accepted metrics for benchmarking NISQ-era quantum computers. As mentioned earlier, one can not rank quantum computers based on the number of qubits alone. Quantum volume gives a rough estimate of the number of effective qubits a quantum computer has based on their performance on the “heavy output generation problem”. The heavy output generation problem is related to the random circuit sampling task used in Google’s quantum supremacy experiment. Quantum volume treats the depth and width of a quantum circuit at the equal footing. Hence, its estimation depends on the largest square-sized circuit, which can successfully implement the heavy output generation problem. The quantum computer’s performance also depends on its software stack, for example, compiler, and thus quantum volume can increase with the improvements in the software stack.

The quantum volume benchmark can be thought analogous to the classical LINPACK benchmark (Dongarra, 1987). Like the LINPACK benchmark, it is architecture-agnostic and provides a single real number metric based on the quantum computer’s performance for a model problem, i.e., heavy output generation problem.

More formally, quantum volume can be defined in the following terms. Given an  $n$  qubit quantum computer

with the largest achievable *model circuit* depth  $d(m)$  for *model circuit* width  $m \in \{1, 2, \dots, n\}$  such that the probability of observing a *heavy output* for a random selection of model circuit is strictly greater than  $2/3$ , the quantum volume  $V_Q$  is defined as (Cross et al., 2019)

$$\log_2 V_Q = \arg \max_m (m, d(m)). \quad (156)$$

Intuitively speaking, quantum volume estimates the largest square random quantum circuit which the quantum computer can successfully implement the so-called heavy output generation problem. To conclude the discussion, it remains to describe the “model circuit” and the “heavy output generation problem”.

The model circuit with depth  $d$  and width  $m$  for estimating quantum volume is given by  $d$ -layered sequence  $U = U^{(d)}U^{(d-1)}\dots U^{(1)}$  where layer  $t$  consists of random permutations  $\pi_t \in S_m$  applied to qubit labels, followed by the tensor product of Haar-random two-qubit unitaries from  $SU(4)$ . If the model circuit width  $m$  is odd, one of the qubits is left idle in every layer. See Figure 6 for a pictorial description.

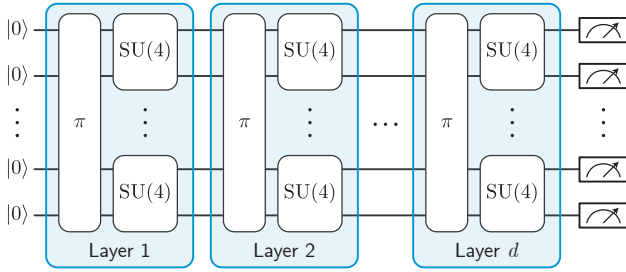


Figure 6 Model circuit for the quantum volume benchmark. Each layer consists of random permutations of qubit labels, followed by application of two-qubit haar-random unitaries. Inspired by (Cross et al., 2019).

Given a model circuit  $U$  with width  $m$ , the ideal output distribution over bit strings  $x \in \{0, 1\}^m$  is given by  $P_U(x) = |\langle x|U|0\rangle|^2$ .

One can arrange the probabilities for various bitstrings in ascending order in a set  $\mathbb{P} = \{p_0 \leq p_1 \leq \dots \leq p_{2^m-1}\}$ . The median of the set  $\mathbb{P}$  is given by  $p_{med} = \frac{p_{2^m-1} + p_{2^m-2}}{2}$ . The Heavy outputs are defined as  $H_U = \{x \in \{0, 1\}^m | P_U(x) > p_{med}\}$ . The goal of the heavy output problem is to sample a set of strings such that at least  $2/3$  are heavy output. For an ideal quantum circuit, the expected heavy output probability asymptotically tends to  $\sim 0.85$ . For a completely depolarized device, it is  $\sim 0.5$ .

For a completely depolarized device, it is  $\sim 0.5$ . On the target system, one implements  $\tilde{U}$  by using a circuit compiler with native gate set such that  $1 - F_{avg}(U, \tilde{U}) \leq \epsilon \leq 1$  for some approximation error  $\epsilon$ , where  $F_{avg}$  is average gate fidelity, as defined in Ref. (Horodecki et al., 1999). The role of circuit compiler is crucial in the aforementioned step. Suppose the observed distribution for the

implemented circuit  $\tilde{U}$  of the model circuit  $U$  is  $g_u(x)$ . The probability of sampling heavy output is given by

$$h_U = \sum_{x \in H_U} q_U(x). \quad (157)$$

For a randomly selected circuit of depth  $d$ , the probability of sampling a heavy output is given by

$$h_d = \int_U h_U dU. \quad (158)$$

The term  $d(m)$  in Eq. (156) is equal to the largest depth  $d$  for model circuit of width  $m \in \{1, 2, \dots, n\}$  such that  $h_d > \frac{2}{3}$ .

The quantum volume benchmark requires simulation of the model circuit’s heavy output generation problem on a classical computer. It hence is not a scalable method as the number of quantum volume increases. Moreover, the special treatment for square circuits is not entirely justified. Investigations are needed to devise other interesting benchmarks. A benchmark for rectangular circuits has also been proposed in the literature (Blume-Kohout and Young, 2020).

At the time of writing, Honeywell’s quantum computer has achieved  $\log_2 V_Q = 7$  (Honeywell, 2020), and IBM has a quantum computer with  $\log_2 V_Q = 6$  on a 27-qubit system (IBM, 2020).

### C. Cross-entropy benchmarking

The linear cross-entropy benchmarking is a statistical test used by Google in their quantum supremacy experiment (Arute et al., 2019; Neill et al., 2018). It measures how often high-probability bitstrings are sampled in an experimental scenario. Suppose we perform a sampling task and obtain bitstrings  $\{x_j\}_j$  via measurement on a given  $m$ -qubit circuit  $\mathcal{C}_E$ . The linear cross-entropy benchmarking fidelity is given by

$$\mathcal{F}_{XEB} = 2^m \langle P(x_j) \rangle_j - 1. \quad (159)$$

Here, the average  $\langle \cdot \rangle_j$  is over the experimentally observed bitstrings  $\{x_j\}_j$  and  $P(x_j)$  denotes the probability of observing bitstring  $x_j$  for the ideal circuit version of  $\mathcal{C}_E$ . In other words,  $P(x_j)$  denotes the ideal probability of the generated sample  $x_j$ . Since one can not have an ideal circuit in practice,  $P(x_j)$  are calculated using a classical computer simulation of the ideal circuit.  $\mathcal{F}_{XEB}$  compares how often a bitstring  $x_j$  is observed experimentally with its classically simulated ideal probability. For the ideal case,  $\mathcal{F}_{XEB}$  approaches unity for a large number of qubits. On the other hand, it is equal to zero for uniform distribution. As the circuit’s nose grows,  $\mathcal{F}_{XEB}$  decreases and approaches zero. Since the probabilities  $P(x_j)$  are calculated via classical simulation; it renders the computation of  $\mathcal{F}_{XEB}$  intractable in the supremacy regime. The



classical hardness of spoofing linear cross-entropy benchmarking was studied by Aranson and Gunn (Aaronson and Gunn, 2019), where they suggested the absence of any efficient classical algorithm for the aforementioned task.

#### D. Application benchmarks

While hardware benchmarks, such as randomized benchmarking or quantum volume, provide valuable insight into the performance of quantum devices, they may not well represent or predict the performance of VQAs which employ structured circuits. Application benchmarks were developed to complement hardware benchmarks and provide a more complete picture of both the performance and (near-term) utility of quantum devices. These benchmarks consist executing experimental demonstrations of VQA instances that can be compared to classically computed exact results. Examples of application benchmarks can be found in Refs. (Arute *et al.*, 2020a; Benedetti *et al.*, 2019a; Dallaire-Demers and Killoran, 2018; Karamlou *et al.*, 2020). In particular, (Arute *et al.*, 2020a) demonstrated the largest VQE experiment to date that also achieved chemical accuracy.

As a specific example of an application benchmark, quantum circuits that diagonalize spin Hamiltonians have been proposed in recent years (Cervera-Liarta, 2018; Schmoll and Orús, 2017; Verstraete *et al.*, 2009). By comparing the results obtained from the quantum device with the analytical solution, one can discern the performance of the computation for a specific purpose experiment. Small experiments have shown that gate fidelities and decoherence times alone do not provide a complete picture of the noise model (Cervera-Liarta, 2018).

In that direction, authors of (Dallaire-Demers *et al.*, 2020) proposed a figure-of-merit called the *effective fermionic length* to quantify the performance of a NISQ device in which the application-at-hand is estimating the energy density of the one-dimensional Fermi-Hubbard model over increasing chain lengths. Theoretically, as the chain length increases, the energy density should approach the infinite chain limit. In practice, the NISQ device will accrue some level of noise and decoherence, which will cause the computed energy density to diverge past some chain length. The maximum chain length after which noise and decoherence start degrading the algorithm performance reveals the “limit” of the quantum device in carrying out related algorithms. Ref. (Dallaire-Demers *et al.*, 2020) abstracts this idea to redefine an application benchmark as a way to systematically test the limits of a quantum processor using exactly solvable VQA instances that can also be scaled to larger system sizes (e.g. chain length in Ref. (Dallaire-Demers *et al.*, 2020) or number of preprocessing steps in Ref. (Karamlou *et al.*, 2020)).

Generative models such as the QCBM (see Sec. V.B.2) can serve as benchmarks for NISQ devices (Hamilton *et al.*, 2019; Leyton-Ortega *et al.*, 2019). Here, the measurement output of hardware efficient variational ansätze are used to represent different types of distributions and study the effect of noise and hardware limitations on the result.

In addition to VQAs, one can analyze more fundamental benchmarks, such as the ability of NISQ devices to violate local-realism by means of Mermin inequalities (Alsina and Latorre, 2016) or the entanglement power of the devices by trying to construct maximal entangled states (Cervera-Liarta *et al.*, 2019; Wang *et al.*, 2018a).

### VII. QUANTUM SOFTWARE TOOLS

A quantum computer is a hybrid device composed by a quantum hardware and a classical software that controls it by sending a list of instructions and processing the results of the computation. This hybrid nature is accentuated in the NISQ era, as explained in the current review. Thus, the classical subroutines are part of the core in state-of-the-art NISQ algorithms and a language to communicate with the quantum device is a bare necessity. On top of that, almost all progress in quantum algorithms is made in a virtual environment. Good quantum simulators are also essential to perform proof-of-concept simulations, before or until the algorithm is applicable on real devices.

Figure 7 represents diagrammatically the typical workflow of a NISQ algorithm. The individual parts of the problem, such as the objective function to optimize, the quantum circuit design or the initialization parameters, are translated into quantum circuits by a classical pre-computation step. The syntax of this language includes the quantum gates, qubit initialization, objective function definition, etc. The theoretical circuit is then compiled to fulfill experimental limitations such as qubit connectivity, native quantum gate set or circuit depth. To accomplish this task, compilers that allow for circuit simplification (see Sec. IV.F), error mitigation libraries, and noise models (for simulation purposes) might be useful. After this pre-processing step, the algorithm is ready to enter into the quantum-classical loop. The quantum circuit can be run in a quantum simulator or real hardware. In the latter case, an assembly language (Cross *et al.*, 2017; Khammassi *et al.*, 2018; Killoran *et al.*, 2019; Smith *et al.*, 2016) will translate the quantum circuit into a set of instructions for the device. After the qubits are measured, the result is post-processed. Either the algorithm finishes or the result is sent to the classical optimizer that computes the next loop variables (e.g. for VQA).

We define a *quantum software package* as a library or a set of libraries written in a classical programming language (e.g. python or C++) that allows writing a quan-

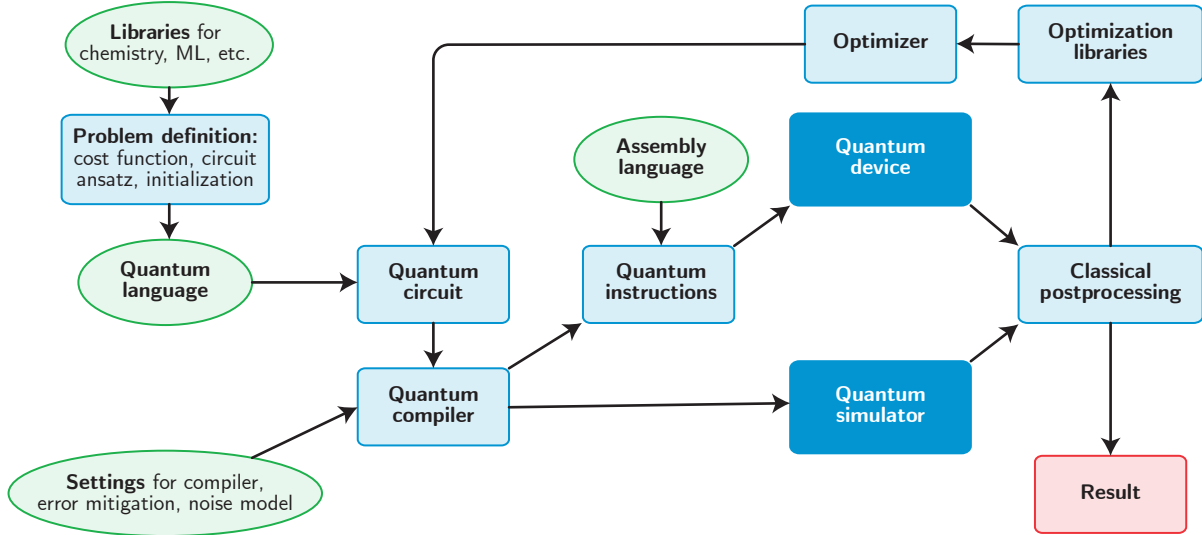


Figure 7 Schematic representation of a standard NISQ programming workflow (color online). Green circular boxes represent the libraries and languages used for designing, optimizing and running a quantum algorithm in a real quantum device or in a simulator. External libraries can be used to define the problem or to improve the performance of the algorithm (through simplifying the circuit or using error-mitigation techniques). An assembly language will be needed to translate the theoretical algorithm to a set of physical operations on the quantum hardware. Classical postprocessing is necessary to manipulate the result of the computation and to either obtain the final result or send the provisional one to a classical optimizer (VQA).

tum program composed by qubit states, quantum gates and measurement operations. There can be many differences between these packages. For instance, we distinguish between those languages that also contain a standalone quantum simulator from those that use external ones. The possibility of running the quantum algorithm in a real device is also a significant difference.

In some cases, these languages are open-source and can be used with a simulator as a backend. The proliferation of all these languages, simulators and devices has also created a necessity for some multi-platform languages. These are those that can use other quantum software packages as a backend, reducing the programming efforts substantially by unifying the language syntax. Finally, some of these languages include built-in sub-libraries suited for particular applications, from chemistry to QML, or particular well-known algorithms such as VQE or QAOA.

Due to the broad applications of NISQ algorithms, specific libraries used in other fields beyond quantum computation can also be required. These include quantum chemistry libraries (McClellan *et al.*, 2020a; Smith *et al.*, 2020), external quantum compilers (Amy and Gheorghiu, 2020; Bichsel *et al.*, 2020; JavadiAbhari *et al.*, 2014; Kissinger and van de Wetering, 2020a; QCA Lab, 2020; Wille *et al.*, 2020), machine learning packages (Abadi *et al.*, 2016; Broughton *et al.*, 2020) and error mitigation package (LaRose *et al.*, 2020).

We provide a list of some open-source libraries suited for NISQ computation in Tab. VII. This list represents

just a snapshot of the state-of-the-art of the quantum software ecosystem as new tools are being developed and some projects are being abandoned. An updated list of quantum software resources can be found in Ref. (QOSF, 2020; Quantum Computing Reports, 2020) and a detailed comparison analysis between some of these languages in Ref. (Garhwal *et al.*, 2019; Gay, 2006; Heim *et al.*, 2020; Nguyen *et al.*, 2020). Table VIII listed some of the external libraries, most of them integrated in the quantum software packages.

## VIII. OUTLOOK

In the last decade, quantum computing has experienced notable progress in applications, experimental demonstrations, and theoretical results. The number of papers in quantum computation, particularly in NISQ applications, is growing almost exponentially. Several reasons explain this community drive, one of those being tremendous improvements in quantum hardware. However, experimental advances alone do not fully explain why nowadays many universities and research centres have groups working on quantum computation. The globalization and democratization of quantum computing have at least partially been possible due to open source policies emerging in academia and the private sector.

Quantum computing is a relatively young field in science and, as such, there is plenty of room for pioneer-

ing research and discoveries. This fact, together with the theoretical, practical, and experimental challenges (several of them covered in this review), has strengthened the motivation for an open-source strategy in the field. Nowadays, many universities and research centers subscribe to an open-source policy that pushes towards the opening and publications of all computational tools, data, and programs used in their research. These policies have proved valuable for rapid scientific development as well as for democratizing community knowledge. This way of thinking has percolated through academia walls and it has been introduced into several private companies, not just for its own advantage, but also because it facilitates the continuous healthy flow of good quantum computing researchers to themselves (and, in some cases, resulting in more startups). Consequently, there is a rich open quantum computing ecosystem composed of universities, institutes, big corporations, startups, and uncountable individual enthusiasts.

