REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
05-07-2023	Final Report	15-Jan-2020 - 14-Mar-2023

4. TITLE AND SUBTITLE
Final Report: Quantum Algorithms for Machine Learning, Optimization, and Simulation

5a. CONTRACT NUMBER
W911NF-20-1-0014

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
IBM Corporation (T.J. Watson Research Lal
1101 Kitchawan Road

Yorktown Heights, NY 10598 -0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
76331-PE-OC.11

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Andrew Cross
UU	UU	UU	UU		19b. TELEPHONE NUMBER 191-494-5288

RPPR Final Report
as of 06-Jul-2023

Agency Code: 21XD

Proposal Number: 76331PEQC **Agreement Number: W911NF-20-1-0014**
INVESTIGATOR(S):

Name: Andrew Cross
Email: awcross@us.ibm.com
Phone Number: 19149452887
Principal: Y

Name: Sergey Bravyi
Email: sbravyi@us.ibm.com
Phone Number: 9149451580
Principal:

Organization: **IBM Corporation (T.J. Watson Research Laboratory)**
Address: 1101 Kitchawan Road, Yorktown Heights, NY 105980001
Country: USA
DUNS Number: 084006741 EIN: 130871985

Report Date: 14-Apr-2023 Date Received: 05-Jul-2023
Final Report for Period Beginning 15-Jan-2020 and Ending 14-Mar-2023
Title: Quantum Algorithms for Machine Learning, Optimization, and Simulation
Begin Performance Period: 15-Jan-2020 **End Performance Period:** 14-Mar-2023
Report Term: 0-Other

Submitted By: Amy Palmer Email: amy2@us.ibm.com
Phone: (720) 818-0651

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 0 **STEM Participants:**

Major Goals: please see the uploaded pdf

Accomplishments: please see the uploaded pdf

Training Opportunities: Nothing to Report

Results Dissemination: please see the uploaded pdf

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: please see the uploaded pdf

PARTICIPANTS:

Participant Type: PD/PI
Participant: Sergey Bravyi
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Co PD/PI
Participant: Andrew Cross
Person Months Worked: 1.00 **Funding Support:**

Project Contribution:
National Academy Member: N

Participant Type: Co-Investigator
Participant: Kristan Temme
Person Months Worked: 10.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Co-Investigator
Participant: Srinivasan Arunachalam
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Co PD/PI
Participant: David Gosset
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Co PD/PI
Participant: Robert Koenig
Person Months Worked: 15.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Co-Investigator
Participant: Giacomo Nannicini
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
National Academy Member: N

ARTICLES:

RPPR Final Report
as of 06-Jul-2023

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 2-Awaiting Publicat
Journal: Physical Review Letters
Publication Identifier Type: Publication Identifier: https://doi.org/10.1103/PhysRevLett.125.260
Volume: 125 Issue: First Page #: 260505
Date Submitted: 5/21/21 12:00AM Date Published: 12/3/20 5:00AM
Publication Location:
Article Title: Obstacles to State Preparation and Variational Optimization from Symmetry Protection
Authors: Eugene Tang, Robert Koenig, Alexander Kliesch, Sergey Bravyi
Keywords: quantum optimization algorithms
Abstract: Over the past few years several quantum machine learning algorithms were proposed that promise quantum speed-ups over their classical counterparts. Most of these learning algorithms either assume quantum access to data -- making it unclear if quantum speed-ups still exist without making these strong assumptions, or are heuristic in nature with no provable advantage over classical algorithms. In this paper, we establish a rigorous quantum speed-up for supervised classification using a general-purpose quantum learning algorithm that only requires classical access to data. Our quantum classifier is a conventional support vector machine that uses a fault-tolerant quantum computer to estimate a kernel function. Data samples are mapped to a quantum feature space and the kernel entries can be estimated as the transition amplitude of a quantum circuit.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 2-Awaiting Publicat
Journal: Proceedings of 61st Annual IEEE Symposium on Foundation of Computer Science (FOCS 2020)
Publication Identifier Type: DOI Publication Identifier: 10.1109/FOCS46700.2020.00069
Volume: Issue: First Page #:
Date Submitted: 5/21/21 12:00AM Date Published:
Publication Location:
Article Title: Sample-efficient learning of quantum many-body systems
Authors: Mehdi Soleimanifar, Tomotaka Kuwahara, Srinivasan Arunachalam, Anurag Anshu
Keywords: quantum Hamiltonian complexity, machine learning, quantum algorithms
Abstract: Over the past few years several quantum machine learning algorithms were proposed that promise quantum speed-ups over their classical counterparts. Most of these learning algorithms either assume quantum access to data -- making it unclear if quantum speed-ups still exist without making these strong assumptions, or are heuristic in nature with no provable advantage over classical algorithms. In this paper, we establish a rigorous quantum speed-up for supervised classification using a general-purpose quantum learning algorithm that only requires classical access to data. Our quantum classifier is a conventional support vector machine that uses a fault-tolerant quantum computer to estimate a kernel function. Data samples are mapped to a quantum feature space and the kernel entries can be estimated as the transition amplitude of a quantum circuit.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 06-Jul-2023

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 2-Awaiting Publicat
Journal: Integer Programming and Combinatorial Optimization (IPCO) 2021
Publication Identifier Type: Other Publication Identifier: Proceedings of IPCO 2021
Volume: 22 Issue: First Page #:
Date Submitted: 5/21/21 12:00AM Date Published:
Publication Location:
Article Title: Fast quantum subroutines for the simplex method
Authors: Giacomo Nannicini
Keywords: quantum optimization algorithms
Abstract: Over the past few years several quantum machine learning algorithms were proposed that promise quantum speed-ups over their classical counterparts. Most of these learning algorithms either assume quantum access to data -- making it unclear if quantum speed-ups still exist without making these strong assumptions, or are heuristic in nature with no provable advantage over classical algorithms. In this paper, we establish a rigorous quantum speed-up for supervised classification using a general-purpose quantum learning algorithm that only requires classical access to data. Our quantum classifier is a conventional support vector machine that uses a fault-tolerant quantum computer to estimate a kernel function. Data samples are mapped to a quantum feature space and the kernel entries can be estimated as the transition amplitude of a quantum circuit.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Nature Physics
Publication Identifier Type: Publication Identifier:
Volume: 17 Issue: First Page #: 1013
Date Submitted: 2/4/22 12:00AM Date Published:
Publication Location:
Article Title: A rigorous and robust quantum speed-up in supervised machine learning
Authors: Yunchao Liu, Srinivasan Arunachalam, Kristan Temme
Keywords: quantum machine learning, kernel methods
Abstract: Over the past few years several quantum machine learning algorithms were proposed that promise quantum speed-ups over their classical counterparts. Most of these learning algorithms either assume quantum access to data -- making it unclear if quantum speed-ups still exist without making these strong assumptions, or are heuristic in nature with no provable advantage over classical algorithms. In this paper, we establish a rigorous quantum speed-up for supervised classification using a general-purpose quantum learning algorithm that only requires classical access to data. Our quantum classifier is a conventional support vector machine that uses a fault-tolerant quantum computer to estimate a kernel function. Data samples are mapped to a quantum feature space and the kernel entries can be estimated as the transition amplitude of a quantum circuit.
Distribution Statement: 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 06-Jul-2023