Another product of the symbiosis between academia and the private sector is *cloud quantum computing*. Companies are offering access to their hardware remotely, in some cases at zero-cost for their small prototypes and simulators. On the one hand, scientists and quantum computing enthusiasts around the world have the opportunity to experience real quantum devices from their homes. On the other hand, this increases the chances of finding real-world applications in quantum computation and solving the current challenges of this field. The proliferation of open-source quantum computing languages, simulators, and tools (detailed in Sec. VII) have burgeoned many user communities. Various international initiatives have been set up to attract quantum computing talent, and the private sector’s involvement is ramping up. Several non-profit initiatives are also encouraging the use and development of these tools (Fund, 2020; QOSF, 2020).

Experimental realizations of quantum computation, although in the early stages, have interested many communities in this quantum information subfield. Healthy competition has also arisen between the classical and quantum computing branches. Classical computational scientists have put their efforts into moving the quantum advantage frontier further, raising the bar to claim that a quantum algorithm shows a significant speed-up. Along that direction, an off-shoot is an effort in “*dequantization*”, first exhibited in the case of recommendation systems, to devise quantum-inspired classical algorithms that are nearly as fast as their quantum counterpart (Tang, 2019). Such attempts at “dequantization” have eliminated examples of speed-up for some problems in linear algebra. So far, dequantized machine learning algorithms have been developed for recommendation systems (Tang, 2019), principal component analysis and supervised clustering (Tang, 2018), stochastic regression (Gilyén *et al.*, 2018) and low-rank linear systems

(Arrazola *et al.*, 2019; Chia *et al.*, 2018).

Since the NISQ devices are inherently noisy, analysis similar to (Napp *et al.*, 2019; Zhou *et al.*, 2020b; Zlokapa *et al.*, 2020) will be required to find out how much noise a NISQ algorithm can endure until its classical simulation becomes efficient. This is crucial in order to understand the boundary where quantum computers provide an advantage. Investigating the potential of NISQ algorithms using ideas from quantum foundations such as contextuality and entanglement are helpful in the respect (Bharti *et al.*, 2020b; Deutsch, 2020). More theoretical results similar to (Biamonte, 2019; Bravyi *et al.*, 2020a; Farhi and Harrow, 2016; Lloyd, 2018) may also prove valuable. It is also imperative to develop strategies that help us bypass complicated measurements involving controlled multi-qubit unitaries (Mitarai and Fujii, 2019). For machine learning tasks, ideas similar to (Harrow, 2020) would be valuable.

Another fascinating frontier that needs to be investigated in the next few years, we believe, is quantum and classical certification schemes for quantum devices and quantum computation (Eisert *et al.*, 2020). The intractability of quantum computation by classical devices poses the challenge to verify the correct functioning of the quantum devices as well as the correctness of the final output (Eisert *et al.*, 2020). The existence of multiple quantum computing platforms requires new methodologies and figures of merit to benchmark and compare these devices. Other works are being proposed in that direction (Kottmann *et al.*, 2020c; Kyaw *et al.*, 2020b), as well as the development of benchmarking measures discussed in Sec. VI. Ideas from complexity theory (Mahadev, 2018; Metger and Vidick, 2020) and quantum foundations (Bharti *et al.*, 2019b,c) could be valuable in this direction.

At the moment of documenting this review, there is no known demonstration of industrially relevant quantum advantage. Quantum computing is still in its early days, and so far a useful quantum computer is missing. The potential of NISQ devices is not fully understood, and a lot of rigorous research is required to release the power of the early quantum computers. However, a number of experiments overcoming classical computational resources have been performed and, many theoretical and practical tools are being used and developed, as explained in Sec. IV.

## A. NISQ goals

We expect experimental pursuit in the NISQ era would focus on the design of quantum hardware with a larger number of qubits and gates with lower error rates for executing deeper circuits. Along the way, the objective is to demonstrate quantum advantage for practical use cases. If the NISQ paradigm is not powerful enough to exhibit

any quantum advantage, theoretical pursuits would be required to understand its limitations. The prime direction of the NISQ and near-term era is to engineer the best possible solution with the limited quantum resources available. The tools and techniques invented during this period could be valuable in the *fault-tolerant* era as well.

To conduct a successful demonstration of quantum advantage, the right blend of following three crucial components is required:

- (1) *Hardware development:* The design of quantum computers with more qubits, lesser error rates, longer coherence times, and more connectivity between the qubits will be one of the top priorities in the NISQ era. A way to scale up the number of qubits present in a quantum platform is to design a novel qubit which has built-in autonomous quantum error correction down to the hardware level (Chamberland *et al.*, 2020; Paz and Zurek, 1998) or protected novel qubit (Douçot and Ioffe, 2012; Kyaw, 2019; Nataf and Ciuti, 2011) which is robust against specific noises in the hardware. As a quantum processor size grows, there is a tremendous need to store quantum information (Kyaw *et al.*, 2014, 2015a,b) during quantum information processing.

Even miniaturizing microwave circulator onto the superconducting chip (Chapman *et al.*, 2017; Mahoney *et al.*, 2017) can be seen as a mean to scale up the quantum platform though it has nothing to do with novel qubit design. The circulator is an invaluable device for noise management and signal routing for superconducting circuit experiments. Currently, commercial ferrite junction circulators are bulky whose physical dimension is close to the microwave wavelength applied.

In short, intensive research in new qubits developments, quantum optimal control and material discovery will be indispensable for both universal programmable quantum computers or special-purpose ones.

- (2) *Algorithm design:* To harness the potential of noisy but powerful quantum devices, we expect breakthroughs on the algorithm frontier. Algorithms with realistic assumptions as the ones mentioned in Sec. IV.F regarding device capabilities will be favored. To lessen the effect of noise, progress towards the design of error mitigation algorithms is expected. Efforts have to be made to develop algorithms that harness the problem's structure in the best possible manner and map it to the given hardware in efficient ways, such as in Sec. III.D. Variational quantum algorithms with better expressibility and trainability will also be helpful.

- (3) *Application problem:* We have discussed the existing applications of NISQ devices in many areas in Sec. V. Collaborations between experts with domain knowledge from these fields and quantum algorithm researchers will be required more and more to develop the field and integrate quantum computation into industrial workflows. New collaborations might reveal difficult problems for classical computers that are well suited for NISQ era quantum devices. It is not clear yet which applications will be the first ones to witness quantum advantage, though there is plenty of speculation and opinions.

## B. Long-term goal: fault-tolerant quantum computing

Noise is regarded as one of the most prominent threats to a quantum computer's practical realization. In 1995, Peter Shor established that by encoding quantum information redundantly using extra qubits, one could circumvent the effect of noise (Shor, 1995). The quantum information is spread over multiple physical qubits to generate a logical qubit (Calderbank and Shor, 1996; Gottesman, 1997; Knill and Laflamme, 1997; Shor, 1995). Most of the transformative algorithms such as Shor's factoring algorithm, Grover search algorithm, and HHL require error-corrected qubits for their execution. Soon after Shor's error-correcting code, many others were developed. Some of the famous error-correcting codes are stabilizer codes and topological error-correcting code (Fowler *et al.*, 2012; Terhal, 2015). While the stabilizer code utilizes extra qubits to protect the logical qubit, topological codes employ a set of qubits positioned on a surface, such as a torus, in a lattice structure.

Over the years, quantum error correction has evolved as a subfield of quantum computation and has transformed from a theoretical pursuit to a practical possibility. The process of detecting and correcting errors can be, itself, prone to noise. Thus error correction alone does not guarantee the prospect of storing or processing quantum information for an arbitrarily long period. The aforesaid issue can be, however, tackled utilizing *Quantum Fault-Tolerant threshold theorem*. Informally speaking, it is possible to execute arbitrarily large quantum computation by arbitrarily suppressing the quantum error rate, given the noise in the individual quantum gates are below a certain threshold (Aharonov and Ben-Or, 2008). If one wants to simulate an ideal circuit of size  $N$ , the size of the noisy quantum circuit for fault-tolerant quantum computation scales  $O(N(\log N)^c)$ , for some constant  $c$ , given the noisy circuit is subjected to stochastic noise strength  $p < p_c$  for some noise threshold  $p_c$  (Terhal, 2015). This theorem rises some practically relevant questions such as *i)* How high is  $p_c$ ; *ii)* what is the value of the constant  $c$ ; and *iii)* what is the value of the multiplicative constant in  $O(\cdot)$ . These questions



determine the practicality of any fault-tolerant quantum computation scheme (Terhal, 2015).

Looking forward, lowering the noise level will be a critical challenge. Though the problem is demanding, significant progress has been made recently at the algorithmic as well as hardware frontier (Campagne-Ibarcq *et al.*, 2020; Lidar and Brun, 2013; Noh and Chamberland, 2020; Terhal, 2015). Recently, Monroe and Brown’s groups have confirmed the first-ever fault-tolerant operation on a logical qubit (Egan *et al.*, 2020).

We are at an exciting juncture in the history of computing. Completely new kinds of computers that were once only figments of imagination are rapidly becoming reality. The NISQ era offers fantastic opportunities to current and future researchers to explore the theoretical limits of the devices and discover practical and exciting applications in the near-term. Theoretical investigations and experimental challenges will help us comprehend quantum devices’ power and build better algorithms. The success of the field lies in the hands of the researchers and practitioners of the area, so we encourage everyone with interest to join the effort.

## ACKNOWLEDGEMENTS

A.A.-G. acknowledges the generous support from Google, Inc. in the form of a Google Focused Award. This work was supported by the U.S. Department of Energy under Award No. DESC0019374 and the U.S. Office of Naval Research (ONS506661). A.A.-G. also acknowledges support from the Canada Industrial Research Chairs Program and the Canada 150 Research Chairs Program. T.H. is supported by a Samsung GRP project and the UK Hub in Quantum Computing and Simulation, part of the UK National Quantum Technologies Programme with funding from UKRI EPSRC grant EP/T001062/1. L.-C.K and K.B acknowledge the financial support from the National Research Foundation and the Ministry of Education, Singapore. K.B. thanks Naresh Boddu and Rahul Jain for interesting discussions.

## REFERENCES

- Aaronson, S., and A. Arkhipov (2011), [Proceedings of the forty-third annual ACM symposium on Theory of computing](#) , 333.
- Aaronson, S., and D. Gottesman (2004), *Phys. Rev. A* **70**, 052328.
- Aaronson, S., and S. Gunn (2019), [arXiv:1910.12085](#) .
- Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.* (2016), [12th {USENIX} symposium on operating systems design and implementation \({OSDI} 16\)](#) , 265.
- Ackley, D. H., G. E. Hinton, and T. J. Sejnowski (1985), *Cogn. Sci.* **9** (1), 147 .
- Adams, C. S., J. D. Pritchard, and J. P. Shaffer (2019), *J. Phys. B* **53** (1), 012002.
- Aharonov, D., and M. Ben-Or (2008), *SIAM J. Comput.* .
- Aharonov, D., W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev (2008), *SIAM review* **50** (4), 755.
- Aïmeur, E., G. Brassard, and S. Gambs (2013), *Mach. Learn.* **90** (2), 261.
- Akshay, V., H. Philathong, M. E. S. Morales, and J. D. Biamonte (2020a), *Phys. Rev. Lett.* **124**, 090504.
- Akshay, V., H. Philathong, I. Zacharov, and J. Biamonte (2020b), [arXiv:2007.09148](#) .
- Albarrán-Arriagada, F., J. Retamal, E. Solano, and L. Lamata (2020), *Mach. Learn.: Sci. Technol.* **1** (1), 015002.
- Albash, T., and D. A. Lidar (2018), *Rev. Mod. Phys.* **90** (1), 015002.
- Albash, T., T. F. Rønnow, M. Troyer, and D. A. Lidar (2015), *The Eur. Phys. J. Spec. Top.* **224** (1), 111.
- Alcazar, J., V. Leyton-Ortega, and A. Perdomo-Ortiz (2020), *Mach. Learn.: Sci. Technol.* **1**, 035003.
- Aleiner, I., F. Arute, K. Arya, J. Atalaya, R. Babbush, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, A. Bourassa, *et al.* (2020), [arXiv:2012.00921](#) .
- Alsina, D., and J. I. Latorre (2016), *Phys. Rev. A* **94** (1), 012314.
- Altman, E. R. (2019), [arXiv:1910.03033](#) .
- Amaral, B., and M. T. Cunha (2018), *On graph approaches to contextuality and their role in quantum theory* (Springer).
- Amari, S.-I. (1998), *Neural computation* **10** (2), 251.
- Ambainis, A. (2012), “Variable time amplitude amplification and quantum algorithms for linear algebra problems,” <https://t.ly/30Uz>.
- Amico, L., M. Boshier, G. Birkel, A. Minguzzi, C. Miniatura, L.-C. Kwek, D. Aghamalyan, V. Ahufinger, N. Andrei, A. Arnold, *et al.* (2020), [arXiv:2008.04439](#) .
- Amin, M. H., E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko (2018), *Phys. Rev. X* **8** (2), 10.1103/physrevx.8.021050.
- Amy, M., and V. Gheorghiu (2020), *Quantum Sci. Technol.* .
- Amy, M., D. Maslov, M. Mosca, and M. Roetteler (2013), *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **32** (6), 818.
- Amy, M., and M. Mosca (2019), *IEEE Transactions on Information Theory* **65** (8), 4771.
- Anand, A., M. Degroote, and A. Aspuru-Guzik (2020a), [arXiv:2012.00101](#) .
- Anand, A., J. Romero, M. Degroote, and A. Aspuru-Guzik (2020b), [arXiv:2006.01976](#) .
- Anderson, D. G. (1965), *J. ACM* **12** (4), 547.
- Anshuetz, E., J. Olson, A. Aspuru-Guzik, and Y. Cao (2019), *International Workshop on Quantum Technology and Optimization Problems* , 74.
- Argüello-Luengo, J., A. González-Tudela, T. Shi, P. Zoller, and J. I. Cirac (2019), *Nature* **574** (7777), 215.
- Arora, S., and B. Barak (2009), *Computational complexity: a modern approach* (Cambridge University Press).
- Arrasmith, A., M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles (2020a), [arXiv:2011.12245](#) .
- Arrasmith, A., L. Cincio, R. D. Somma, and P. J. Coles (2020b), [arXiv:2004.06252](#) .
- Arrasmith, A., L. Cincio, A. T. Sornborger, W. H. Zurek, and P. J. Coles (2019), *Nat. commun.* **10** (1), 1.