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 0-Other
Journal: arXiv:2009.11270
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: Simpler (classical) and faster (quantum) algorithms for Gibbs partition functions
Authors: Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, Pawel Wocjan
Keywords: quantum simulations, quantum walks, simulated annealing
Abstract: We consider the problem of approximating the partition function of a classical Hamiltonian using simulated annealing. This requires the computation of a cooling schedule, and the ability to estimate the mean of the Gibbs distributions at the corresponding inverse temperatures. We propose classical and quantum algorithms for these two tasks, achieving two goals: (i) we simplify the seminal work of Stefankovic, Vempala and Vigoda, improving their running time and almost matching that of the current classical state of the art; (ii) we quantize our new simple algorithm, improving upon the best known algorithm for computing partition functions of many problems, due to Harrow and Wei. A key ingredient of our method is the paired-product estimator of Huber. The proposed quantum algorithm has two advantages over the classical algorithm: it has quadratically faster dependence on the spectral gap of the Markov chains as well as the precision, and it computes a shorter cooling schedule.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 0-Other
Journal: PRX Quantum
Publication Identifier Type: DOI Publication Identifier: https://doi.org/10.1103/PRXQuantum.3.0103
Volume: 3 Issue: First Page #: 010309
Date Submitted: 2/4/22 12:00AM Date Published:
Publication Location:
Article Title: Doubling the size of quantum simulators by entanglement forging
Authors: Andrew Eddins, Mario Motta, Tanvi P. Gujarati, Sergey Bravyi, Antonio Mezzacapo, Charles Hadfield, S
Keywords: quantum chemistry simulation
Abstract: Quantum computers are promising for simulations of chemical and physical systems, but the limited capabilities of today's quantum processors permit only small, and often approximate, simulations. Here we present a method, classical entanglement forging, that harnesses classical resources to capture quantum correlations and double the size of the system that can be simulated on quantum hardware. Shifting some of the computation to classical post-processing allows us to represent ten spin-orbitals on five qubits of an IBM Quantum processor to compute the ground state energy of the water molecule in the most accurate simulation to date. We discuss conditions for applicability of classical entanglement forging and present a roadmap for scaling to larger problems.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 06-Jul-2023

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Advances in Neural Information Processing Systems 34 (2021)
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 2/4/22 12:00AM Date Published:
Publication Location:
Article Title: Private learning implies quantum stability
Authors: Srinivasan Arunachalam, Yihui Quek, John Smolin
Keywords: quantum learning
Abstract: We present two quantum interior point methods for semidefinite optimization problems, building on recent advances in quantum linear system algorithms. The first scheme, more similar to a classical solution algorithm, computes an inexact search direction and is not guaranteed to explore only feasible points; the second scheme uses a nullspace representation of the Newton linear system to ensure feasibility even with inexact search directions. The second is a novel scheme that might seem impractical in the classical world, but it is well-suited for a hybrid quantum-classical setting. We show that both schemes converge to an optimal solution of the semidefinite optimization problem under standard assumptions. By comparing the theoretical performance of classical and quantum interior point methods with respect to various input parameters, we show that our second scheme obtains a speedup over classical algorithms in terms of the dimension of the problem.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 0-Other
Journal: arXiv:2110.15466
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: On the complexity of quantum partition functions
Authors: Sergey Bravyi, Anirban Chowdhury, David Gosset, Pawel Wocjan
Keywords: quantum algorithms, quantum Hamiltonian complexity
Abstract: The partition function and free energy of a quantum many-body system determine its physical properties in thermal equilibrium. Here we study the computational complexity of approximating these quantities for n-qubit local Hamiltonians. First, we report a classical algorithm with poly(n) runtime which approximates the free energy of a given 2-local Hamiltonian provided that it satisfies a certain denseness condition. Secondly, we establish polynomial-time equivalence between the problem of approximating the free energy of local Hamiltonians and three other natural quantum approximate counting problems, including the problem of approximating the number of witness states accepted by a QMA verifier. Finally, we summarize state-of-the-art classical and quantum algorithms for approximating the free energy and show how to improve their runtime and memory footprint.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support:

RPPR Final Report
as of 06-Jul-2023

Publication Type: Journal Article Peer Reviewed: **Publication Status:** 0-Other
Journal: arXiv:2112.06025
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: Date Published:
Publication Location:
Article Title: Quantum Interior Point Methods for Semidefinite Optimization
Authors: Brandon Augustino, Giacomo Nannicini, Tamás Terlaky, Luis F. Zuluaga
Keywords: quantum algorithms, optimization, semidefinite programs
Abstract: We present two quantum interior point methods for semidefinite optimization problems, building on recent advances in quantum linear system algorithms. The first scheme, more similar to a classical solution algorithm, computes an inexact search direction and is not guaranteed to explore only feasible points; the second scheme uses a nullspace representation of the Newton linear system to ensure feasibility even with inexact search directions. The second is a novel scheme that might seem impractical in the classical world, but it is well-suited for a hybrid quantum-classical setting. We show that both schemes converge to an optimal solution of the semidefinite optimization problem under standard assumptions. By comparing the theoretical performance of classical and quantum interior point methods with respect to various input parameters, we show that our second scheme obtains a speedup over classical algorithms in terms of the dimension of the problem.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support:

Partners

,

I certify that the information in the report is complete and accurate:
Signature: Sergey Bravyi
Signature Date: 7/5/23 1:45PM

Quantum Algorithms for Machine Learning, Optimization, and Simulation (QAMLOS)

PIs: Sergey Bravyi (sbravyi@us.ibm.com), Andrew Cross (awcross@us.ibm.com)

Contract number W911NF-20-1-0014

Senior Personnel:
Srinivasan Arunachalam (IBM)
David Gosset (University of Waterloo, Perimeter Institute for Theoretical Physics)
Robert Koenig (Technical University of Munich)
Giacomo Nannicini (IBM, presently at University of Southern California)
Kristan Temme (IBM)