- Arrazola, J. M., and T. R. Bromley (2018), *Phys. Rev. Lett.* **121** (3), 030503.
- Arrazola, J. M., T. R. Bromley, and P. Rebentrost (2018), *Phys. Rev. A* **98** (1), 012322.
- Arrazola, J. M., A. Delgado, B. R. Bardhan, and S. Lloyd (2019), [arXiv:1905.10415](#).
- Arunachalam, S., and R. de Wolf (2017), [arXiv:1701.06806](#).
- Arute, F., K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.* (2019), *Nature* **574** (7779), 505.
- Arute, F., K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, *et al.* (2020a), *Science* **369** (6507), 1084.
- Arute, F., K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, *et al.* (2020b), [arXiv:2004.04197](#).
- Aspuru-Guzik, A., A. D. Dutoi, P. J. Love, and M. Head-Gordon (2005), *Science* **309** (5741), 1704.
- Avkhadiiev, A., P. Shanahan, and R. Young (2020), *Phys. Rev. Lett.* **124** (8), 080501.
- Ayanzadeh, R., S. Mousavi, M. Halem, and T. Finin (2019), [arXiv:1901.00088](#).
- Babbush, R., P. J. Love, and A. Aspuru-Guzik (2014), *Sci. Rep.* **4** (1), 6603.
- Ball, R. (2005), *Phys. Rev. Lett.* **95** (17), 176407.
- Banchi, L., and G. E. Crooks (2020), [arXiv:2005.10299](#).
- Banchi, L., M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola (2020a), *Sci. Adv.* **6** (23), eaax1950.
- Banchi, L., N. Quesada, and J. M. Arrazola (2020b), [arXiv:2004.04770](#).
- Bang, J., and S. Yoo (2014), *J. Korean Phys. Soc* **65** (12), 2001.
- Barenco, A., C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter (1995), *Phys. Rev. A* **52**, 3457.
- Barison, S., D. E. Galli, and M. Motta (2020), [arXiv:2011.08137](#).
- Barkoutsos, P. K., G. Nannicini, A. Robert, I. Tavernelli, and S. Woerner (2020), *Quantum* **4**, 256.
- Barron, G. S., B. T. Gard, O. J. Altman, N. J. Mayhall, E. Barnes, and S. E. Economou (2020), [arXiv:2003.00171](#).
- Bastidas, V., T. Haug, C. Gravel, L.-C. Kwek, W. Munro, and K. Nemoto (2020), [arXiv:2009.00823](#).
- Bastidas, V. M., T. H. Kyaw, J. Tangpanitanon, G. Romero, L.-C. Kwek, and D. G. Angelakis (2018), *New J. Phys.* **20** (9), 093004.
- Batista, C. D., and G. Ortiz (2004), *Adv. Phys.* **53** (1), 1.
- de Beaudrap, N., X. Bian, and Q. Wang (2019), [arXiv:1911.09039](#).
- Beckey, J. L., M. Cerezo, A. Sone, and P. J. Coles (2020), [arXiv:2010.10488](#).
- Benedetti, M., D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, and A. Perdomo-Ortiz (2019a), *npj Quantum Inf.* **5** (1), 1.
- Benedetti, M., E. Lloyd, S. Sack, and M. Fiorentini (2019b), *Quantum Sci. Technol.* **4** (4), 043001.
- Benedetti, M., J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz (2017), *Phys. Rev. X* **7** (4), 041052.
- Benedetti, M., J. Realpe-Gómez, and A. Perdomo-Ortiz (2018), *Quantum Sci. Technol.* **3** (3), 034007.
- Bengtsson, A., P. Vikstål, C. Warren, M. Svensson, X. Gu, A. F. Kockum, P. Krantz, C. Križan, D. Shiri, I.-M. Svensson, G. Tancredi, G. Johansson, P. Delsing, G. Ferrini, and J. Bylander (2020), *Phys. Rev. Appl.* **14**, 034010.
- Benioff, P. (1980), *J. Stat. Phys.* **22** (5), 563.
- Bergholm, V., J. Izaac, M. Schuld, C. Gogolin, M. S. Alam, S. Ahmed, J. M. Arrazola, C. Blank, A. Delgado, S. Janghigiri, *et al.* (2018), [arXiv:1811.04968](#).
- Bharti, K. (2020), [arXiv:2009.11001](#).
- Bharti, K., and T. Haug (2020a), [arXiv:2010.05638](#).
- Bharti, K., and T. Haug (2020b), [arXiv:2011.06911](#).
- Bharti, K., T. Haug, V. Vedral, and L.-C. Kwek (2019a), [arXiv:1912.10783](#).
- Bharti, K., T. Haug, V. Vedral, and L.-C. Kwek (2020a), [arXiv:2003.11224](#).
- Bharti, K., T. Haug, V. Vedral, and L.-C. Kwek (2020b), *AVS Quantum Sci.* **2**, 034101.
- Bharti, K., M. Ray, A. Varvitsiotis, A. Cabello, and L.-C. Kwek (2019b), [arXiv:1911.09448](#).
- Bharti, K., M. Ray, A. Varvitsiotis, N. A. Warsi, A. Cabello, and L.-C. Kwek (2019c), *Phys. Rev. Lett.* **122**, 250403.
- Biamonte, J. (2019), [arXiv:1903.04500](#).
- Biamonte, J., P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd (2017), *Nature* **549** (7671), 195.
- Bichsel, B., M. Baader, T. Gehr, and M. Vechev (2020), *Proceedings PLDI 2020*, 286.
- Björklund, A., B. Gupt, and N. Quesada (2019), *ACM J. Exp. Algorithmics* **24**.
- Blauberger, M., and R. De Visser (2008), *J. Phys. A Math. Theor.* **41** (39), 395307.
- Blatt, R., and C. F. Roos (2012), *Nat. Phys.* **8** (4), 277.
- Bloch, I., J. Dalibard, and S. Nascimbene (2012), *Nat. Phys.* **8** (4), 267.
- Blume-Kohout, R., and K. C. Young (2020), *Quantum* **4**, 362.
- Bondarenko, D., and P. Feldmann (2020), *Phys. Rev. Lett.* **124**, 130502.
- Bonet-Monroig, X., R. Babbush, and T. E. O'Brien (2020), *Phys. Rev. X* **10** (3), 031064.
- Bonet-Monroig, X., R. Sagastizabal, M. Singh, and T. O'Brien (2018), *Phys. Rev. A* **98** (6), 062339.
- Botea, A., A. Kishimoto, and R. Marinescu (2018), SOCS.
- Boulund, A., W. van Dam, H. Joorati, I. Kerenidis, and A. Prakash (2020), [arXiv:2011.06492](#).
- Brady, L. T., C. L. Baldwin, A. Bapat, Y. Kharkov, and A. V. Gorshkov (2020), [arXiv:2003.08952](#).
- Brady, L. T., and W. van Dam (2016), *Phys. Rev. A* **94** (3), 032309.
- Bravo-Prieto, C. (2020), [arXiv:2010.06599](#).
- Bravo-Prieto, C., D. García-Martín, and J. I. Latorre (2020), *Phys. Rev. A* **101** (6), 062310.
- Bravo-Prieto, C., R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles (2019), [arXiv:1909.05820](#).
- Bravyi, S., D. P. DiVincenzo, D. Loss, and B. M. Terhal (2008), *Phys. Rev. Lett.* **101** (7), 070503.
- Bravyi, S., D. Gosset, R. Koenig, and M. Tomamichel (2020a), *Nat. Phys.* **16** (10), 1040.
- Bravyi, S., A. Kliesch, R. Koenig, and E. Tang (2019), [arXiv:1910.08980](#).
- Bravyi, S., A. Kliesch, R. Koenig, and E. Tang (2020b), [arXiv:2011.13420](#).
- Bravyi, S. B., and A. Y. Kitaev (2002), *Ann. Phys.* **298** (1), 210.
- Breuer, H.-P., F. Petruccione, *et al.* (2002), *The theory of open quantum systems* (Oxford University Press on Demand).
- Briegel, H. J., D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest (2009), *Nat. Phys.* **5** (1), 19.

- Bromley, T. R., J. M. Arrazola, S. Jahangiri, J. Izaac, N. Quesada, A. D. Gran, M. Schuld, J. Swinerton, Z. Zabaneh, and N. Killoran (2020), *Quantum Sci. Technol.* **5** (3), 034010.
- Broughton, M., G. Verdon, T. McCourt, A. J. Martinez, J. H. Yoo, S. V. Isakov, P. Massey, M. Y. Niu, R. Halavati, E. Peters, *et al.* (2020), [arXiv:2003.02989](#).
- Brown, W. G., and B. Eastin (2018), *Phys. Rev. A* **97** (6), 062323.
- Brunner, N., D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner (2014), *Rev. Mod. Phys.* **86**, 419.
- Burges, C. J. (2002), *Microsoft Research* 10.1.1.67.7626.
- Butcher, J. C. (1987), *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods* (Wiley-Interscience).
- Bylaska, E. J., D. Song, N. P. Bauman, K. Kowalski, D. Claudino, and T. S. Humble (2020), [arXiv:2009.00080](#).
- Cade, C., L. Mineh, A. Montanaro, and S. Stanisic (2020), *Phys. Rev. B* **102** (23), 235122.
- Cai, Z. (2020), *Phys. Rev. Appl.* **14** (1), 014059.
- Calderbank, A. R., and P. W. Shor (1996), *Phys. Rev. A* **54** (2), 1098.
- Campagne-Ibarcq, P., A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, *et al.* (2020), *Nature* **584** (7821), 368.
- Cao, Y., J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, *et al.* (2019), *Chem. Rev.* **119** (19), 10856.
- Cárdenas-López, F. A., L. Lamata, J. C. Retamal, and E. Solano (2018), *PloS one* **13** (7), e0200455.
- Carignan-Dugas, A., J. J. Wallman, and J. Emerson (2015), *Phys. Rev. A* **92** (6), 060302.
- Carleo, G., I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová (2019), *Rev. Mod. Phys.* **91** (4), 045002.
- Carleo, G., and M. Troyer (2017), *Science* **355** (6325), 602.
- Cerezo, M., A. Poremba, L. Cincio, and P. J. Coles (2020a), *Quantum* **4**, 248.
- Cerezo, M., K. Sharma, A. Arrasmith, and P. J. Coles (2020b), [arXiv:2004.01372](#).
- Cerezo, M., A. Sone, T. Volkoff, L. Cincio, and P. J. Coles (2020c), [arXiv:2001.00550](#).
- Cervera-Lierta, A. (2018), *Quantum* **2**, 114.
- Cervera-Lierta, A., J. S. Kottmann, and A. Aspuru-Guzik (2020), [arXiv:2009.13545](#).
- Cervera-Lierta, A., J. I. Latorre, and D. Goyeneche (2019), *Phys. Rev. A* **100**, 022342.
- Chamberland, C., K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller, *et al.* (2020), [arXiv:2012.04108](#).
- Chang, D. E., V. Vuletić, and M. D. Lukin (2014), *Nat. Photon.* **8** (9), 685.
- Chapman, B. J., E. I. Rosenthal, J. Kerckhoff, B. A. Moores, L. R. Vale, J. Mates, G. C. Hilton, K. Lalumiere, A. Blais, and K. Lehnert (2017), *Phys. Rev. X* **7** (4), 041043.
- Chen, J., H. I. Nurdin, and N. Yamamoto (2020), [arXiv:2001.09498](#).
- Chen, S. Y. C., C. H. H. Yang, J. Qi, P. Y. Chen, X. Ma, and H. S. Goan (2020), *IEEE Access* **8**, 141007.
- Cheng, S., J. Chen, and L. Wang (2018), *Entropy* **20** (8), 583.
- Chia, N.-H., H.-H. Lin, and C. Wang (2018), [arXiv:1811.04852](#).
- Chien, R. W., and J. D. Whitfield (2020), [arXiv:2009.11860](#).
- Childs, A. M., R. Kothari, and R. D. Somma (2017), *SIAM J. Comput.* **46** (6), 1920.
- Chivilikhin, D., A. Samarin, V. Ulyantsev, I. Iorsh, A. R. Oganov, and O. Kyriienko (2020), [arXiv:2007.04424](#).
- Choi, J., S. Oh, and J. Kim (2020), [arXiv:2004.11229](#).
- Choi, J.-y., S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross (2016), *Science* **352** (6293), 1547.
- Chong, F. T., D. Franklin, and M. Martonosi (2017), *Nature* **549** (7671), 180.
- Choquette, A., A. Di Paolo, P. K. Barkoutsos, D. Sénéchal, I. Tavernelli, and A. Blais (2020), [arXiv:2008.01098](#).
- Chowdhury, A. N., G. H. Low, and N. Wiebe (2020), [arXiv:2002.00055](#).
- Ciliberto, C., M. Herbster, A. D. Ialongo, M. Pontil, A. Rocchetto, S. Severini, and L. Wossnig (2018), *Proc. Math. Phys. Eng. Sci.* **474** (2209), 20170551.
- Cirstoiu, C., Z. Holmes, J. Iosue, L. Cincio, P. J. Coles, and A. Sornborger (2020), *npj Quantum Inf.* **6** (1), 1.
- Cohen, J., A. Khan, and C. Alexander (2020), [arXiv:2007.01430](#).
- Colless, J. I., V. V. Ramasesh, D. Dahlen, M. S. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. De Jong, and I. Siddiqi (2018), *Phys. Rev. X* **8** (1), 011021.
- Commeau, B., M. Cerezo, Z. Holmes, L. Cincio, P. J. Coles, and A. Sornborger (2020), [arXiv:2009.02559](#).
- Cook, J., S. Eidenbenz, and A. Bäertschi (2019), [arXiv:1910.13483](#).
- Cornelissen, A. (2018), *Master thesis* (TU Delft).
- Cortes, C., and V. Vapnik (1995), *Mach. Learn.* **20** (3), 273.
- Cowtan, A., S. Dilkes, R. Duncan, A. Krajenbrink, W. Simmons, and S. Sivarajah (2019a), [arXiv:1902.08091](#).
- Cowtan, A., S. Dilkes, R. Duncan, W. Simmons, and S. Sivarajah (2019b), [arXiv:1906.01734](#).
- Coyle, B., M. Henderson, J. C. J. Le, N. Kumar, M. Paini, and E. Kashefi (2020a), *Quantum Sci. Technol.*
- Coyle, B., D. Mills, V. Danos, and E. Kashefi (2020b), *npj Quantum Inf.* **6** (1), 1.
- Crawford, D., A. Levit, N. Ghadermarzy, J. S. Oberoi, and P. Ronagh (2016), [arXiv:1612.05695](#).
- Crawford, D., A. Levit, N. Ghadermarzy, J. S. Oberoi, and P. Ronagh (2019a), [arXiv:1612.05695](#).
- Crawford, O., B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley (2019b), [arXiv:1908.06942](#).
- Crooks, G. E. (2018), [arXiv:1811.08419](#).
- Cross, A. W., L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta (2019), *Phys. Rev. A* **100** (3), 032328.
- Cross, A. W., L. S. Bishop, J. A. Smolin, and J. M. Gambetta (2017), [arXiv:1707.03429](#).
- Cross, A. W., E. Magesan, L. S. Bishop, J. A. Smolin, and J. M. Gambetta (2016), *npj Quantum Inf.* **2** (1), 1.
- Dalla Torre, E. G., S. Diehl, M. D. Lukin, S. Sachdev, and P. Strack (2013), *Phys. Rev. A* **87** (2), 023831.
- Dallaire-Demers, P.-L., and N. Killoran (2018), *Phys. Rev. A* **98** (1), 012324.
- Dallaire-Demers, P.-L., J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik (2019), *Quantum Sci. Technol.* **4** (4), 045005.
- Dallaire-Demers, P.-L., M. Stęchły, J. F. Gonthier, N. T. Bashige, J. Romero, and Y. Cao (2020), [arXiv:2003.01862](#).