Final Report

Research Term: January 15, 2020 till March 15, 2023

Abstract: In this report we survey quantum algorithms developed by the team in the areas of machine learning, optimization, and simulation of physical systems. First, we consider purely classical computational tasks such as classification problems for data with group structure, learning of Boolean functions, combinatorial optimization problems including MaxCut and graph k-coloring, linear programming, and statistical inference based on Markov Chain Monte Carlo algorithms. For each of these problems we give a quantum algorithm improving upon the state-of-the-art. Secondly, we consider problems relevant for chemistry and material science where the goal is to simulate certain properties of a quantum many-body system. This includes ground state preparation for Hamiltonians describing electronic structure of molecules and systems with a topological quantum order that host non-abelian anyons. We show how to tackle these simulation problems more efficiently by combining classical and quantum computational resources. Finally, we investigate computational complexity of simulating quantum many-body systems at the thermal equilibrium and improve existing algorithms for simulation of open quantum systems whose dynamics is governed by the Lindblad master equation.

Contents

1 Introduction

Quantum computing has remained a theoretical possibility for many years. Recent advances in hardware enabled demonstrations of quantum circuits at a scale which is already at the border of what can be simulated with the existing classical computers. It is therefore of central importance to assess the potential impact of near-term quantum computers on real problems of interest. The primary goal of this program was to design and analyze quantum machine learning, optimization, and simulation algorithms with the emphasis on near-term applications. Our approach is aimed at directly addressing the fact that near-term quantum computers are subject to decoherence and lack error correction capabilities. These computers are limited to shallow circuit depths and are unlikely to have the ability to load vast amounts of data from a classical memory into a quantum memory. In this setting, how could a near-term quantum algorithm look like?

Many current proposals for attaining a quantum computational advantage make several assumptions about how the data is provided to the quantum computer. Some algorithms require the availability of so-called quantum RAM (QRAM), where a data set can be loaded

in a coherent superposition [1]. In particular an n-qubit QRAM provides coherent access to 2^n data points in superposition. This appears to open up the tantalizing possibility of processing large amounts of data with only logarithmic complexity. However, this appearance may be misleading for several reasons [2]. The process of converting classical data to quantum data in the QRAM negates any quantum speed-up gained at a later stage. So is there still a place where quantum algorithms can improve on the current state-of-the-art?

For a quantum algorithm to provide a meaningful advantage in machine learning we accept that all input data is given to the programmer in a classical form and that the programmer will most likely have to look at every sample used in the algorithm, just as in the classical case. We adopt quantum kernels framework [3] for designing hybrid quantum-classical learning algorithms. It is aimed at avoiding the use of quantum RAM, yet still retaining a quantum advantage. The team has demonstrated that certain classically hard classification problems in supervised learning can indeed be solved in polynomial time using the quantum kernels framework [4, 5]. To this end we introduce a family of covariant quantum kernels for data with group structure, see Section 2.

The success of machine learning algorithms in analyzing high-dimensional data, has resulted in a surge of interest in applying these algorithms for studying quantum many-body systems whose description requires dealing with an exponentially large state space [6]. One important problem in this direction is the quantum Hamiltonian learning problem [7, 8, 9, 10, 11, 12]. Here the goal is to obtain an approximate description of the Hamiltonian of a quantum many-body system given samples from its thermal Gibbs state. The classical analog of this problem, known as learning graphical models or Boltzmann machines, is a well-studied question in machine learning and statistics [13, 14, 15, 16, 17]. We present the first efficient quantum algorithm for learning Hamiltonians with local interactions [18], see Section 2. This algorithm was further improved in [19] achieving sample complexity $O(\log n)$ and gate complexity $O(n)$ for learning a local Hamiltonian describing a system of n qubits. Our work paves the way toward a more rigorous application of machine learning techniques to quantum many-body problems.

Boolean functions are the bread and butter of computational complexity theory. Among their uses is to serve as simple toy models for complex problems, allowing rigorous proofs. For instance, learning to classify data may be done effectively but heuristically with a neural network, but the PAC learning framework [20] can prove theorems. We propose simple toy models of learning Boolean functions with the aid of quantum computation [21]. In particular, we give quantum learning algorithms with $poly(n)$ query and gate complexity for learning low-degree Boolean polynomials with n variables encoded into n-qubit phase states, see Section 2.

QAOA optimization algorithms pioneered by Farhi et al. [22] provide a natural arena where near-term quantum devices may have an impact. However, recent work [23] indicates that QAOA based on bounded-depth variational circuits do not compare favorably to analogous "local" classical optimizers. On the other hand, it has been shown that multi-level QAOA based on deep variational circuits can replicate the quantum speedup achieved by Grover search algorithm [24]. Unfortunately, multi-level QAOA may be beyond the reach of near-term devices. Furthermore, its performance on practically relevant problems is hard to assess due to the difficulty of classical simulation and the lack of analytic tools.

Instead, here we study QAOA with $O(1)$ levels and examine whether its computational

power can be enhanced using a novel technique which we call variable elimination [25]. Specifically, we seek to mimic classical optimization algorithms based on rounding of linear programming relaxations. Instead of rounding individual variables obtained by solving a suitable relaxation of the problem, we propose rounding correlations among variables obtained from the optimal QAOA solution. Applying the correlation rounding method to cost functions describing MaxCut and Max-k-Cut optimization problems allows one to eliminate variables reducing a problem instance with n variables to the one with $n-1$ or fewer variables [25, 26]. This leads to a recursive version of QAOA described in details in Section 3. We also present a twisted version of QAOA in which an extra classical processing step is folded into the definition of the cost function Hamiltonian [27]. The twisted QAOA comes with rigorous performance guarantees, see Section 3.

A key development that may lead to faster quantum optimization algorithms is the discovery of quantum linear system algorithms that can be much faster than their classical counterpart under suitable conditions [28, 29]. Here we study applications of quantum linear system algorithms for classical convex optimization. More specifically, we study the simplex method for solving linear programming problems — one of the most important algorithms in the history of optimization, that is still widely used today, see Section 3. For problems with a well-conditioned sparse constraint matrix we give a quantum pricing algorithm that achieves a provable asymptotic speedup over best known classical algorithms [30]. Importantly, this speedup does not depend on the data being available in a "quantum form" such as QRAM; the input of our quantum subroutines is the natural classical description of the problem, and the output is the index of the variables that should leave or enter the basis.

Simulating quantum systems is an especially hard task for classical computers, making the realization of quantum computers potentially revolutionary for the study of chemistry, materials science, and fundamental physics. However, techniques like quantum phase estimation, which promises accurate chemical simulations, require hardware well beyond the present state of the art. While hardware capabilities continue to steadily advance, limitations on both quantity and quality of qubits are giving rise to a new family of algorithms that leverage additional classical resources to assist quantum computations[31, 32, 33, 34, 35, 36, 37, 38, 39]. Section 4 reports new examples of such classically enhanced quantum simulation algorithms.

First, we propose entanglement forging method for simulating ground state properties of quantum many-body systems [40]. It works by partitioning a problem into weakly interacting clusters and classically correlating the simulation results measured on each cluster. A range of important systems naturally possess suitable partitions, including low-energy eigenstates of chemical [41] and lattice-model [42, 43, 44] Hamiltonians, systems embedded in a quantum bath [45, 46], and static correlations associated with chemical bond-breaking processes [47, 48, 49].

Secondly, we consider simulation of exotic phases of matter harboring non-abelian anyons. In contrast to ordinary particles such as bosons or fermions whose state can be fully specified by the position (or momentum) of each particle, a quantum state describing non-abelian anyons features non-local topological degrees of freedom. The latter serve as a quantum error correcting code with a macroscopic distance. Braiding a pair of non-abelian anyons can enact a non-trivial unitary transformation on the topological degrees of freedom. This motivates proposals for using non-abelian anyons for reliable quantum information processing [50]. We present constant-depth quantum circuits for preparing the ground state of topologically

ordered systems and realizing braiding of non-abelian anyons [51]. To overcome known no-go results, our approach relies on dynamic circuits [52]. These are quantum circuits that combine regular unitary gates with mid-circuit measurement, real-time classical computing, and feed-forward such that each gate in the circuit can be classically controlled by the outcomes of all previous measurement outcomes. These results demonstrate that the ability to run quantum and classical computation in parallel can help reducing quantum circuit depth required for simulation of non-abelian anyons. We note that several groups have made similar proposals culminating in the recent experimental demonstration of a non-abelian braiding statistics [53], see Section 4 for the literature review.

Many properties of a quantum many-body system in thermal equilibrium are determined by the Helmholtz free energy proportional to the logarithm of the partition function. The ability to calculate the free energy and its derivatives with respect to the temperature and Hamiltonian parameters such as external fields is instrumental for mapping out the phase diagram of the system and predicting physical properties of each phase. Accordingly, the problem of estimating the free energy has been extensively studied in the physics community. Here we examine this problem through the lens of quantum Hamiltonian complexity theory, see Section 4. We show that a suitably formalized version of the free energy problem is computationally equivalent to Quantum Approximate Counting — estimating the number of accepting witnesses for a QMA verifier [54]. We establish a surprising connection between the free energy problem and the one of computing the Kronecker coefficients of the symmetric group [55]. We also give an efficient quantum algorithm for computing the Kronecker coefficients in a certain special case.

These and several related results by the team members are described in greater detail in the following sections that cover machine learning algorithms (Section 2), optimization problems (Section 3), and simulation of physical systems (Section 4). Each section is equipped with a short introduction serving as a guide for more technical sub-sections focusing on specific computational problems and quantum algorithms. A complete list of publications, preprints, and patents that resulted from the program can be found in Section 5.

2 Quantum machine learning algorithms

Much attention has been drawn towards establishing a quantum advantage in machine learning due to its wide applicability [56, 57, 58, 59, 60]. In this direction there have been several quantum algorithms for machine learning tasks that promise polynomial and exponential speed-ups. A family of such quantum algorithms assumes that classical data is encoded in amplitudes of a quantum state, which uses a number of qubits that is only logarithmic in the size of the dataset. These quantum algorithms are therefore able to achieve exponential speed-ups over classical approaches [61, 62, 63, 64, 65, 66, 67, 68, 69, 70]. However, it is not known whether data can be efficiently provided this way in practically relevant settings. This raises the question of whether the advantage comes from the quantum algorithm, or from the way data is provided [2]. Indeed, recent works have shown that if classical algorithms have an analogous sampling access to data, then some of the proposed exponential speed-ups no longer exist [71, 72, 73, 74, 75, 76].

Consequently a different class of quantum algorithms has been developed which only

assumes classical access to data. Most of these algorithms use variational circuits for learning, where a candidate circuit is selected from a parameterized circuit family via classical optimization [77, 78, 79, 80, 81]. Although friendly to experimental implementation, these algorithms are heuristic in nature since no formal evidence has been provided which shows that they have a genuine advantage over classical algorithms. An important challenge is therefore to find one example of such a heuristic quantum machine learning algorithm, which given *classical access* to data can *provably* outperform all classical learners for some learning problem. Our work [4, 5], summarized in Sections 2.1, 2.2 address this challenge using quantum kernel methods.

Section 2.3, based on Ref. [18], gives the first sample-efficient algorithm for the quantum Hamiltonian learning problem. In particular, it shows that polynomially many samples in the number of particles (qudits) are sufficient for learning the parameters of a geometrically local Hamiltonian, given samples of the thermal Gibbs state.

A natural next question is: what are other classes of quantum states ubiquitous in quantum computing can be learned efficiently? Known examples of efficiently learnable states include Matrix Product States describing weakly entangled quantum spin chains [82], output states of Clifford circuits [83], symmetric states invariant under qubit permutations [84], see the recent survey [85] for other examples. Section 2.4, based on Ref. [21], provides a new class of efficiently learnable quantum states: phase states associated with (generalized) Boolean functions. Such states have the form $|\psi_f\rangle = 2^{-n/2} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle$, where $f(x)$ is a low-degree Boolean polynomial. We show that an unknown polynomial f of bounded degree can be learned by performing measurements on $poly(n)$ copies of the phase state $|\psi_f\rangle$.

2.1 Provable quantum advantage in machine learning

In the work [4] we show that an exponential quantum speed-up can be obtained via the use of a quantum-enhanced feature space [86, 87], where each data point is mapped nonlinearly to a quantum state and then classified by a linear classifier in the high dimensional Hilbert space. To efficiently learn a linear classifier in feature space from training data, we use the standard kernel method in *support vector machines* (SVMs), a well-known family of supervised classification algorithms [88, 89]. We obtain the kernel matrix by measuring the pairwise inner product of the feature states on a quantum computer, a procedure we refer to as *quantum kernel estimation* (QKE). This kernel matrix is then given to a classical optimizer that efficiently finds the linear classifier that optimally separates the training data in feature space by running a convex quadratic program.