- Das, S., A. J. Wildridge, S. B. Vaidya, and A. Jung (2019), [arXiv:1903.08879](#).
- Dattani, N. S., and N. Bryans (2014), [arXiv:1411.6758](#).
- Dawson, C. M., and M. A. Nielsen (2006), *Quantum Info. Comput.* **6** (1), 81–95.
- De, A. (2015), [arXiv:1509.07905](#).
- Deb, K., S. Agrawal, A. Pratab, and T. Meyarivan (2000), *Proc. of the Parallel Problem Solving from Nature {VI} Conference*, 849.
- Denchev, V. S., S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven (2016), *Phys. Rev. X* **6** (3), 031015.
- Deutsch, D. (1985), *Proc. Math. Phys. Sci.* **400** (1818), 97.
- Deutsch, I. H. (2020), *PRX Quantum* **1** (2), 020101.
- Devitt, S. J., A. M. Stephens, W. J. Munro, and K. Nemoto (2013), *Nat. commun.* **4** (1), 1.
- Dive, B., A. Pitchford, F. Mintert, and D. Burgarth (2018), *Quantum* **2**, 80.
- Dodd, J. L., M. A. Nielsen, M. J. Bremner, and R. T. Thew (2002), *Phys. Rev. A* **65**, 040301.
- Dong, D., C. Chen, H. Li, and T.-J. Tarn (2008), *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38** (5), 1207.
- Dongarra, J. J. (1987), *International Conference on Supercomputing* (Springer).
- Douçot, B., and L. Ioffe (2012), *Rep. Prog. Phys.* **75** (7), 072001.
- Du, Y., M.-H. Hsieh, T. Liu, and D. Tao (2020), *Phys. Rev. Res.* **2** (3), 033125.
- Dumitrescu, E. F., A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski (2018), *Phys. Rev. Lett.* **120** (21), 210501.
- Duncan, R., A. Kissinger, S. Perdrix, and J. Van De Wetering (2020), *Quantum* **4**, 279.
- Dunjko, V., and H. J. Briegel (2018), *Rep. Prog. Phys.* **81** (7), 074001.
- Dunjko, V., J. M. Taylor, and H. J. Briegel (2016), *Phys. Rev. Lett.* **117** (13), 130501.
- Dunjko, V., J. M. Taylor, and H. J. Briegel (2017), *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 282.
- Dür, W., M. Skotiniotis, F. Fröwis, and B. Kraus (2014), *Phys. Rev. Lett.* **112**, 080801.
- Ebadi, S., T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samajdar, H. Pichler, W. W. Ho, *et al.* (2020), [arXiv:2012.12281](#).
- Efthymiou, S., S. Ramos-Calderer, C. Bravo-Prieto, A. Pérez-Salinas, D. García-Martín, A. García-Saez, J. I. Latorre, and S. Carrazza (2020), [arXiv:2009.01845](#).
- Egan, L., D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, *et al.* (2020), [arXiv:2009.11482](#).
- Egger, D. J., C. Gambella, J. Marecek, S. McFaddin, M. Mevissen, R. Raymond, A. Simonetto, S. Woerner, and E. Yndurain (2020a), *IEEE Transactions on Quantum Engineering* **10.1109/TQE.2020.3030314**.
- Egger, D. J., J. Marecek, and S. Woerner (2020b), [arXiv:2009.10095](#).
- Eisert, J., D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi (2019), [arXiv:1910.06343](#).
- Eisert, J., D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi (2020), *Nat. Rev. Phys.* **2** (7), 382.
- Elben, A., B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller (2020), *Phys. Rev. Lett.* **124** (1), 010504.
- Elben, A., B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller (2018), *Phys. Rev. Lett.* **120** (5), 050406.
- Endo, S., S. C. Benjamin, and Y. Li (2018), *Phys. Rev. X* **8** (3), 031027.
- Endo, S., Z. Cai, S. C. Benjamin, and X. Yuan (2020a), [arXiv:2011.01382](#).
- Endo, S., I. Kurata, and Y. O. Nakagawa (2020b), *Phys. Rev. Res.* **2** (3), 033281.
- Endo, S., J. Sun, Y. Li, S. C. Benjamin, and X. Yuan (2020c), *Phys. Rev. Lett.* **125**, 010501.
- Entropica Labs, (2019), “Entropicaqaoa,”.
- Erhard, A., J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt (2019), *Nat. commun.* **10** (1), 1.
- Evangelista, F. A., G. K.-L. Chan, and G. E. Scuseria (2019), *J. Chem. Phys.* **151** (24), 244112.
- Fan, K. (1951), *Proc. Natl. Acad. Sci. U. S. A.* **37** (11), 760.
- Farhi, E., J. Goldstone, and S. Gutmann (2002), [arXiv:0201031](#).
- Farhi, E., J. Goldstone, and S. Gutmann (2014), [arXiv:1411.4028](#).
- Farhi, E., and A. W. Harrow (2016), [arXiv:1602.07674](#).
- Farhi, E., and H. Neven (2018), [arXiv:1802.06002](#).
- Ferguson, R. R., L. Dellantonio, K. Jansen, A. A. Balushi, W. Dür, and C. A. Muschik (2020), [arXiv:2010.13940](#).
- Fermann, J. T., and E. F. Valeev (2020), [arXiv:2007.12057](#).
- Festa, P., P. M. Pardalos, M. G. Resende, and C. C. Ribeiro (2002), *Optim. Methods Softw.* **17** (6), 1033.
- Feynman, R. P. (1982), *Int. J. Theor. Phys.* **21**, 467.
- Fingerhuth, M., T. Babej, and C. Ing (2018), [arXiv:1810.13411](#).
- Fishman, M., S. R. White, and E. M. Stoudenmire (2020), [arXiv:2007.14822](#).
- Fitzpatrick, M., N. M. Sundaresan, A. C. Li, J. Koch, and A. A. Houck (2017), *Phys. Rev. X* **7** (1), 011016.
- Fleming, C., and B. Hu (2012), *Ann. Phys.* **327** (4), 1238.
- Fletcher, A. S., P. W. Shor, and M. Z. Win (2008), *IEEE Transactions on Information Theory* **54** (12), 5705.
- Fletcher, R. (2000), *Mathematics of Computation* (John Wiley & Sons, Ltd, Chichester, West Sussex England).
- Fowler, A. G., M. Mariantoni, J. M. Martinis, and A. N. Cleland (2012), *Phys. Rev. A* **86** (3), 032324.
- Foxen, B., C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Z. Chen, K. Satzinger, R. Barends, *et al.* (2020), *Phys. Rev. Lett.* **125** (12), 120504.
- França, D. S., and A. Hashagen (2018), *J. Phys. A Math. Theor.* **51** (39), 395302.
- Fujii, K., and K. Nakajima (2017), *Phys. Rev. Appl.* **8** (2), 024030.
- Fund, U. (2020), “Unitary fund: Because evolution is unitary.” <https://unitary.fund/>.
- Gaitan, F. (2020), *npj Quantum Inf.* **6** (1), 1.
- Galicía, A., B. Ramon, E. Solano, and M. Sanz (2020), *Phys. Rev. Res.* **2** (3), 033103.
- Gambetta, J. M., A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto, T. A. Ohki, *et al.* (2012), *Phys. Rev. Lett.* **109** (24), 240504.
- Gambetta, J. M., F. Motzoi, S. Merkel, and F. K. Wilhelm (2011), *Phys. Rev. A* **83** (1), 012308.



- Ganzhorn, M., D. Egger, P. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, M. Roth, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, and S. Filipp (2019), *Phys. Rev. Appl.* **11** (4), 044092.
- García-Álvarez, L., J. Casanova, A. Mezzacapo, I. L. Egusquiza, L. Lamata, G. Romero, and E. Solano (2015), *Phys. Rev. Lett.* **114**, 070502.
- Garcia-Saez, A., and J. Latorre (2018), [arXiv:1806.02287](#).
- Gard, B. T., L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes (2020), *npj Quantum Inf.* **6** (1), 1.
- Garhwal, S., M. Ghorani, and A. Ahmad (2019), *Arch. Computat. Methods Eng.*, 1.
- Gay, S. J. (2006), *Math. Struct. Comp. Sci.* **16** (4), 581.
- Georgescu, I. M., S. Ashhab, and F. Nori (2014), *Rev. Mod. Phys.* **86** (1), 153.
- Ghosh, S., A. Opala, M. Matuszewski, T. Paterek, and T. C. Liew (2019), *npj Quantum Inf.* **5** (1), 1.
- Gilyén, A., S. Lloyd, and E. Tang (2018), [arXiv:1811.04909](#).
- Giovannetti, V., S. Lloyd, and L. Maccone (2008), *Phys. Rev. Lett.* **100**, 160501.
- Giurgica-Tiron, T., Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng (2020), [arXiv:2005.10921](#).
- Goemans, M. X., and D. P. Williamson (1995), *J. ACM* **42** (6), 1115–1145.
- Gokhale, P., O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong (2019), [arXiv:1907.13623](#).
- Gokhale, P., and F. T. Chong (2019), [arXiv:1908.11857](#).
- Gokhale, P., A. Javadi-Abhari, N. Earnest, Y. Shi, and F. T. Chong (2020), [arXiv:2004.11205](#).
- Goldman, N., and J. Dalibard (2014), *Phys. Rev. X* **4** (3), 031027.
- Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio (2014), *Adv. Neural Inform. Process. Syst.*, 2672.
- Gosset, D., V. Kliuchnikov, M. Mosca, and V. Russo (2013), [arXiv:1308.4134](#).
- Gottesman, D. (1997), [arXiv:9705052](#).
- Gottlieb, D., and S. A. Orszag (1977), *Numerical analysis of spectral methods: theory and applications* (SIAM).
- Gould, S. H. (2012), *Variational methods for eigenvalue problems: an introduction to the methods of Rayleigh, Ritz, Weinstein, and Aronszajn* (Courier Corporation).
- de la Grand'rive, P. D., and J.-F. Hullo (2019), [arXiv:1908.02210](#).
- Grant, E., L. Wossnig, M. Ostaszewski, and M. Benedetti (2019), *Quantum* **3**, 214.
- Gray, J. (2018), *J. Open Source Softw.* **3** (29), 819.
- Greenbaum, D. (2015), [arXiv:1509.02921](#).
- Greenberger, D. M., M. A. Horne, A. Shimony, and A. Zeilinger (1990), *Am. J. Phys.* **58** (12), 1131.
- Gregorcic, G. (2001), *matrix* **2** (1), 2.
- Grimsley, H. R., D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall (2019a), *J. Chem. Theory Comput.*.
- Grimsley, H. R., S. E. Economou, E. Barnes, and N. J. Mayhall (2019b), *Nat. commun.* **10** (1), 1.
- Gross, C., and I. Bloch (2017), *Science* **357** (6355), 995.
- Guerreschi, G. G., J. Hogaboam, F. Baruffa, and N. P. Sawaya (2020), *Quantum Sci. Technol.* **5** (3), 034007.
- de Guise, H., O. Di Matteo, and L. L. Sánchez-Soto (2018), *Phys. Rev. A* **97**, 022328.
- Gurvits, L. (2003), *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, 10.
- Hadfield, S., Z. Wang, B. O’Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas (2019), *Algorithms* **12** (2), 10.3390/a12020034.
- Hadfield, S., Z. Wang, E. G. Rieffel, B. O’Gorman, D. Venturelli, and R. Biswas (2017), *Proceedings of the Second International Workshop on Post Moores Era Supercomputing*, 15–21.
- Häffner, H., C. F. Roos, and R. Blatt (2008), *Phys. Rep.* **469** (4), 155.
- Hallberg, K. A. (2006), *Adv. Phys.* **55** (5-6), 477.
- Halperin, E., D. Livnat, and U. Zwick (2004), *J. Algorithms* **53** (2), 169.
- Hamamura, I., and T. Imamichi (2020), *npj Quantum Inf.* **6** (1), 1.
- Hamilton, C. S., R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex (2017), *Phys. Rev. Lett.* **119** (17), 170501.
- Hamilton, K. E., E. F. Dumitrescu, and R. C. Pooser (2019), *Phys. Rev. A* **99** (6), 062323.
- Hanks, M., M. P. Estarellas, W. J. Munro, and K. Nemoto (2020), *Phys. Rev. X* **10**, 041030.
- Hansen, N., S. D. Müller, and P. Koumoutsakos (2003), *Evol. Comput.* **11** (1), 1.
- Harper, R., and S. T. Flammia (2017), *Quantum Sci. Technol.* **2** (1), 015008.
- Harrow, A., and J. Napp (2019), [arXiv:1901.05374](#).
- Harrow, A. W. (2020), [arXiv:2004.00026](#).
- Harrow, A. W., A. Hassidim, and S. Lloyd (2009), *Phys. Rev. Lett.* **103** (15), 150502.
- Härtle, R., C. Benesch, and M. Thoss (2008), *Phys. Rev. B* **77** (20), 205314.
- Hartmann, M. J. (2016), *J. Opt.* **18** (10), 104005.
- Hashagen, A., S. Flammia, D. Gross, and J. Wallman (2018), *Quantum* **2**, 85.
- Hastings, M. B. (2019), [arXiv:1905.07047](#).
- Hastings, M. B. (2020), [arXiv:2005.03791](#).
- Haug, T., and K. Bharti (2020), [arXiv:2011.14737](#).
- Haug, T., R. Dumke, L.-C. Kwek, C. Miniatura, and L. Amico (2021), *Phys. Rev. Research* **3**, 013034.
- Hauke, P., H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver (2020), *Rep. Prog. Phys.* **83** (5), 054401.
- Hauke, P., D. Marcos, M. Dalmonte, and P. Zoller (2013), *Phys. Rev. X* **3**, 041018.
- Havlíček, V., A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta (2019), *Nature* **567** (7747), 209.
- Headley, D., T. Müller, A. Martin, E. Solano, M. Sanz, and F. K. Wilhelm (2020), [arXiv:2002.12215](#).
- Heim, B., M. Soeken, S. Marshall, C. Granade, M. Roetteler, A. Geller, M. Troyer, and K. Svore (2020), *Nat. Rev. Phys.*, 1.
- Helsen, J., X. Xue, L. M. Vandersypen, and S. Wehner (2019), *npj Quantum Inf.* **5** (1), 1.
- Helstrom, C. W. (1969), *J. Stat. Phys.* **1** (2), 231.
- Hen, I., and M. S. Sarandy (2016), *Phys. Rev. A* **93**, 062312.
- Hen, I., and F. M. Spedalieri (2016), *Phys. Rev. Appl.* **5**, 034007.
- Hensgens, T., T. Fujita, L. Janssen, X. Li, C. Van Diepen, C. Reichl, W. Wegscheider, S. D. Sarma, and L. M. Vandersypen (2017), *Nature* **548** (7665), 70.
- Herbst, M. F. (2018), *Development of a modular quantum-chemistry framework for the investigation of novel basis functions*, *Ph.D. thesis* (Ruprecht-Karls-Universität Heidelberg).

- Heya, K., K. M. Nakanishi, K. Mitarai, and K. Fujii (2019), [arXiv:1904.08566](#).
- Heyfron, L. E., and E. T. Campbell (2018), [Quantum Sci. Technol.](#) **4** (1), 015004.
- Higgott, O., D. Wang, and S. Brierley (2019), [Quantum](#) **3**, 156.
- Ho, W. W., and T. H. Hsieh (2019), [SciPost Phys.](#) **6**, 29.
- Ho, W. W., C. Jonay, and T. H. Hsieh (2019), [Phys. Rev. A](#) **99**, 052332.
- Hodson, M., B. Ruck, H. Ong, D. Garvin, and S. Dulman (2019), [arXiv:1911.05296](#).
- Hofmann, T., B. Schölkopf, and A. J. Smola (2008), [Ann. Stat.](#), 1171.
- Holmes, Z., A. Arrasmith, B. Yan, P. J. Coles, A. Albrecht, and A. T. Sornborger (2020), [arXiv:2009.14808](#).
- Holmes, Z., K. Sharma, M. Cerezo, and P. J. Coles (2021), [arXiv:2101.02138](#).
- Honeywell, (2020), “Achieving Quantum Volume 128 on the Honeywell Quantum Computer,” <https://t.ly/rCuv>, [Accessed: 25-October-2020].
- Horodecki, M., P. Horodecki, and R. Horodecki (1999), [Phys. Rev. A](#) **60** (3), 1888.
- Houck, A. A., H. E. Türeci, and J. Koch (2012), [Nat. Phys.](#) **8** (4), 292.
- Håstad, J. (2001), [J. ACM](#) **48** (4), 798–859.
- Hu, L., S.-H. Wu, W. Cai, Y. Ma, X. Mu, Y. Xu, H. Wang, Y. Song, D.-L. Deng, C.-L. Zou, *et al.* (2019), [Sci. Adv.](#) **5** (1), eaav2761.
- Hu, Z., R. Xia, and S. Kais (2020), [Sci. Rep.](#) **10** (1), 1.
- Huang, C.-J., H. Ma, Q. Yin, J.-F. Tang, D. Dong, C. Chen, G.-Y. Xiang, C.-F. Li, and G.-C. Guo (2020a), [Phys. Rev. A](#) **102**, 032412.
- Huang, H.-L., Y. Du, M. Gong, Y. Zhao, Y. Wu, C. Wang, S. Li, F. Liang, J. Lin, Y. Xu, *et al.* (2020b), [arXiv:2010.06201](#).
- Huang, H.-Y., K. Bharti, and P. Rebentrost (2019), [arXiv:1909.07344](#).
- Huang, H.-Y., R. Kueng, and J. Preskill (2020c), [Nat. Phys.](#) **16**, 1050.
- Huang, K., Z.-A. Wang, C. Song, K. Xu, H. Li, Z. Wang, Q. Guo, Z. Song, Z.-B. Liu, D. Zheng, *et al.* (2020d), [arXiv:2009.12827](#).
- Huang, L., and L. Wang (2017), [Phys. Rev. B](#) **95** (3), 035105.
- Huembeli, P., and A. Dauphin (2020), [arXiv:2008.02785](#).
- Huggins, W. J., J. Lee, U. Baek, B. O’Gorman, and K. B. Whaley (2020), [New J. Phys.](#) **22**, 073009.
- Huggins, W. J., J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush (2019), [arXiv:1907.13117](#).
- Huh, J., G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik (2015), [Nat. Photonics](#) **9** (9), 615.
- Huh, J., S. Mostame, T. Fujita, M.-H. Yung, and A. Aspuru-Guzik (2014), [New J. Phys.](#) **16** (12), 123008.
- Huyer, W., and A. Neumaier (2008), [ACM Transactions on Mathematical Software](#) **35** (2), 1.
- IBM, (2020), “IBM delivers its highest quantum volume to date, expanding the computational power of its ibm cloud-accessible quantum computers,” <https://t.ly/hoPX>, [Accessed: 25-October-2020].
- IBMQ, (2019), “Qiskit: An open-source framework for quantum computing,” 10.5281/zenodo.2562110.
- Inagaki, T., Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. L. McMahon, T. Umeki, K. Enbutsu, *et al.* (2016), [Science](#) **354** (6312), 603.
- Izmaylov, A. F., M. Díaz-Tinoco, and R. A. Lang (2020), [Phys. Chem. Chem. Phys.](#) **22**, 12980.
- Izmaylov, A. F., T.-C. Yen, R. A. Lang, and V. Vertelecky (2019a), [J. Chem. Theory Comput.](#) **16** (1), 190.
- Izmaylov, A. F., T.-C. Yen, and I. G. Ryabinkin (2019b), [Chem. Sci.](#) **10** (13), 3746.
- Jaderberg, B., A. Agarwal, K. Leonhardt, M. Kiffner, and D. Jaksch (2020), [Quantum Sci. Technol.](#) **5**, 034015.
- Jahangiri, S., J. M. Arrazola, and A. Delgado (2020a), [arXiv:2012.09231](#).
- Jahangiri, S., J. M. Arrazola, N. Quesada, and N. Killoran (2020b), [Phys. Rev. E](#) **101**, 022134.
- JavadiAbhari, A., S. Patil, D. Kudrow, J. Heckey, A. Lvov, F. T. Chong, and M. Martonosi (2014), [Proceedings of the 11th ACM Conference on Computing Frontiers](#), 1.
- Jensen, P. W. K., L. B. Kristensen, J. S. Kottmann, and A. Aspuru-Guzik (2020), [Quantum Sci. Technol.](#) **6** (1), 015004.
- Jerbi, S., L. M. Trenkwalder, H. P. Nautrup, H. J. Briegel, and V. Dunjko (2019), [arXiv:1910.12760](#).
- Johansson, J., P. Nation, and F. Nori (2012), [Comput. Phys. Commun.](#) **183** (8), 1760.
- Johnson, P. D., J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik (2017), [arXiv:1711.02249](#).
- Jones, N. C., R. Van Meter, A. G. Fowler, P. L. McMahon, J. Kim, T. D. Ladd, and Y. Yamamoto (2012), [Phys. Rev. X](#) **2** (3), 031007.
- Jones, T. (2020), [arXiv:2011.02991](#).
- Jones, T., and S. C. Benjamin (2020), [Quantum Sci. Technol.](#) **5**, 034012.
- Jones, T., A. Brown, I. Bush, and S. C. Benjamin (2019a), [Sci. Rep.](#) **9**, 10736.
- Jones, T., S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin (2019b), [Phys. Rev. A](#) **99** (6), 062304.
- Joó, B., C. Jung, N. H. Christ, W. Detmold, R. G. Edwards, M. Savage, and P. Shanahan (2019), [Eur. Phys. J. A](#) **55** (11), 199.
- Jordan, P., and O. Klein (1927), [Z. Phys.](#) **45** (11-12), 751.
- Jordan, P., and E. P. Wigner (1928), [Z. Phys.](#) **47** (631), 14.
- Jørgensen, P. (2012), *Second quantization-based methods in quantum chemistry* (Elsevier).
- Joshi, M. K., A. Elben, B. Vermersch, T. Brydges, C. Maier, P. Zoller, R. Blatt, and C. F. Roos (2020), [Phys. Rev. Lett.](#) **124**, 240505.
- Jouzdani, P., and S. Bringuier (2020), [arXiv:2008.11347](#).
- Kaiser, D. (2011), *How the hippies saved physics: science, counterculture, and the quantum revival* (WW Norton & Company).
- Kako, S., T. Leleu, Y. Inui, F. Khoyrater, S. Reifenstein, and Y. Yamamoto (2020), [Adv. Quantum Technol.](#), 2000045.
- Kandala, A., A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta (2017), [Nature](#) **549** (7671), 242.
- Kandala, A., K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta (2019), [Nature](#) **567** (7749), 491.
- Karamlou, A. H., W. A. Simon, A. Katabarwa, T. L. Scholten, B. Peropadre, and Y. Cao (2020), [arXiv:2012.07825](#).
- Kaubruegger, R., P. Silvi, C. Kokail, R. van Bijnen, A. M. Rey, J. Ye, A. M. Kaufman, and P. Zoller (2019), [Phys. Rev. Lett.](#) **123** (26), 260505.
- Kempe, J., A. Kitaev, and O. Regev (2006), [SIAM J. Comput.](#) **35** (5), 1070.