The advantage of our quantum learner stems from its ability to recognize classically intractable complex patterns using the feature map. We prove an end-to-end quantum advantage based on this intuition, where our quantum classifier is guaranteed to achieve high accuracy for a classically hard classification problem. We show that under a suitable quantum feature map, the classical data points, which are indistinguishable from having random labels by efficient classical algorithms, are linearly separable with a large margin in high-dimensional Hilbert space. Based on this property, we then combine ideas from classic results on the generalization of soft margin classifiers [90, 91, 92, 93] to rigorously bound the misclassification error of the SVM-QKE algorithm. The optimization for large margin classifiers in the SVM program is crucial in our proof, as it allows us to learn the optimal

separator in the exponentially large feature space, while also making our quantum classifier robust against additive sampling errors.

Our classification problem that shows the exponential quantum speed-up is constructed based on the discrete logarithm problem (DLP). We prove that no efficient classical algorithm can achieve an accuracy that is inverse-polynomially better than random guessing, assuming the widely-believed classical hardness of DLP. In computational learning theory, the use of one-way functions for constructing classically hard learning problems is a well-known technique [94]. Rigorous separations between quantum and classical learnability have been established using this idea in the quantum oracular and PAC model [95, 57], as well as in the classical generative setting [96]. There the quantum algorithms are constructed specifically to solve the problems for showing a separation, and in general are not applicable to other learning problems. Based on different complexity-theoretic assumptions, evidences of an exponential quantum speed-up were shown for a quantum generative model [97], where the overall performance is not guaranteed.

Quantum kernel estimation. The core component in quantum kernel methods [86, 87] that leads to its ability to outperform classical learners is the *quantum feature map*, which maps a classical data point $x \in \mathbb{R}^d$ non-linearly to a n-qubit quantum state $\Phi(x) = |\phi(x)\rangle\langle\phi(x)|$. Here the feature state is the projector onto $|\phi(x)\rangle = U(x)\,|0^n\rangle$ and is constructed by a parameterized circuit family $\{U(x)\}$. In practice, given an arbitrary dataset, one can heuristically choose a parameterized circuit family to construct the feature map, and then run the training procedure described below to learn a linear classifier in feature space. Therefore, quantum kernel methods represent a family of general-purpose supervised learning algorithms that is widely applicable, just like classical kernels being ubiquitous in real world applications [98].

In learning algorithms, feature maps play the role of pattern recognition: the intrinsic labeling patterns for data, which are hard to recognize in the original space, become easy to identify once mapped to a high-dimensional feature space. The idea of applying a high-dimensional feature map to reduce a complex pattern recognition problem to linear classification is not new, and has been the foundation of a family of supervised learning algorithms called support vector machines (SVMs) [89]. Consider a general feature map $\Phi : \mathcal{X} \to \mathcal{H}$ that maps data to a feature space $\mathcal{H}$ associated with an inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. To find a linear classifier in $\mathcal{H}$, we consider the convex quadratic program

$$\min_{w,\xi} \frac{1}{2}\|w\|_2^2 + \frac{\lambda}{2}\sum_{i=1}^{m}\xi_i^2 \quad \text{s.t. } y_i \cdot \langle\Phi(x_i), w\rangle_{\mathcal{H}} \geq 1 - \xi_i \tag{1}$$

where $\xi_i \geq 0$. Here $\lambda > 0$ is a constant, $\{(x_i, y_i)\}_{i=1}^m \subseteq \mathbb{R}^d \times \{\pm 1\}$ is the training set, w is a vector normal to a hyperplane in $\mathcal{H}$ which defines a linear classifier $y = \text{sign}\left(\langle\Phi(x), w\rangle_{\mathcal{H}}\right)$, and ξ_i are slack variables used in the soft margin constraints. Intuitively, this program optimizes for the hyperplane that maximally separates +1/-1 labeled data. Note that (1) is efficient in the dimension of $\mathcal{H}$. However, once we map to a high-dimensional feature space, it takes exponential time to find the optimal hyperplane. The main insight which leads to the success of SVMs is that this problem can be solved by running the dual program of Eq. (1). It reads

$$\max_{\alpha \in \mathbb{R}_+^m} F(\alpha) \tag{2}$$

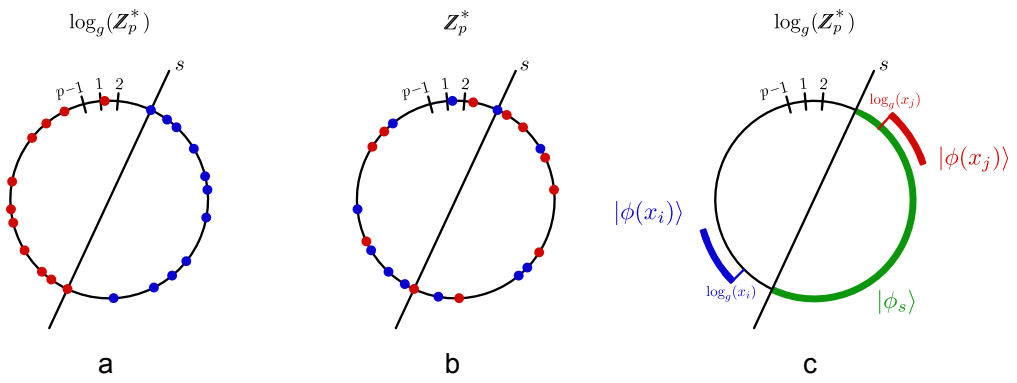


Figure 1: Learning the concept class $\mathcal{C}$ by a quantum feature map. (**a**) After taking the discrete log of the data samples, they become separated in log space by the concept s. (**b**) However, in the original data space, the data samples look like randomly labeled and cannot be learned by an efficient classical algorithm. (**c**) Using the quantum feature map, each $x \in \mathbb{Z}_p^*$ is mapped to a quantum state $|\phi(x)\rangle\langle\phi(x)|$, which corresponds to a uniform superposition of an interval in log space starting with $\log_g x$. This feature map creates a large margin, as the $+1$ labeled example (red interval) has high overlap with a separating hyperplane (green interval), while the -1 labeled example (blue interval) has zero overlap.

with

$$F(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \frac{1}{2\lambda} \sum_{i=1}^{m} \alpha_i^2, \tag{3}$$

where $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle_{\mathcal{H}}$ is the *kernel matrix*. This dual program, which returns a linear classifier defined as $y(x) = \text{sign}\left(\sum_{i=1}^{m} \alpha_i y_i K(x, x_i)\right)$, is equivalent to the original program as guaranteed by strong duality. Effectively, this means that we can do optimization in the high-dimensional feature space efficiently, as long as the kernel $K(x_i, x_j)$ can be efficiently computed.

The same insight can be applied to quantum feature maps: to utilize the full power of the quantum feature space, it suffices to compute the Hilbert-Schmidt inner products $\text{Tr}[\Phi^\dagger(x_i)\Phi(x_j)] = |\langle \phi(x_i)|\phi(x_j)\rangle|^2$ between the feature states. To estimate such an inner product using a quantum computer, we simply apply $U^\dagger(x_j)U(x_i)$ on input $|0^n\rangle$, and estimate the probability of the 0^n output. We call such a procedure quantum kernel estimation (QKE), highlighting the fact that this procedure is subject to small additive noise in each kernel entry due to finite sampling. The overall procedure for learning with quantum feature map is now clear. On input a set of m labeled training examples $S = \{(x_i, y_i)\}_{i=1}^m$, run QKE to obtain the $m \times m$ kernel matrix, then run the dual SVM Eq. (2) on a classical computer to obtain the solution α. To classify a new example x, run QKE to obtain $K(x, x_i)$ for each $i = 1, \ldots, m$, then return

$$y = \text{sign}\left(\sum_{i=1}^{m} \alpha_i y_i K(x, x_i)\right). \tag{4}$$

Throughout the entire SVM-QKE algorithm, QKE is the only subroutine that requires a

quantum computer, while all other optimization steps can be performed classically. A detailed description of the algorithm can be found in [4].

The main differences between quantum kernel methods and classical SVMs are two-fold. First, as quantum computers are exponentially more powerful than classical computers, we expect quantum feature maps to be more *expressive* than classical counterparts, which leads to better performance in pattern recognition and classification tasks. Second, quantum kernels are subject to small additive noise from the estimation procedure, while classical kernels can be computed exactly. Therefore, a demonstration of quantum advantage with the SVM-QKE algorithm needs to take advantage of the expressivity of quantum feature maps, while remaining robust against finite sampling noise.

While quantum kernel methods are very general and flexible, it is unclear that they have any provable advantage over classical algorithms. Indeed, previous analysis [86, 87] only establish evidence that parts of the algorithm cannot be efficiently simulated classically (for example, by definition we know that the quantum kernel estimation procedure is BQP-complete in general), but this does not imply that quantum kernel methods can have an advantage against all classical algorithms in a classification task. Here we provide a solution to this challenge by constructing a classification problem and rigorously prove that the SVM-QKE algorithm can solve it in polynomial time, while no classical polynomial time algorithm can perform better than random guessing for this task.

We shall now define a learning problem that we use to establish a quantum advantage. Recall that the task of supervised classification is to assign a label $y \in \{-1, 1\}$ to a datum $x \in \mathcal{X}$ from data space $\mathcal{X}$ according to some unknown decision rule f (usually referred to as a *concept*), by learning from labeled examples $S = \{(x_i, y_i)\}_{i=1,\ldots,m}$ that are generated from this concept, $y_i = f(x_i)$. Given the training set S, an efficient learner needs to compute a classifier f^* in time that is polynomial in the size of S, with the goal of achieving high *test accuracy*,

$$\mathrm{acc}_f(f^*) = \Pr_{x\sim\mathcal{X}}\left[f^*(x) = f(x)\right], \tag{5}$$

the probability of agreeing with f on unseen examples. Here we assume that the datum x is sampled uniformly random from $\mathcal{X}$, in both training and testing, and the size of S is polynomial in the data dimension.

An important ingredient of machine learning is prior knowledge, i.e., additional information given to the learning algorithm besides the training set. In standard computational learning theory [94, 99], this is modeled as a concept class – a (often exponentially large) set of labeling rules, and the target concept is promised to be chosen from the concept class. A concept class $\mathcal{C}$ is efficiently learnable if for every $f \in \mathcal{C}$, an efficient algorithm can achieve a high test accuracy (say, at least 0.99) by learning from examples labeled according to f with high success probability.

Classically hard learning problem. Our concept class that separates quantum and classical learnability is based on the discrete logarithm problem (DLP). For a large prime number p and a generator g of $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$, it is a widely-believed conjecture that no classical algorithm can compute $\log_g x$ on input $x \in \mathbb{Z}_p^*$, in time polynomial in $n = \lceil \log_2(p) \rceil$, the number of bits needed to represent p. Meanwhile, DLP can be solved by Shor's quantum algorithm [100] in polynomial time.

Based on DLP, we define our concept class $\mathcal{C} = \{f_s\}_{s\in\mathbb{Z}_p^*}$ over the data space $\mathcal{X} = \mathbb{Z}_p^*$ as

follows,

$$f_s(x) = \begin{cases} +1, & \text{if } \log_g x \in [s, s + \frac{p-3}{2}], \\ -1, & \text{else.} \end{cases} \tag{6}$$

Each concept $f_s : \mathbb{Z}_p^* \to \{-1, 1\}$ maps half the elements in $\mathbb{Z}_p^*$ to $+1$ and half of them to -1.

To see why the discrete logarithm is important in our definition, note that if we change $\log_g x$ to x in Eq. (6), then learning the concept class $\mathcal{C}$ is a trivial problem. Indeed, if we imagine the elements of $\mathbb{Z}_p^*$ as lying on a circle, then each concept f_s corresponds to a direction for cutting the circle in two halves (Fig. 1a). Therefore, the training set of labeled examples can be separated as two distinct clusters, where one cluster is labeled $+1$ and the other labeled -1. To classify a new example, a learning algorithm can simply decide based on which cluster is closer to the example. On the other hand, due to the classical intractability of DLP, the training samples for the concept class $\mathcal{C}$ look like randomly labeled from the viewpoint of a classical learner (Fig. 1b). In fact, we can prove that the best a classical learner can do is randomly guess the label for new examples, which achieves 50% test accuracy. These results are summarized below.

Theorem 1. *Assuming the classical hardness of DLP, no efficient classical algorithm can achieve* $\frac{1}{2} + \frac{1}{\mathsf{poly}(n)}$ *test accuracy for* $\mathcal{C}$.