- Khairy, S., R. Shaydulin, L. Cincio, Y. Alexeev, and P. Balaprakash (2019), [arXiv:1911.04574](#).
- Khammassi, N., G. G. Guerreschi, I. Ashraf, J. W. Hogaboam, C. G. Almudever, and K. Bertels (2018), [arXiv:1805.09607](#).
- Khatri, S., R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles (2019), [Quantum](#) **3**, 140.
- Kieferová, M., and N. Wiebe (2017), [Phys. Rev. A](#) **96**, 062327.
- Killoran, N., J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook (2019), [Quantum](#) **3**, 129.
- Kirby, W. M., and P. J. Love (2019), [Phys. Rev. Lett.](#) **123** (20), 200501.
- Kirby, W. M., A. Tranter, and P. J. Love (2020), [arXiv:2011.10027](#).
- Kissinger, A., and J. van de Wetering (2019), [arXiv:1903.10477](#).
- Kissinger, A., and J. van de Wetering (2020a), [Proceedings QPL 2019](#) **318**, 229.
- Kissinger, A., and J. van de Wetering (2020b), [Phys. Rev. A](#) **102**, 022406.
- Kitaev, A. Y. (2003), [Ann. Phys.](#) **303** (1), 2.
- Kivlichan, I. D., J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush (2018), [Phys. Rev. Lett.](#) **120**, 110501.
- Kivlichan, I. D., N. Wiebe, R. Babbush, and A. Aspuru-Guzik (2017), [J. Phys. A Math. Theor.](#) **50** (30), 305301.
- Kjaergaard, M., M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver (2020), [Annu. Rev. Condens. Matter Phys.](#) **11**, 369.
- Klco, N., E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage (2018), [Phys. Rev. A](#) **98**, 032331.
- Klema, V., and A. Laub (1980), [IEEE Transactions on automatic control](#) **25** (2), 164.
- Knill, E., and R. Laflamme (1997), [Phys. Rev. A](#) **55** (2), 900.
- Koczor, B., and S. C. Benjamin (2019), [arXiv:1912.08660](#).
- Koczor, B., and S. C. Benjamin (2020), [arXiv:2008.13774](#).
- Koczor, B., S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin (2020), [New J. Phys.](#) **22**, 083038.
- Kokail, C., C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, *et al.* (2019), [Nature](#) **569** (7756), 355.
- Koppenhöfer, M., C. Bruder, and A. Roulet (2020), [Phys. Rev. Res.](#) **2** (2), 023026.
- Koren, Y., R. Bell, and C. Volinsky (2009), [Computer](#) **42** (8), 30.
- Kosut, R. L., and D. A. Lidar (2009), [Quantum Inf. Process.](#) **8** (5), 443.
- Kosut, R. L., A. Shabani, and D. A. Lidar (2008), [Phys. Rev. Lett.](#) **100** (2), 020502.
- Kottmann, J. S. (2018), *Coupled-Cluster in Real Space*, [Ph.D. thesis](#) (Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät).
- Kottmann, J. S., S. Alperin-Lea, T. Tamayo-Mendoza, A. Cervera-Lierta, C. Lavigne, T.-C. Yen, V. Verteletskyi, P. Schleich, A. Anand, M. Degroote, *et al.* (2020a), [arXiv:2011.03057](#).
- Kottmann, J. S., A. Anand, and A. Aspuru-Guzik (2020b), [arXiv:2011.05938](#).
- Kottmann, J. S., and F. A. Bischoff (2017), [J. Chem. Theory Comput.](#) **13** (12), 5945.
- Kottmann, J. S., M. Krenn, T. H. Kyaw, S. Alperin-Lea, and A. Aspuru-Guzik (2020c), [arXiv:2006.03075](#).
- Kottmann, J. S., P. Schleich, T. Tamayo-Mendoza, and A. Aspuru-Guzik (2021), [J. Phys. Chem. Lett.](#) **12** (1), 663.
- Krantz, P., M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver (2019), [Appl. Phys. Rev.](#) **6** (2), 021318.
- Krenn, M., M. Erhard, and A. Zeilinger (2020a), [Nat Rev Phys](#) **2**, 649–661.
- Krenn, M., J. Kottmann, N. Tischler, and A. Aspuru-Guzik (2020b), [arXiv:2005.06443](#).
- Kreula, J., S. R. Clark, and D. Jaksch (2016), [Sci. Rep.](#) **6**, 32940.
- Krinner, S., P. Kurpiers, B. Royer, P. Marnard, I. Tsitsilin, J.-C. Besse, A. Remm, A. Blais, and A. Wallraff (2020), [Phys. Rev. Appl.](#) **14** (4), 044039.
- Kruse, R., C. S. Hamilton, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex (2019), [Phys. Rev. A](#) **100** (3), 032326.
- Kubo, K., Y. O. Nakagawa, S. Endo, and S. Nagayama (2020), [arXiv:2012.04429](#).
- Kunitsa, A. A., and S. Hirata (2020), [Phys. Rev. E](#) **101**, 013311.
- Kusumoto, T., K. Mitarai, K. Fujii, M. Kitagawa, and M. Negoro (2019), [arXiv:1911.12021](#).
- Kyaw, T., S. Felicetti, G. Romero, E. Solano, and L. C. Kwek (2014), [Proc. SPIE. XII](#) **9225**, 92250B.
- Kyaw, T. H. (2019), *Towards a Scalable Quantum Computing Platform in the Ultrastrong Coupling Regime* (Springer).
- Kyaw, T. H., V. M. Bastidas, J. Tangpanitanon, G. Romero, and L.-C. Kwek (2020a), [Phys. Rev. A](#) **101** (1), 012111.
- Kyaw, T. H., S. Felicetti, G. Romero, E. Solano, and L.-C. Kwek (2015a), [Sci. Rep.](#) **5**, 8621.
- Kyaw, T. H., D. A. Herrera-Martí, E. Solano, G. Romero, and L.-C. Kwek (2015b), [Phys. Rev. B](#) **91** (6), 064503.
- Kyaw, T. H., T. Menke, S. Sim, N. P. Sawaya, W. D. Oliver, G. G. Guerreschi, and A. Aspuru-Guzik (2020b), [arXiv:2006.03070](#).
- Kyriienko, O., A. E. Paine, and V. E. Elfving (2020), [arXiv:2011.10395](#).
- Kyriienko, O., and A. S. Sørensen (2018), [Phys. Rev. Appl.](#) **9** (6), 064029.
- Kübler, J. M., A. Arrasmith, L. Cincio, and P. J. Coles (2020), [Quantum](#) **4**, 10.22331/q-2020-05-11-263.
- Lacroix, N., C. Hellings, C. K. Andersen, A. Di Paolo, A. Remm, S. Lazar, S. Krinner, G. J. Norris, M. Gabureac, J. Heinsoo, *et al.* (2020), [PRX Quantum](#) **1** (2), 110304.
- Lamata, L. (2017), [Sci. Rep.](#) **7** (1), 1.
- Lamata, L., U. Alvarez-Rodriguez, J. D. Martín-Guerrero, M. Sanz, and E. Solano (2018a), [Quantum Sci. Technol.](#) **4** (1), 014007.
- Lamata, L., A. Parra-Rodriguez, M. Sanz, and E. Solano (2018b), [Adv. Phys.: X](#) **3** (1), 1457981.
- Landsman, K. A., C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe (2019), [Nature](#) **567** (7746), 61.
- Lang, R. A., I. G. Ryabinkin, and A. F. Izmaylov (2020), [arXiv:2002.05701](#).
- LaRose, R., A. Mari, P. J. Karalekas, N. Shammah, and W. J. Zeng (2020), [arXiv:2009.04417](#).
- LaRose, R., A. Tikku, É. O’Neel-Judy, L. Cincio, and P. J. Coles (2019), [npj Quantum Inf.](#) **5** (1), 1.
- Las Heras, U., U. Alvarez-Rodriguez, E. Solano, and M. Sanz (2016), [Phys. Rev. Lett.](#) **116**, 230504.
- Lau, J. W. Z., K. Bharti, T. Haug, and L. C. Kwek (2021), [arXiv:2101.07677](#).

- Lavrijsen, W., and the scikit-quant contributors (2020), “scikit-quant,” <https://github.com/scikit-quant/scikit-quant>.
- Lavrijsen, W., A. Tudor, J. Müller, C. Iancu, and W. de Jong (2020), [arXiv:2004.03004](https://arxiv.org/abs/2004.03004).
- Lechner, W. (2020), *IEEE Transactions on Quantum Engineering* **1**, 1.
- Lechner, W., P. Hauke, and P. Zoller (2015), *Sci. Adv.* **1** (9), e1500838.
- Lee, C.-K., J. W. Z. Lau, L. Shi, and L. C. Kwek (2021), [arXiv:2101.06879](https://arxiv.org/abs/2101.06879).
- Lee, J., W. J. Huggins, M. Head-Gordon, and K. B. Whaley (2018), *J. Chem. Theory Comput.* **15** (1), 311.
- Lenstra, J. K., and A. R. Kan (1975), *Journal of the Operational Research Society* **26** (4), 717.
- Leyton-Ortega, V., A. Perdomo-Ortiz, and O. Perdomo (2019), [arXiv:1901.08047](https://arxiv.org/abs/1901.08047).
- Li, A. C., F. Petruccione, and J. Koch (2016), *Phys. Rev. X* **6** (2), 021037.
- Li, G., Y. Ding, and Y. Xie (2019), *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, 1001.
- Li, J., X. Yang, X. Peng, and C.-P. Sun (2017a), *Phys. Rev. Lett.* **118** (15), 150503.
- Li, L., M. Fan, M. Coram, P. Riley, and S. Leichenauer (2020), *Phys. Rev. Res.* **2**, 023074.
- Li, R., U. Alvarez-Rodriguez, L. Lamata, and E. Solano (2017b), *Quantum Meas. Quantum Metrol.* **4** (1), 1.
- Li, R. Y., R. Di Felice, R. Rohs, and D. A. Lidar (2018), *npj Quantum inf.* **4** (1), 1.
- Li, Y., and S. C. Benjamin (2017), *Phys. Rev. X* **7** (2), 021050.
- Lidar, D. A. (2014), *Adv. Chem. Phys.* **154**, 295.
- Lidar, D. A., and T. A. Brun (2013), *Quantum error correction* (Cambridge university press).
- Liu, D. C., and J. Nocedal (1989), *Math. Program.* **45** (1-3), 503.
- Liu, J., Z. Li, and J. Yang (2020a), [arXiv:2012.07047](https://arxiv.org/abs/2012.07047).
- Liu, J., Y. Qi, Z. Y. Meng, and L. Fu (2017), *Phys. Rev. B* **95**, 041101.
- Liu, J., and Y. Xin (2020), *J. High Energ. Phys.* **2020**, 11.
- Liu, J.-G., and L. Wang (2018), *Phys. Rev. A* **98** (6), 062324.
- Liu, J.-G., Y.-H. Zhang, Y. Wan, and L. Wang (2019), *Phys. Rev. Res.* **1** (2), 023025.
- Liu, Z., L.-M. Duan, and D.-L. Deng (2020b), [arXiv:2008.05488](https://arxiv.org/abs/2008.05488).
- Lloyd, S. (1996), *Science* **273**, 1073.
- Lloyd, S. (2018), [arXiv:1812.11075](https://arxiv.org/abs/1812.11075).
- Lloyd, S., M. Mohseni, and P. Rebentrost (2013), [arXiv:1307.0411](https://arxiv.org/abs/1307.0411).
- Lloyd, S., M. Schuld, A. Ijaz, J. Izaac, and N. Killoran (2020), [arXiv:2001.03622](https://arxiv.org/abs/2001.03622).
- Lloyd, S., and C. Weedbrook (2018), *Phys. Rev. Lett.* **121** (4), 040502.
- Lockwood, O., and M. Si (2020a), *NeurIPS*.
- Lockwood, O., and M. Si (2020b), *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* **16** (1), 245.
- Lu, D., K. Li, J. Li, H. Katiyar, A. J. Park, G. Feng, T. Xin, H. Li, G. Long, A. Brodutch, *et al.* (2017), *npj Quantum Inf.* **3** (1), 1.
- Lubasch, M., J. Joo, P. Moinier, M. Kiffner, and D. Jaksch (2020), *Phys. Rev. A* **101** (1), 010301.
- Lucas, A. (2014), *Front. Phys.* **2**, 5.
- Luo, X.-Z., J.-G. Liu, P. Zhang, and L. Wang (2020), *Quantum* **4**, 341.
- Lykov, D. (2020), “Q-Tensor,” <https://github.com/danlkv/qtensor>.
- Lyu, C., V. Montenegro, and A. Bayat (2020), *Quantum* **4**, 324.
- Ma, Z., P. Gokhale, T.-X. Zheng, S. Zhou, X. Yu, L. Jiang, P. Maurer, and F. T. Chong (2020), [arXiv:2010.08702](https://arxiv.org/abs/2010.08702).
- Magann, A. B., C. Arenz, M. D. Grace, T.-S. Ho, R. L. Kosut, J. R. McClean, H. A. Rabitz, and M. Sarovar (2021), *Phys. Rev. X Quantum* **2**, 010101.
- Magesan, E., J. M. Gambetta, and J. Emerson (2011), *Phys. Rev. Lett.* **106** (18), 180504.
- Magesan, E., J. M. Gambetta, and J. Emerson (2012), *Phys. Rev. A* **85**, 042311.
- Mahadev, U. (2018), *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, 259.
- Mahoney, A., J. Colless, S. Pauka, J. Hornibrook, J. Watson, G. Gardner, M. Manfra, A. Doherty, and D. Reilly (2017), *Phys. Rev. X* **7** (1), 011007.
- Manin, Y. (1980), *Sovetskoye Radio, Moscow* **128**.
- Manne, A. S. (1960), *Oper. Res.* **8** (2), 219.
- Marandi, A., Z. Wang, K. Takata, R. L. Byer, and Y. Yamamoto (2014), *Nat. Photonics* **8** (12), 937.
- Mari, A., T. R. Bromley, and N. Killoran (2020), [arXiv:2008.06517](https://arxiv.org/abs/2008.06517).
- Markowitz, H. (1952), *The Journal of Finance* **7** (1), 77.
- Marrero, C. O., M. Kieferová, and N. Wiebe (2020), [arXiv:2010.15968](https://arxiv.org/abs/2010.15968).
- Martin, A., L. Lamata, E. Solano, and M. Sanz (2020), *Phys. Rev. Res.* **2** (1), 013012.
- Martinez, E. A., C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, *et al.* (2016), *Nature* **534** (7608), 516.
- Marzec, M. (2016), *Handbook of High-Frequency Trading and Modeling in Finance*, 73.
- Massey, P., J. A. Clark, and S. Stepney (2004), *Genetic and Evolutionary Computation Conference*, 569.
- Massey, P., J. A. Clark, and S. Stepney (2006), *Evol. Comput.* **14** (1), 21.
- McArdle, S., S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan (2020), *Rev. Mod. Phys.* **92** (1), 015003.
- McArdle, S., T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan (2019a), *npj Quantum Inf.* **5** (1), 1.
- McArdle, S., A. Mayorov, X. Shan, S. Benjamin, and X. Yuan (2019b), *Chem. Sci.* **10** (22), 5725.
- McArdle, S., X. Yuan, and S. Benjamin (2019c), *Phys. Rev. Lett.* **122** (18), 180501.
- McCaskey, A. J., D. I. Lyakh, E. F. Dumitrescu, S. S. Powers, and T. S. Humble (2020), *Quantum Sci. Technol.* **5** (2), 024002.
- McClean, J., N. Rubin, K. Sung, I. D. Kivlichan, X. Bonet-Monroig, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, *et al.* (2020a), *Quantum Sci. Technol.*.
- McClean, J. R., S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven (2018), *Nat. Commun.* **9** (1), 1.
- McClean, J. R., M. P. Harrigan, M. Mohseni, N. C. Rubin, Z. Jiang, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven (2020b), [arXiv:2008.08615](https://arxiv.org/abs/2008.08615).
- McClean, J. R., Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven (2020c), *Nat. Commun.* **11** (1), 636.
- McClean, J. R., M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong (2017), *Phys. Rev. A* **95** (4), 042308.