Our proof of classical hardness of learning $\mathcal{C}$ is based on an average-case hardness result for discrete log by Blum and Micali [101]. They showed that computing the most significant bit of $\log_g x$ for $\frac{1}{2} + \frac{1}{\mathsf{poly}(n)}$ fraction of $x \in \mathbb{Z}_p^*$ is as hard as solving DLP. We then reduce our concept class learning problem to DLP using this result, by showing that if an efficient learner can achieve $\frac{1}{2} + \frac{1}{\mathsf{poly}(n)}$ test accuracy for $\mathcal{C}$, then it can be used to construct an efficient classical algorithm for DLP, which proves Theorem 1.

2.2 Covariant quantum kernels for data with group structure

A natural question is whether the quantum advantage established in the previous section can be observed in practically relevant settings. Our work [5] takes steps to address this quesdtion by identifying a class of learning problems that provide a natural fit for QKE and generalize the result of the previous section. What these learning problems have in common is that the data space $\mathcal{X}$ is a subset of some group G. The study of data with group structure has a long tradition in statistics [102]. Important learning problems such as ranking [102, 103] can be expressed as the learning of permutations [104, 105]. Other examples are learning problems in coset spaces such as partial rankings, Q-sort data, error correcting codes and homogeneous spaces [102]. We consider a general class of feature map circuits that we call covariant feature maps and that can be used for data space with a group structure. The corresponding quantum feature maps are intimately related to covariant measurements [106]. The covariant feature map is defined relative to a unitary representation $D : G \to U(2^n)$ and a fiducial state $|\psi\rangle$ of n qubits such that the datum $x \in G$ is mapped to $\Phi(x) = D_x|\psi\rangle\langle\psi|D_x^\dagger$. The covariant quantum kernel is then estimated as the fidelity $K(x, x') = |\langle\psi|D_x^\dagger D_{x'}|\psi\rangle|^2$. We assume the fiducial state $|\psi\rangle = V|0^n\rangle$ can be prepared by applying an efficient quantum circuit V. Likewise, it is important to also assume that the representation D_x of any element $x \in G$ can be implemented efficiently on a quantum computer. In this case, the QKE

routine reduces to estimating the transition amplitude $K(x,y) = |\langle 0^n|VD_x^\dagger D_y V|0^n\rangle|^2$, and the feature map circuit becomes $U(x) = D_x V$. The kernel as defined here is left-invariant under the group action. The learning problem from the previous section corresponds to a special case $G = \mathbb{Z}_p^*$ (integer multiplication modulo p) and reduces to the discrete logarithm problem. The corresponding kernel is group covariant while the fiducial state is the uniform superposition of group elements obtained from applications of the generator to a suitable interval.

While the aforementioned fiducial state produces a kernel that can lead to an efficient learning algorithm for the concept class defined in Eq. (6), a fiducial state ψ that is given by one of computational basis states would lead to an identity kernel, matrix which is well-known to have extremely poor performance. This illustrates that the choice of ψ is essential for the performance of the quantum kernel. If sufficient structural knowledge about the problem is present, a suitable fiducial state can be chosen a priori. However, we also want a method to optimize the fiducial state subject to the available data, if no prior knowledge is available. The objective of this optimization will depend on the learning problem. We consider a binary classification problem with SVMs. For other types of kernel functions, objectives have been proposed [107, 108] that are motivated by quantum information theoretic insights. Here, we will follow a method commonly used in the classical literature referred to as kernel alignment [109, 110]. We consider fiducial states of the form $|\psi_\lambda\rangle = V_\lambda|0^n\rangle$, where V_λ is a variational quantum circuit on n qubits that depends on a vector of variational parameters λ. This yields a parametrized quantum kernel $K_\lambda(x,y)$ with feature map circuit $D_x V_\lambda$. The parameters λ are optimized with kernel alignment according to

$$\lambda = \arg\min_\lambda \max_\alpha F(\alpha, \lambda) \tag{7}$$

where $F(\alpha,\lambda)$ is defined similarly to the dual objective function in Eq. (3), except that the kernel K is replaced by K_λ. The minimization over λ ranges over the chosen class of variational circuits V_λ. The maximization over α is subject to constraints of the feasible set $0 \leq \alpha_i \leq C$, where C is the box parameter, and $\sum_{i=1}^m y_i\alpha_i = 0$. We present a stochastic algorithm for this optimization problem, which is an iterative algorithm with kernel matrices evaluated on a quantum processor and continuous parameters updated with classical optimization routines, see [5] for details.

In the quantum experiment we want to benchmark both the accuracy of a learner with access to covariant quantum kernels, as well as the performance of the physical hardware. To this end we introduce a learning problem called labeling cosets with error (LCE) that serves as an abstraction of common learning problems on coset spaces. Learning problems on coset spaces are frequently considered in the literature [102, 104], for example when considering partial rankings [103] or for manifolds that arise as homogeneous spaces [102]. In LCE we are given a (continuous) group G, a subgroup $S \subset G$, and generate data from two left-cosets, $c_\pm S \subset G$ determined by representatives $c_+, c_- \in G$. Every datum taken from cosets is perturbed with a small error ϵ so that the data is not part of the coset any longer. After seeing sufficient data, the learner is asked to classify to which coset a previously unseen datum belongs. We implemented LCE with the groups G and S motivated by our physical hardware, namely, $G = SU(2)^{\otimes n}$ with $n = 27$. The subgroup S was chosen as the stabilizer group of the 27-qubit graph state for the qubit coupling graph describing the device connectivity. Details of this demonstration can be found in [5].

2.3 Learning of quantum many-body systems

Quantum many-body systems consist of many quantum particles (qudits) that locally interact with each other. The interactions between these particles are described by the Hamiltonian of the system. Even though the interactions in the Hamiltonian are local, the state of the whole system can be highly entangled. This is not only true at low temperatures when the system is in the lowest energy eigenstate of its Hamiltonian (the ground state), but remains true even at finite temperatures when the state is a mixture of different eigenstates of the Hamiltonian known as the Gibbs or thermal state. While the underlying fundamental interactions in these systems are long known to be given by Coulomb forces between electrons and nuclei, they are too complicated to be grasped in entirety. Physicists are primarily interested in "effective interactions" that, if accurately chalked out, can be used to describe a variety of properties of the system. How can such effective interactions be learned in a system as complicated as, for example, the high temperature superconductor? Algorithms for Hamiltonian learning can directly address this problem and provide a suitable approximation to the effective interactions.