- McClean, J. R., J. Romero, R. Babbush, and A. Aspuru-Guzik (2016), *New J. Phys.* **18** (2), 023023.
- McKay, D. C., T. Alexander, L. Bello, M. J. Biercuk, L. Bishop, J. Chen, J. M. Chow, A. D. Córcoles, D. Egger, S. Filipp, *et al.* (2018), [arXiv:1809.03452](#).
- McKay, D. C., S. Sheldon, J. A. Smolin, J. M. Chow, and J. M. Gambetta (2019), *Phys. Rev. Lett.* **122**, 200502.
- McMahon, P. L., A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, *et al.* (2016), *Science* **354** (6312), 614.
- Meitei, O. R., B. T. Gard, G. S. Barron, D. P. Pappas, S. E. Economou, E. Barnes, and N. J. Mayhall (2020), [arXiv:2008.04302](#).
- Mengoni, R., and A. Di Pierro (2019), *Quantum Mach. Intell.* **1**, 65.
- Menke, T., F. Häse, S. Gustavsson, A. J. Kerman, W. D. Oliver, and A. Aspuru-Guzik (2019), [arXiv:1912.03322](#).
- Merkel, S. T., J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen (2013), *Phys. Rev. A* **87** (6), 062119.
- Metger, T., and T. Vidick (2020), [arXiv:2001.09161](#).
- Meyer, J. J., J. Borregaard, and J. Eisert (2020), [arXiv:2006.06303](#).
- Mezzacapo, A., U. Las Heras, J. Pedernales, L. DiCarlo, E. Solano, and L. Lamata (2014), *Sci. Rep.* **4** (1), 1.
- Miquel, C., J. P. Paz, M. Saraceno, E. Knill, R. Laflamme, and C. Negrevergne (2002), *Nature* **418** (6893), 59.
- Mitarai, K., and K. Fujii (2019), *Phys. Rev. Res.* **1** (1), 013006.
- Mitarai, K., M. Negoro, M. Kitagawa, and K. Fujii (2018), *Phys. Rev. A* **98** (3), 032309.
- Mitarai, K., T. Yan, and K. Fujii (2019), *Phys. Rev. Appl.* **11**, 044087.
- Mitchell, T. M. (1997), *Machine learning* (McGraw Hill).
- Moll, N., P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, *et al.* (2018), *Quantum Sci. Technol.* **3** (3), 030503.
- Möller, M., and M. Schalkers (2020), *International Conference on Computational Science*, 451.
- Monroe, C., W. Campbell, L.-M. Duan, Z.-X. Gong, A. Gorshkov, P. Hess, R. Islam, K. Kim, N. Linke, G. Pagano, *et al.* (2019), [arXiv:1912.07845](#).
- Montanaro, A. (2016), *npj Quantum Inf.* **2** (1), 1.
- Morales, M. E. S., J. Biamonte, and Z. Zimborás (2019), [arXiv:1909.03123](#).
- Mott, A., J. Job, J.-R. Vlimant, D. Lidar, and M. Spiropulu (2017), *Nature* **550** (7676), 375.
- Motta, M., C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandão, and G. K.-L. Chan (2020), *Nat. Phys.* **16** (2), 205.
- Motzoi, F., J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm (2009), *Phys. Rev. Lett.* **103** (11), 110501.
- Nakaji, K., and N. Yamamoto (2020a), [arXiv:2005.12537](#).
- Nakaji, K., and N. Yamamoto (2020b), [arXiv:2010.13727](#).
- Nakajima, K., K. Fujii, M. Negoro, K. Mitarai, and M. Kitagawa (2019), *Phys. Rev. Appl.* **11** (3), 034021.
- Nakanishi, K. M., K. Fujii, and S. Todo (2020), *Phys. Rev. Res.* **2** (4), 043158.
- Nakanishi, K. M., K. Mitarai, and K. Fujii (2019), *Phys. Rev. Res.* **1** (3), 033062.
- Napp, J., R. L. La Placa, A. M. Dalzell, F. G. Brandao, and A. W. Harrow (2019), [arXiv:2001.00021](#).
- Nataf, P., and C. Ciuti (2011), *Phys. Rev. Lett.* **107** (19), 190402.
- Navarrete-Benlloch, C. (2015), [arXiv:1504.05266](#).
- Negoro, M., K. Mitarai, K. Fujii, K. Nakajima, and M. Kitagawa (2018), [arXiv:1806.10910](#).
- Neill, C., P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, *et al.* (2018), *Science* **360** (6385), 195.
- Nguyen, T., A. Santana, T. Kharazi, D. Claudino, H. Finkel, and A. McCaskey (2020), [arXiv:2010.03935](#).
- Noh, K., and C. Chamberland (2020), *Phys. Rev. A* **101** (1), 012316.
- Nokkala, J., R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini (2020), [arXiv:2006.04821](#).
- Oh, Y.-H., H. Mohammadbagherpoor, P. Dreher, A. Singh, X. Yu, and A. J. Rindos (2019), [arXiv:1911.00595](#).
- Oka, T., and S. Kitamura (2019), *Annual Review of Condensed Matter Physics* **10**, 387.
- Ollitrault, P. J., A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli (2020), *Phys. Rev. Res.* **2**, 043140.
- Orús, R. (2014), *Ann. Phys.* **349**, 117.
- Orus, R., S. Mugel, and E. Lizaso (2019), *Reviews in Physics* **4**, 100028.
- Ostaszewski, M., E. Grant, and M. Benedetti (2019), [arXiv:1905.09692](#).
- Otten, M., and S. K. Gray (2019), *npj Quantum Inf.* **5**, 11.
- Otterbach, J. S., R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, *et al.* (2017), [arXiv:1712.05771](#).
- O'Malley, P. J., R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, *et al.* (2016), *Phys. Rev. X* **6** (3), 031007.
- Pagano, G., A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, L. T. Brady, A. Deshpande, F. Liu, S. Jordan, A. V. Gorshkov, and C. Monroe (2020), *Proceedings of the National Academy of Sciences* **117** (41), 25396, <https://www.pnas.org/content/117/41/25396.full.pdf>.
- Paparo, G. D., V. Dunjko, A. Makmal, M. A. Martin-Delgado, and H. J. Briegel (2014), *Phys. Rev. X* **4** (3), 031002.
- Parra-Rodriguez, A., P. Lougovski, L. Lamata, E. Solano, and M. Sanz (2020), *Phys. Rev. A* **101** (2), 022305.
- Parrish, R. M., E. G. Hohenstein, P. L. McMahon, and T. J. Martínez (2019a), *Phys. Rev. Lett.* **122** (23), 230401.
- Parrish, R. M., J. T. Iosue, A. Ozaeta, and P. L. McMahon (2019b), [arXiv:1904.03206](#).
- Parrish, R. M., and P. L. McMahon (2019), [arXiv:1909.08925](#).
- Paz, J. P., and W. H. Zurek (1998), *Proc. Math. Phys. Eng. Sci.* **454** (1969), 355.
- Pepper, A., N. Tischler, and G. J. Pryde (2019), *Phys. Rev. Lett.* **122**, 060501.
- Perdomo-Ortiz, A., M. Benedetti, J. Realpe-Gómez, and R. Biswas (2018), *Quantum Sci. Technol.* **3** (3), 030502.
- Perdomo-Ortiz, A., N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik (2012), *Sci. Rep.* **2**, 571.
- Pérez-Salinas, A., A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre (2020a), *Quantum* **4**, 226.
- Pérez-Salinas, A., D. García-Martín, C. Bravo-Prieto, and J. I. Latorre (2020b), *Entropy* **22** (4), 436.
- Peruzzo, A., J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien

- (2014), *Nat. Commun.* **5**, 4213.
- Pino, J. M., J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, C. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, *et al.* (2020), [arXiv:2003.01293](https://arxiv.org/abs/2003.01293).
- Pirhooshyaran, M., and T. Terlaky (2020), [arXiv:2012.04046](https://arxiv.org/abs/2012.04046).
- Platt, J. (1998), “Sequential minimal optimization: A fast algorithm for training support vector machines,” <https://t.ly/hoPX>, [Accessed: 16-Januaray-2021].
- Pokharel, B., N. Anand, B. Fortman, and D. A. Lidar (2018), *Phys. Rev. Lett.* **121** (22), 220502.
- Policharla, G.-V., and S. Vinjanampathy (2020), [arXiv:2011.13777](https://arxiv.org/abs/2011.13777).
- Potoček, V., A. P. Reynolds, A. Fedrizzi, and D. W. Corne (2018), [arXiv:1812.04458](https://arxiv.org/abs/1812.04458).
- Powell, M. J. (2009), *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, 26.
- Preskill, J. (2018), *Quantum* **2**, 79.
- Pulay, P. (1980), *Chem. Phys. Lett.* **73** (2), 393.
- Q-Ctrl, (2020), “Q-ctrl open controls,”.
- QCA lab, (2020), “Delft quantum & classical simulator,” <https://github.com/QE-Lab/dqcsim>.
- QCA Lab, (2020), “OpenQL,” <https://github.com/QE-Lab/OpenQL>.
- QCA lab, (2020), “QX Simulator,” <https://github.com/QE-Lab/qx-simulator>.
- QOSF, (2020), “Open-Source Quantum Software Projects,” <https://github.com/qosf/awesome-quantum-software>.
- QOSF, (2020), “Quantum open source foundation,” <https://qosf.org/>.
- Quantum, G. A., *et al.* (2020), *Science* **369** (6507), 1084.
- Quantum, M. (2020), “Microsoft quantum documentation,” <https://docs.microsoft.com/en-us/quantum/>, [Accessed: 22-December-2020].
- Quantum AI team and collab., (2020), “Cirq,”.
- Quantum Computing Reports, (2020), “Quantum computing reports,” <https://quantumcomputingreport.com/tools/>.
- Quesada, N., J. M. Arrazola, and N. Killoran (2018), *Phys. Rev. A* **98**, 062322.
- Ramos-Calderer, S., A. Pérez-Salinas, D. García-Martín, C. Bravo-Prieto, J. Cortada, J. Planagumà, and J. I. Latorre (2019), [arXiv:1912.01618](https://arxiv.org/abs/1912.01618).
- Raussendorf, R., and J. Harrington (2007), *Phys. Rev. Lett.* **98** (19), 190504.
- Rebentrost, P., B. Gupta, and T. R. Bromley (2018), *Phys. Rev. A* **98** (2), 022321.
- Rechenberg, I. (1978), in *Simulationsmethoden in der Medizin und Biologie* (Springer) pp. 83–114.
- Roggero, A., A. C. Li, J. Carlson, R. Gupta, and G. N. Perdue (2020), *Phys. Rev. D* **101** (7), 074038.
- Rohwedder, T. (2010), *An analysis for some methods and algorithms of quantum chemistry*, *Doctoral thesis* (Technische Universität Berlin, Fakultät II - Mathematik und Naturwissenschaften, Berlin).
- Romero, J., and A. Aspuru-Guzik (2019), [arXiv:1901.00848](https://arxiv.org/abs/1901.00848).
- Romero, J., R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik (2018), *Quantum Sci. Technol.* **4** (1), 014008.
- Romero, J., J. P. Olson, and A. Aspuru-Guzik (2017), *Quantum Sci. Technol.* **2** (4), 045001.
- Roos, B. O., P. R. Taylor, and P. E. Sigbahn (1980), *Chemical Physics* **48** (2), 157.
- Rosenberg, G., P. Haghnegahdar, P. Goddard, P. Carr, K. Wu, and M. L. De Prado (2016), *IEEE Journal of Selected Topics in Signal Processing* **10** (6), 1053.
- Roushan, P., C. Neill, J. Tangpanitanon, V. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsforth, *et al.* (2017), *Science* **358** (6367), 1175.
- Rudolph, M. S., N. T. Bashige, A. Katabarwa, S. Johr, B. Peropadre, and A. Perdomo-Ortiz (2020), [arXiv:2012.03924](https://arxiv.org/abs/2012.03924).
- Ryabinkin, I. G., S. N. Genin, and A. F. Izmaylov (2018a), *J. Chem. Theory Comput.* **15** (1), 249.
- Ryabinkin, I. G., A. F. Izmaylov, and S. N. Genin (2021), *Quantum Sci. Technol.*.
- Ryabinkin, I. G., R. A. Lang, S. N. Genin, and A. F. Izmaylov (2020), *J. Chem. Theory Comput.* **16** (2), 1055.
- Ryabinkin, I. G., T.-C. Yen, S. N. Genin, and A. F. Izmaylov (2018b), *J. Chem. Theory Comput.* **14** (12), 6317.
- Sagastizabal, R., X. Bonet-Monroig, M. Singh, M. A. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno, *et al.* (2019a), *Phys. Rev. A* **100** (1), 010302.
- Sagastizabal, R., X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O’Brien, and L. DiCarlo (2019b), *Phys. Rev. A* **100** (1), 010302.
- Saleem, Z. H. (2020), *Int. J. Quantum Inf.* **18** (04), 2050011.
- Santagati, R., J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, *et al.* (2018), *Sci. Adv.* **4** (1), eaap9646.
- Santos, L. F., and L. Viola (2005), *Phys. Rev. A* **72** (6), 062303.
- Sawaya, N. P., and J. Huh (2019), *J. Phys. Chem. Lett.* **10** (13), 3586.
- Sawaya, N. P., T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik, and G. G. Guerreschi (2020), *npj Quantum Inf.* **6** (1), 1.
- Schmoll, P., and R. Orús (2017), *Phys. Rev. B* **95**, 045112.
- Schölkopf, B., R. Herbrich, and A. J. Smola (2001), *Computational Learning Theory*, 416.
- Schuld, M. (2018), *Supervised learning with quantum computers* (Springer).
- Schuld, M., V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran (2019), *Phys. Rev. A* **99** (3), 032331.
- Schuld, M., A. Bocharov, K. M. Svore, and N. Wiebe (2020a), *Phys. Rev. A* **101** (3), 032308.
- Schuld, M., K. Brádler, R. Israel, D. Su, and B. Gupta (2020b), *Phys. Rev. A* **101**, 032314.
- Schuld, M., and N. Killoran (2019), *Phys. Rev. Lett.* **122** (4), 040504.
- Schuld, M., I. Sinayskiy, and F. Petruccione (2014), *Quantum Inf. Process.* **13** (11), 2567.
- Schuld, M., I. Sinayskiy, and F. Petruccione (2016), *Phys. Rev. A* **94** (2), 022342.
- Schuld, M., R. Sweke, and J. J. Meyer (2020c), [arXiv:2008.08605](https://arxiv.org/abs/2008.08605).
- Schwefel, H.-P. (1977), *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie: mit einer vergleichenden Einführung in die Hill-Climbing-und Zufallsstrategie*, Vol. 1 (Springer).
- Seeley, J. T., M. J. Richard, and P. J. Love (2012), *J. Chem. Phys.* **137** (22), 224109.
- Setia, K., and J. D. Whitfield (2018), *J. Chem. Phys.* **148** (16), 164104.

- Sharma, K., M. Cerezo, L. Cincio, and P. J. Coles (2020), [arXiv:2005.12458](#).
- Shin, S. W., G. Smith, J. A. Smolin, and U. Vazirani (2014), [arXiv:1401.7087](#).
- Shor, P. W. (1994), [Proceedings 35th annual symposium on foundations of computer science](#), 124.
- Shor, P. W. (1995), *Phys. Rev. A* **52** (4), R2493.
- Shor, P. W. (1999), *SIAM review* **41** (2), 303.
- Sieberer, L. M., M. Buchhold, and S. Diehl (2016), *Rep. Prog. Phys.* **79** (9), 096001.
- Sim, S., P. D. Johnson, and A. Aspuru-Guzik (2019), *Adv. Quantum Technol.* **2** (12), 1900070.
- Sim, S., J. Romero, J. F. Gonthier, and A. A. Kunitsa (2020), [arXiv:2010.00629](#).
- Siraichi, M. Y., V. F. d. Santos, S. Collange, and F. M. Q. Pereira (2018), [Proceedings of the 2018 International Symposium on Code Generation and Optimization CGO 2018](#), 113–125.
- Situ, H., Z. He, Y. Wang, L. Li, and S. Zheng (2020), *Inf. Sci.*.
- Skolik, A., J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib (2020), [arXiv:2006.14904](#).
- Smith, D. G., L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, *et al.* (2020), *J. Chem. Phys.* **152** (18), 184108.
- Smith, R. S., M. J. Curtis, and W. J. Zeng (2016), [arXiv:1608.03355](#).
- Somma, R., G. Ortiz, E. Knill, and J. Gubernatis (2003), [arXiv:quant-ph/0304063](#).
- Song, C., J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li (2019), *Sci. Adv.* **5** (9), eaaw5686.
- Spagnolo, N., E. Maiorino, C. Vitelli, M. Bentivegna, A. Crespi, R. Ramponi, P. Mataloni, R. Osellame, and F. Sciarrino (2017), *Sci. Rep.* **7** (1), 1.
- Spall, J. C. (1992), *IEEE Transactions on Automatic Control* **37** (3), 332.
- Stair, N. H., R. Huang, and F. A. Evangelista (2020), *J. Chem. Theory Comput.* **16** (4), 2236.
- Stefanucci, G., and R. Van Leeuwen (2013), *Nonequilibrium many-body theory of quantum systems: a modern introduction* (Cambridge University Press).
- Steiger, D. S., T. Häner, and M. Troyer (2018), *Quantum* **2**, 49.
- Stokes, J., J. Izaac, N. Killoran, and G. Carleo (2020), *Quantum* **4**, 269.
- Strano, D., and the Qrack contributors (2020), “Qrack,” <https://github.com/vm6502q/qrack>.
- Streif, M., and M. Leib (2020), *Phys. Rev. A* **102**, 042416.
- Suchsland, P., F. Tacchino, M. H. Fischer, T. Neupert, P. K. Barkoutsos, and I. Tavernelli (2020), [arXiv:2008.10914](#).
- Sun, J., X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin, and S. Endo (2020a), [arXiv:2001.04891](#).
- Sun, Q., T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan (2018), *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8** (1), e1340.
- Sun, Q., X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, and *et al.* (2020b), *J. Chem. Phys.* **153** (2), 024109.
- Sun, S.-N., M. Motta, R. N. Tazhigulov, A. T. K. Tan, G. K.-L. Chan, and A. J. Minnich (2020c), [arXiv:2009.03542](#).
- Sung, K. J., J. Yao, M. P. Harrigan, N. C. Rubin, Z. Jiang, L. Lin, R. Babbush, and J. R. McClean (2020), *Quantum Sci. Technol.* **5** (4), 044008.
- Surján, P. R. (2012), *Second quantized approach to quantum chemistry: an elementary introduction* (Springer Science & Business Media).
- Sutton, R. S., and A. G. Barto (2018), *Reinforcement learning: An introduction* (MIT press).
- Suzuki, M. (1976), *Commun. Math. Phys.* **51** (2), 183.
- Suzuki, Y., Y. Kawase, Y. Masumura, Y. Hiraga, M. Nakadai, J. Chen, K. M. Nakanishi, K. Mitarai, R. Imai, S. Tamiya, *et al.* (2020), [arXiv:2011.13524](#).
- Sweke, R., F. Wilde, J. J. Meyer, M. Schuld, P. K. Fährmann, B. Meynard-Piganeau, and J. Eisert (2020), *Quantum* **4**, 314.
- Takeshita, T., N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean (2020), *Phys. Rev. X* **10** (1), 011004.
- Tang, E. (2018), [arXiv:1811.00414](#).
- Tang, E. (2019), [Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing](#), 217.
- Tang, H. L., E. Barnes, H. R. Grimsley, N. J. Mayhall, and S. E. Economou (2019), [arXiv:1911.10205](#).
- Taube, A. G., and R. J. Bartlett (2006), *Int. J. Quantum Chem.* **106** (15), 3393.
- Temme, K., S. Bravyi, and J. M. Gambetta (2017), *Phys. Rev. Lett.* **119** (18), 180509.
- Terhal, B. M. (2015), *Rev. Mod. Phys.* **87** (2), 307.
- Tilly, J., G. Jones, H. Chen, L. Wossnig, and E. Grant (2020), *Phys. Rev. A* **102**, 062425.
- Torlai, G., G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo (2018), *Nat. Phys.* **14** (5), 447.
- Torrontegui, E., and J. J. García-Ripoll (2019), *Eur. Phys. Lett.* **125** (3), 30004.
- Tóth, G., and I. Apellaniz (2014), *J. Phys. A Math. Theor.* **47** (42), 424006.
- Tovey, C. A. (1984), *Discret. Appl. Math.* **8** (1), 85.
- Trabesinger, A. (2012), *Nat. Phys.* **8** (4), 263.
- Tranter, A., P. J. Love, F. Mintert, and P. V. Coveney (2018), *J. Chem. Theory Comput.* **14** (11), 5617.
- Tranter, A., S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm, and P. J. Love (2015), *Int. J. Quantum Chem.* **115** (19), 1431.
- Utkarsh, B. K. Behera, and P. K. Panigrahi (2020), [arXiv:2002.01351](#).
- Utsunomiya, S., K. Takata, and Y. Yamamoto (2011), *Opt. Express* **19** (19), 18091.
- Valiev, M., E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, *et al.* (2010), *Comput. Phys. Commun.* **181** (9), 1477.
- Van Meter, R., T. D. Ladd, A. G. Fowler, and Y. Yamamoto (2010), *Int. J. Quantum Inf.* **8** (01n02), 295.
- Vandewalle, J., and B. De Moor (1991), *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms* (Springer).
- Veis, L., J. Višňák, H. Nishizawa, H. Nakai, and J. Pittner (2016), *Int. J. Quantum Chem.* **116** (18), 1328.
- Venturelli, D., and A. Kondratyev (2019), *Quantum Mach. Intell.* **1**, 17.
- Verdon, G., M. Broughton, and J. Biamonte (2017), [arXiv:1712.05304](#).
- Verdon, G., M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni (2019a), [arXiv:1907.05415](#).

- Verdon, G., J. Marks, S. Nanda, S. Leichenauer, and J. Hidary (2019b), [arXiv:1910.02071](#).
- Verstraete, F., and J. I. Cirac (2005), *J. Stat. Mech. Theory Exp.* **2005** (09), P09012.
- Verstraete, F., J. I. Cirac, and J. I. Latorre (2009), *Phys. Rev. A* **79**, 032316.
- Verstraete, F., J. J. Garcia-Ripoll, and J. I. Cirac (2004), *Phys. Rev. Lett.* **93** (20), 207204.
- Verstraete, F., V. Murg, and J. I. Cirac (2008), *Adv. Phys.* **57** (2), 143.
- Verteletskyi, V., T.-C. Yen, and A. F. Izmaylov (2020), *J. Chem. Phys.* **152** (12), 124114.
- Vidal, G. (2003), *Phys. Rev. Lett.* **91**, 147902.
- Vidal, G. (2004), *Phys. Rev. Lett.* **93**, 040502.
- Vidal, J. G., and D. O. Theis (2019), [arXiv:1901.11434](#).
- Vikstål, P., M. Grönkvist, M. Svensson, M. Andersson, G. Johansson, and G. Ferrini (2020), *Phys. Rev. Appl.* **14**, 034009.
- Vinjanampathy, S., and J. Anders (2016), *Contemp. Phys.* **57** (4), 545.
- Viola, L., and E. Knill (2005), *Phys. Rev. Lett.* **94** (6), 060502.
- Viola, L., E. Knill, and S. Lloyd (1999), *Phys. Rev. Lett.* **82** (12), 2417.
- Volkoff, T., and P. J. Coles (2020), [arXiv:2005.12200](#).
- Vovrosh, J., K. E. Khosla, S. Greenaway, C. Self, M. Kim, and J. Knolle (2021), [arXiv:2101.01690](#).
- Wallman, J. J., and J. Emerson (2016), *Phys. Rev. A* **94** (5), 052325.
- Wan, K. H., O. Dahlsten, H. Kristjánsson, R. Gardner, and M. Kim (2017), *npj Quantum Inf.* **3** (1), 1.
- Wan, Y., R. Jördens, S. Erickson, J. Wu, R. Bowler, T. Tan, P.-Y. Hou, D. Wineland, A. Wilson, and D. Leibfried (2020), [arXiv:2003.03520](#).
- Wang, D., O. Higgott, and S. Brierley (2019), *Phys. Rev. Lett.* **122** (14), 140504.
- Wang, K., Z. Song, X. Zhao, Z. Wang, and X. Wang (2020a), [arXiv:2012.14311](#).
- Wang, L. (2016), *Phys. Rev. B* **94**, 195105.
- Wang, S., E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles (2020b), [arXiv:2007.14384](#).
- Wang, X., Z. Song, and Y. Wang (2020c), [arXiv:2006.02336](#).
- Wang, Y., Y. Li, Z.-q. Yin, and B. Zeng (2018a), *npj Quantum Inf.* **4** (1), 1.
- Wang, Z., S. Hadfield, Z. Jiang, and E. G. Rieffel (2018b), *Phys. Rev. A* **97**, 022304.
- Wang, Z., A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto (2013), *Phys. Rev. A* **88** (6), 063853.
- Watts, P., M. O'Connor, and J. Vala (2013), *Entropy* **15** (6), 1963.
- Wauters, M. M., G. B. Mbeng, and G. E. Santoro (2020a), [arXiv:2003.07419](#).
- Wauters, M. M., E. Panizon, G. B. Mbeng, and G. E. Santoro (2020b), *Phys. Rev. Res.* **2**, 033446.
- Wecker, D., M. B. Hastings, and M. Troyer (2015), *Phys. Rev. A* **92** (4), 042303.
- Wen, J., X. Kong, S. Wei, B. Wang, T. Xin, and G. Long (2019), *Phys. Rev. A* **99** (1), 012320.
- van de Wetering, J. (2020), [arXiv:2012.13966](#).
- White, S. R. (1992), *Phys. Rev. Lett.* **69**, 2863.
- Wierichs, D., C. Gogolin, and M. Kastoryano (2020), [arXiv:2004.14666](#).
- Wierstra, D., T. Schaul, T. Glasmachers, Y. Sun, J. Peters, and J. Schmidhuber (2014), *The J. Mach. Learn. Res.* **15** (1), 949.
- Wille, R., L. Burgholzer, and A. Zulehner (2019), *2019 56th ACM/IEEE Design Automation Conference (DAC)*, 1.
- Wille, R., S. Hillmich, and L. Burgholzer (2020), *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, 1.
- Williams, C. P., and A. G. Gray (1999), *Quantum Computing and Quantum Communications*, 113.
- Woitzik, A. J. C., P. K. Barkoutsos, F. Wudarski, A. Buchleitner, and I. Tavernelli (2020), *Phys. Rev. A* **102** (4), 042402.
- Wright, K., K. Beck, S. Debnath, J. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. Pienti, M. Chmielewski, C. Collins, et al. (2019), *Nat. commun.* **10** (1), 1.
- Xiang, D., X. Wang, C. Jia, T. Lee, and X. Guo (2016), *Chem. Rev.* **116** (7), 4318.
- Xu, X., S. C. Benjamin, and X. Yuan (2019a), [arXiv:1911.05759](#).
- Xu, X., J. Sun, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan (2019b), [arXiv:1909.03898](#).
- Yamamoto, N. (2019), [arXiv:1909.05074](#).
- Yamamoto, Y., T. Leleu, S. Ganguli, and H. Mabuchi (2020), *Appl. Phys. Lett.* **117**, 160501.
- Yang, Z.-C., A. Rahmani, A. Shabani, H. Neven, and C. Chamon (2017), *Phys. Rev. X* **7**, 021027.
- Yao, J., M. Bukov, and L. Lin (2020a), *Proceedings of The First Mathematical and Scientific Machine Learning Conference Proceedings of Machine Learning Research*, **107**, 605.
- Yao, J., P. Köttering, H. Gundlach, L. Lin, and M. Bukov (2020b), [arXiv:2012.06701](#).
- Yao, N. Y., L. Jiang, A. V. Gorshkov, P. C. Maurer, G. Giedke, J. I. Cirac, and M. D. Lukin (2012), *Nat. Commun.* **3** (1), 1.
- Yao, Y.-X., N. Gomes, F. Zhang, T. Iadecola, C.-Z. Wang, K.-M. Ho, and P. P. Orth (2020c), [arXiv:2011.00622](#).
- Yen, T.-C., and A. F. Izmaylov (2020), [arXiv:2007.01234](#).
- Yen, T.-C., V. Verteletskyi, and A. F. Izmaylov (2020), *J. Chem. Theory Comput.* **16** (4), 2400.
- Yordanov, Y. S., D. R. Arvidsson-Shukur, and C. H. Barnes (2020), *Phys. Rev. A* **102** (6), 062612.
- Yoshioka, N., Y. O. Nakagawa, K. Mitarai, and K. Fujii (2020), *Phys. Rev. Res.* **2** (4), 043289.
- Yserentant, H. (2010), *Regularity and Approximability of Electronic Wave Functions*, Lecture Notes in Mathematics (Springer Berlin Heidelberg).
- Yu, S., F. Albarrán-Arriagada, J. C. Retamal, Y.-T. Wang, W. Liu, Z.-J. Ke, Y. Meng, Z.-P. Li, J.-S. Tang, E. Solano, et al. (2019), *Adv. Quantum Technol.* **2** (7-8), 1800074.
- Yuan, X., S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin (2019), *Quantum* **3**, 191.
- Yung, M.-H., J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano (2014), *Sci. Rep.* **4** (1), 3589.
- Zanardi, P., and M. Rasetti (1999), *Phys. Lett. A* **264**, 94.
- Zeng, J., Y. Wu, J.-G. Liu, L. Wang, and J. Hu (2019), *Phys. Rev. A* **99** (5), 052306.
- Zhang, J., T. H. Kyaw, D. Tong, E. Sjöqvist, and L.-C. Kwek (2015), *Sci. Rep.* **5**, 18414.
- Zhang, S., Y. Lu, K. Zhang, W. Chen, Y. Li, J.-N. Zhang, and K. Kim (2020a), *Nat. Commun.* **11**, 587.



- Zhang, Y.-H., P.-L. Zheng, Y. Zhang, and D.-L. Deng (2020b), [Phys. Rev. Lett. \*\*125\*\*, 170501](#).
- Zhang, Z.-J., T. H. Kyaw, J. Kottmann, M. Degroote, and A. Aspuru-Guzik (2021), [Quantum Sci. and Technol.](#) .
- Zhang, Z.-J., J. Sun, X. Yuan, and M.-H. Yung (2020c), [arXiv:2011.05283](#) .
- Zhao, A., A. Tranter, W. M. Kirby, S. F. Ung, A. Miyake, and P. J. Love (2020a), [Phys. Rev. A \*\*101\*\* \(6\), 062322](#).
- Zhao, T., G. Carleo, J. Stokes, and S. Veerapaneni (2020b), [Mach. Learn.: Sci. Technol. \*\*2\*\* \(2\), 02LT01](#).
- Zhong, H.-S., H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, *et al.* (2020), [Science \*\*370\*\* \(6523\), 1460](#).
- Zhou, L., S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin (2020a), [Phys. Rev. X \*\*10\*\*, 021067](#).
- Zhou, Y., E. M. Stoudenmire, and X. Waintal (2020b), [arXiv:2002.07730](#) .
- Zlokapa, A., S. Boixo, and D. Lidar (2020), [arXiv:2005.02464](#) .
- Zoufal, C., A. Lucchi, and S. Woerner (2019), [npj Quantum Inf. \*\*5\*\* \(1\), 1](#).
- Zoufal, C., A. Lucchi, and S. Woerner (2020), [arXiv:2006.06004](#) .
- Zulehner, A., A. Paler, and R. Wille (2018), [IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems \*\*38\*\* \(7\), 1226](#).
- Zulehner, A., and R. Wille (2019), [Proceedings of the 24th Asia and South Pacific Design Automation Conference](#) , 185.

## TABLES OF APPLICATIONS

Algorithm/Application	Proposed implementations
Variational quantum eigensolver (VQE) and related solvers	
VQE	(McClean <i>et al.</i> , 2016; Peruzzo <i>et al.</i> , 2014; Wecker <i>et al.</i> , 2015)
Adaptive VQE	(Grimsley <i>et al.</i> , 2019b; Ryabinkin <i>et al.</i> , 2018b; Sim <i>et al.</i> , 2020)
IQAE	(Bharti, 2020; Bharti and Haug, 2020a)
Krylov approaches	(Huggins <i>et al.</i> , 2020; Jouzdani and Bringuier, 2020; Stair <i>et al.</i> , 2020)
Imaginary time evolution	(Bharti and Haug, 2020b; McArdle <i>et al.</i> , 2019a; Motta <i>et al.</i> , 2020; Sun <i>et al.</i> , 2020c)
VQE for excited states	
Folded spectrum	(Peruzzo <i>et al.</i> , 2014; Ryabinkin <i>et al.</i> , 2018a)
Orthogonally constrained VQE	(Higgott <i>et al.</i> , 2019; Kottmann <i>et al.</i> , 2020b; Lee <i>et al.</i> , 2018)
Subspace expansion and linear-response based	(McClean <i>et al.</i> , 2017; Ollitrault <i>et al.</i> , 2020; Takeshita <i>et al.</i> , 2020)
Subspace-search VQE	(Nakanishi <i>et al.</i> , 2019)
Multistate contracted VQE	(Parrish <i>et al.</i> , 2019a)
Fourier transform of evolutions	(Aleiner <i>et al.</i> , 2020; Roushan <i>et al.</i> , 2017)
WAVES	(Santagati <i>et al.</i> , 2018)
Adiabatically-Assisted	(Garcia-Saez and Latorre, 2018; McClean <i>et al.</i> , 2016)
Hamiltonian simulation	
Variational quantum simulator	(Endo <i>et al.</i> , 2020c; Kubo <i>et al.</i> , 2020; Li and Benjamin, 2017; McArdle <i>et al.</i> , 2019a; Yuan <i>et al.</i> , 2019)
Subspace variational quantum simulator	(Heya <i>et al.</i> , 2019)
Variational fast forwarding	(Cirstoiu <i>et al.</i> , 2020; Commeau <i>et al.</i> , 2020)
Quantum assisted simulator	(Bharti and Haug, 2020b)
Quantum information scrambling and thermalization	
Scrambling	(Holmes <i>et al.</i> , 2020; Joshi <i>et al.</i> , 2020; Landsman <i>et al.</i> , 2019)
Thermal state	(Verdon <i>et al.</i> , 2019b)
Open quantum systems	
Generalized variational quantum simulator	(Endo <i>et al.</i> , 2020c; Liu <i>et al.</i> , 2020b; Yuan <i>et al.</i> , 2019)
Generalized quantum assisted simulator	(Haug and Bharti, 2020)
Trotter simulation	(Hu <i>et al.</i> , 2020; Koppenhöfer <i>et al.</i> , 2020)
State preparation	
Non-equilibrium steady state	(Yoshioka <i>et al.</i> , 2020)(Endo <i>et al.</i> , 2020b; Jaderberg <i>et al.</i> , 2020; Kreula <i>et al.</i> , 2016)
Gibbs-state	(Chowdhury <i>et al.</i> , 2020; Endo <i>et al.</i> , 2020c; Haug and Bharti, 2020)
Many-body ground state	(Ho and Hsieh, 2019; Ho <i>et al.</i> , 2019; Wauters <i>et al.</i> , 2020a)
Quantum autoencoder	
Quantum autoencoder	(Bondarenko and Feldmann, 2020; Bravo-Prieto, 2020; Huang <i>et al.</i> , 2020a,b; Pepper <i>et al.</i> , 2019; Romero <i>et al.</i> , 2017)
Quantum computer-aided design	
Optical setups	(Kottmann <i>et al.</i> , 2020c)
Superconducting circuits	(Kyaw <i>et al.</i> , 2020b)

Table I NISQ algorithms for Many-body physics and chemistry applications from Sec. V.A.

Algorithm/Application	Proposed implementations
Supervised learning	
Quantum kernel methods	(Havlíček <i>et al.</i> , 2019; Kusumoto <i>et al.</i> , 2019; Schuld <i>et al.</i> , 2020b; Schuld and Killoran, 2019)
Variational quantum classifiers (VQC)	(Farhi and Neven, 2018; Lloyd <i>et al.</i> , 2020; Mitarai <i>et al.</i> , 2018; Pérez-Salinas <i>et al.</i> , 2020a; Schuld <i>et al.</i> , 2020a,c; Vidal and Theis, 2019)
Encoding strategies in VQA	(Cervera-Lierta <i>et al.</i> , 2020; Mitarai <i>et al.</i> , 2019)
Quantum reservoir computing	(Chien and Whitfield, 2020; Fujii and Nakajima, 2017; Ghosh <i>et al.</i> , 2019; Mitarai <i>et al.</i> , 2018; Nakajima <i>et al.</i> , 2019; Negoro <i>et al.</i> , 2018; Nokkala <i>et al.</i> , 2020)
Supervised QUBO classifier	(Li <i>et al.</i> , 2018)
Unsupervised learning	
Quantum Boltzmann machines (QBM)	(Amin <i>et al.</i> , 2018; Kieferová and Wiebe, 2017; Zoufal <i>et al.</i> , 2020)
Quantum circuit Born machines (QCBM)	(Alcazar <i>et al.</i> , 2020; Benedetti <i>et al.</i> , 2019a; Coyle <i>et al.</i> , 2020a; Hamilton <i>et al.</i> , 2019; Leyton-Ortega <i>et al.</i> , 2019; Liu and Wang, 2018; Rudolph <i>et al.</i> , 2020)
Quantum generative adversarial networks (QGAN)	(Dallaire-Demers and Killoran, 2018; Hu <i>et al.</i> , 2019; Lloyd and Weedbrook, 2018; Romero and Aspuru-Guzik, 2019; Situ <i>et al.</i> , 2020; Zeng <i>et al.</i> , 2019)
Reinforcement learning	
Reinforcement learning	(Albarrán-Arriagada <i>et al.</i> , 2020; Cárdenas-López <i>et al.</i> , 2018; Chen <i>et al.</i> , 2020; Crawford <i>et al.</i> , 2016; Jerbi <i>et al.</i> , 2019; Lamata, 2017; Lockwood and Si, 2020a,b; Yu <i>et al.</i> , 2019)

Table II NISQ algorithms for machine learning applications from Sec. V.B.

Algorithm/Application	Proposed implementations
Max cut	(Bravyi <i>et al.</i> , 2019; Farhi <i>et al.</i> , 2014; Hastings, 2019; Headley <i>et al.</i> , 2020; Otterbach <i>et al.</i> , 2017)
Max clique	(Arrazola and Bromley, 2018; Banchi <i>et al.</i> , 2020a)
Maximum independent set	(Choi <i>et al.</i> , 2020; Saleem, 2020; Utkarsh <i>et al.</i> , 2020)
Max hafnian	(Arrazola <i>et al.</i> , 2018)
Vertex cover	(Cook <i>et al.</i> , 2019)
Exact cover	(Bengtsson <i>et al.</i> , 2020; Garcia-Saez and Latorre, 2018; Vikstål <i>et al.</i> , 2020)
Knapsack	(de la Grand'rive and Hullo, 2019)
Graph multi-coloring	(Oh <i>et al.</i> , 2019)

Table III NISQ algorithms for combinatorial optimization from Sec. V.C.

Algorithm/Application	Proposed implementations
Factoring	(Anshuetz <i>et al.</i> , 2019; Karamlou <i>et al.</i> , 2020)
SVD	(Bravo-Prieto <i>et al.</i> , 2020; Wang <i>et al.</i> , 2020c)
Linear systems	(Bravo-Prieto <i>et al.</i> , 2019; Huang <i>et al.</i> , 2019; Xu <i>et al.</i> , 2019b)
Non-linear differential equations	(Gaitan, 2020; Haug and Bharti, 2020; Kyriienko <i>et al.</i> , 2020; Lubasch <i>et al.</i> , 2020)

Table IV NISQ algorithms for numerical solvers applications from Sec. V.D.

Algorithm/Application	Proposed implementations
Portfolio optimization	(Bouland <i>et al.</i> , 2020; Cohen <i>et al.</i> , 2020; Egger <i>et al.</i> , 2020a; Marzec, 2016; Rosenberg <i>et al.</i> , 2016; Venturelli and Kondratyev, 2019)
Fraud detection	(Egger <i>et al.</i> , 2020a,b; Zoufal <i>et al.</i> , 2020)
Option pricing	(Kubo <i>et al.</i> , 2020)

Table V NISQ algorithms for finance applications from Sec. V.E

Algorithm/Application	Proposed implementations
Quantum foundations	
Bell inequalities	(Alsina and Latorre, 2016)
Contextuality	(Kirby and Love, 2019; Kirby <i>et al.</i> , 2020)
Variational consistent history (VCH)	(Arrasmith <i>et al.</i> , 2019)
Quantum optimal control	
Quantum optimal control	(Dive <i>et al.</i> , 2018; Li <i>et al.</i> , 2017a; Lu <i>et al.</i> , 2017; Magann <i>et al.</i> , 2021; Policharla and Vinjanampathy, 2020)
Quantum metrology	
Quantum metrology	(Beckey <i>et al.</i> , 2020; Kaubruegger <i>et al.</i> , 2019; Koczor <i>et al.</i> , 2020; Ma <i>et al.</i> , 2020; Meyer <i>et al.</i> , 2020)
Fidelity estimation	
Fidelity estimation	(Cerezo <i>et al.</i> , 2020a)
Quantum error correction (QEC)	
Quantum variational error corrector (QVECTOR)	(Johnson <i>et al.</i> , 2017)
Variational circuit compiler for QEC	(Xu <i>et al.</i> , 2019a)
Nuclear physics	
Nuclear physics	(Avkhadiiev <i>et al.</i> , 2020; Dumitrescu <i>et al.</i> , 2018; Hauke <i>et al.</i> , 2013; Klco <i>et al.</i> , 2018; Kokail <i>et al.</i> , 2019; Liu and Xin, 2020; Martinez <i>et al.</i> , 2016; Roggero <i>et al.</i> , 2020)
Entanglement properties	
Schmidt decomposition	(Bravo-Prieto <i>et al.</i> , 2019; Wang <i>et al.</i> , 2020a)
Multipartite entanglement	(Pérez-Salinas <i>et al.</i> , 2020b)
Entanglement spectrum	(Cerezo <i>et al.</i> , 2020b; LaRose <i>et al.</i> , 2019)

Table VI NISQ algorithm for other applications listed in Sec. V.F.



TABLE OF SOFTWARE PACKAGES

Name	Language	Hardware	Multi-platform	Built-in applications
Cirq ( <a href="#">Quantum AI team and collab., 2020</a> )	python	Yes	No	Chemistry, ML, Noise characterization, Optimization
DQCSim ( <a href="#">QCA lab, 2020</a> )	python, C++, Rust	No	No	–
IQS ( <a href="#">Guerreschi et al., 2020</a> )	C++	No	No	QAOA
<i> Lib)</i> ( <a href="#">Möller and Schalkers, 2020</a> )	python, C++	Yes**	Yes	–
PennyLane ( <a href="#">Bergholm et al., 2018</a> )	python	Yes**	Yes	ML, Optimization
ProjectQ ( <a href="#">Steiger et al., 2018</a> )	python	Yes**	Yes	Fermionic simulation, Optimization
pyquil/Forest ( <a href="#">Smith et al., 2016</a> )	python, Lisp	Yes	No	VQE, QAOA, Noise characterization, Optimization
QDK ( <a href="#">Quantum, 2020</a> )	python, C#, Q#	Yes**	No	Chemistry, Optimization
Qibo ( <a href="#">Efthymiou et al., 2020</a> )	python	Yes*	Yes	VQE, QAOA, Adiabatic evolution, Optimization
Qiskit ( <a href="#">IBMQ, 2019</a> )	python	Yes	No	Chemistry, ML, Optimization, Finance, Noise characterization
QTensor ( <a href="#">Lykov, 2020</a> )	python, C++	No	No	Tensor-Network simulator, quantum circuit simulator, QAOA
QuEST ( <a href="#">Jones and Benjamin, 2020</a> ; <a href="#">Jones et al., 2019a</a> )	C++, Mathematica	No	Yes	–
Quimb ( <a href="#">Gray, 2018</a> )	python	No	No	Tensor-Network simulator, quantum circuit simulator
Qulacs ( <a href="#">Suzuki et al., 2020</a> )	python, C++	No	No	Simulator, noise characterization
StrawberryFields ( <a href="#">Killoran et al., 2019</a> )	python	Yes*	No	GBS
Tequila ( <a href="#">Kottmann et al., 2020a</a> )	python	Yes**	Yes	Chemistry, ML, Noise characterization, Optimization
XACC ( <a href="#">McCaskey et al., 2020</a> )	python, C++	Yes**	Yes	Optimization, VQE, QAOA, RBM and other algorithms
Yao ( <a href="#">Luo et al., 2020</a> )	Julia	No	No	–

Table VII List of open-source quantum software packages. These packages are designed using common computing languages such as python, C++ or Julia. Some of them can be used in real quantum hardware, either because the developers are also building these devices or because the package include other quantum packages. Some of them can translate their code to other quantum packages or simulators. \*Hardware not publicly available yet. \*\*Not on own hardware but can be run in a hardware backend.

TABLE OF EXTERNAL LIBRARIES

Name	Language	Application
OpenFermion ( <a href="#">McClellan et al., 2020a</a> )	python	Chemistry
psi4 ( <a href="#">Smith et al., 2020</a> )	python	Chemistry
PySCF ( <a href="#">Sun et al., 2018, 2020b</a> )	python	Chemistry
NWChem ( <a href="#">Valiev et al., 2010</a> )	Fortran 77, C	Chemistry
EntropicaQAOA ( <a href="#">Entropica Labs, 2019</a> )	python	QAOA
TensorFlowQ ( <a href="#">Broughton et al., 2020</a> )	python	QML
TensorFlow ( <a href="#">Abadi et al., 2016</a> )	python	ML
Mitig ( <a href="#">LaRose et al., 2020</a> )	python	Error Mitigation
pyzx ( <a href="#">Kissinger and van de Wetering, 2020a</a> )	python	Compiler
iTensor ( <a href="#">Fishman et al., 2020</a> )	Julia, C++	Tensor Networks
OpenQL ( <a href="#">QCA Lab, 2020</a> )	python, C++	Compiler
JKQ ( <a href="#">Wille et al., 2020</a> )	C++	Simulator and Compiler
ScaffCC ( <a href="#">JavadiAbhari et al., 2014</a> )	Scaffold	Compiler
staq ( <a href="#">Amy and Gheorghiu, 2020</a> )	C++	Compiler
Silq ( <a href="#">Bichsel et al., 2020</a> )	D	Compiler
QX simulator ( <a href="#">QCA lab, 2020</a> )	python, C++	Simulator
QRACK ( <a href="#">Strano and the Qrack contributors, 2020</a> )	C++	Simulator
QuTip ( <a href="#">Johansson et al., 2012</a> )	python	Quantum Info. SDK
Q-Ctrl ( <a href="#">Q-Ctrl, 2020</a> )	python	Quantum Control

Table VIII External open-source libraries useful for the NISQ era. These libraries have applications in chemistry, machine learning, circuit compilation and quantum control. Some of them are integrated in the quantum software packages listed in Table VII.