Hamiltonian Learning Problem Suppose Λ is a regular D-dimensional lattice with n sites. We consider k-local geometrically local Hamiltonian H acting on n qudits located at sites of Λ. In general, we can parameterize such Hamiltonian as

$$H(\mu) = \sum_{\ell=1}^{m} \mu_\ell E_\ell$$

where μ_ℓ are real coefficients and E_ℓ are linearly independent Hermitian operators acting non-trivially on at most k qudits such that the support of E_ℓ has a constant diameter r. We assume that k, D, r are constants independent of n. For example, E_ℓ could be tensor products of k single-qubit Pauli operators acting on spatially contiguous blocks of sites. We let the vector $\mu = (\mu_1, \ldots, \mu_m)$ be the vector of interaction coefficients. In our setup, without loss of generality we assume the Hamiltonian is traceless, i.e., for the identity operator $E_\ell = I$, the coefficient $\mu_\ell = 0$. At a inverse-temperature β, the qudits are in the Gibbs state defined as

$$\rho_\beta(\mu) = \frac{e^{-\beta H(\mu)}}{\mathrm{tr}[e^{-\beta H(\mu)}]}.$$

In the learning problem, we are given multiple copies of $\rho_\beta(\mu)$ and can perform arbitrary local measurements on them. In particular, we can obtain all the k-local marginals of $\rho_\beta(\mu)$ denoted by

$$e_\ell = \mathrm{tr}[\rho_\beta(\mu) E_\ell] \quad \text{for } \ell \in [m].$$

The goal is to learn the coefficients μ_ℓ of the Hamiltonian H using the result of these measurements. We call this the Hamiltonian Learning Problem. This problem has been the focus of many recent theoretical and experimental works [7, 8, 9, 10, 11, 12]. The classical analogue of this problem plays the central role in machine learning and modern statistical inference, known as learning graphical models or Boltzmann machines (aka Ising models). Classically, understanding the learnability of Boltzmann machines was initiated by the works of Hinton and others in the 80s [111, 112]. In the past few years, there has been renewed interest in

this subject and has seen significant progress resulting in efficient *provable* learning algorithms for graphical models with optimal sample and time complexity especially for sparse and bounded-degree graphs [13, 14, 15, 16, 17]. Thus far, a rigorous analysis of the quantum Hamiltonian learning problem with guaranteed sample complexity has been lacking.

Main result. Let $H(\mu)$ be a geometrically local Hamiltonian defined above. We are given N copies of the Gibbs state $\rho_\beta(\mu)$ at a fixed inverse-temperature β. Our goal is to obtain an estimate $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_m)$ of the coefficients μ_ℓ such that with probability at least $1 - \delta$,

$$\|\mu - \hat{\mu}\|_2 \leq \epsilon,$$

where $\|\mu - \hat{\mu}\|_2$ is the ℓ_2-norm of the difference of μ and $\hat{\mu}$.

Our main result is a sample-efficient algorithm for the Hamiltonian Learning Problem that requires

$$N = \mathcal{O}\left(\frac{e^{\beta^c}}{\beta^{\tilde{c}}\epsilon^2} \cdot m^3 \cdot \log\left(\frac{m}{\delta}\right)\right) \tag{8}$$

copies of the Gibbs state $\rho_\beta(\mu)$, where $c, \tilde{c} \geq 1$ are constants that depend on the geometry of the Hamiltonian. For geometrically local Hamiltonians the number of interaction terms m scales as $O(n)$. Hence, the result implies a sample complexity polynomial in the number of qudits.

The number of samples in Eq. (8) increases as $\beta \to \infty$ or $\beta \to 0$. As the temperature increases ($\beta \to 0$), the Gibbs state approaches the maximally mixed state independent of the choice of parameters μ. At low temperatures ($\beta \to \infty$), the Gibbs state is in the vicinity of the ground space, which for instance, could be a product state $|0\rangle^{\otimes n}$ for the various choices of μ. In either cases, more sample are required to distinguish the parameters μ.

Ideally, one would like to have an algorithm for the Hamiltonian Learning Problem that requires small number of samples, but also has a polynomial running time. Satisfying both these constraints for all inverse temperatures β even in the classical learning problems is quite challenging, as it requires approximating the partition function which is NP-hard [113]. There are recent results on efficiently computing the partition function of quantum many-body systems under various assumptions [114, 115, 116] that may enable efficient Hamiltonian learning in some special cases.

Proof overview. Let $S(\sigma, \eta) \equiv \mathrm{Tr}[\sigma \log \sigma] - \mathrm{Tr}[\sigma \log \eta]$ be the relative entropy of density matrices σ and η. Let $S(\sigma) = -\mathrm{Tr}[\sigma \log \sigma]$ be the von Neumann entropy. Recall that $S(\sigma, \eta) \geq 0$ for all σ, η with the equality iff $\sigma = \eta$. Substituting $\sigma = \rho_\beta(\mu)$ and $\eta = \rho_\beta(\lambda)$ one infers that

$$S(\rho_\beta(\mu)) \leq \log Z_\beta(\lambda) + \beta \mathrm{Tr}[\rho_\beta(\mu) H(\lambda)] \tag{9}$$

for any $\lambda \in \mathbb{R}^m$ with the equality iff $\lambda = \mu$. Here

$$Z_\beta(\lambda) = \mathrm{Tr}\left[e^{-\beta H(\lambda)}\right]$$

is the quantum partition function. Expressing the expected value of $H(\lambda)$ in Eq. (9) in terms of the marginals $e_\ell = \mathrm{Tr}[\rho_\beta(\mu) E_\ell]$ one concludes that

$$\mu = \arg\min_{\lambda \in \mathbb{R}^m} \log Z_\beta(\lambda) + \beta \sum_{\ell=1}^{m} \lambda_\ell e_\ell. \tag{10}$$

If one knows the exact marginals e_ℓ and leaves aside the computational complexity, one can find the desired coefficients μ by solving the optimization problem Eq. (10). Given N samples of the Gibbs state $\rho_\beta(\mu)$, the best one can hope for is to approximate the marginals e_ℓ within the statistical error of order $N^{-1/2}$. We denote these approximate marginals $\hat{e}_\ell$. Instead of solving the problem Eq. (10), our strategy is to solve its empirical version

$$\hat{\mu} = \arg\min_{\lambda\in\mathbb{R}^m} \log Z_\beta(\lambda) + \beta \sum_{\ell=1}^{m} \lambda_\ell \hat{e}_\ell, \tag{11}$$

The main technical challenge is proving that the optimization problem Eq. (10) is robust in against the statistical error in the marginals e_ℓ in the sense that $|\hat{\mu} - \mu|$ is small whenever $|\hat{e}_\ell - e_\ell|$ is small for all ℓ. To this end we show that the logarithm of the quantum partition function considered as a function of λ exhibits the strong convexity property, namely,

$$\nabla^2 \log Z_\beta(\lambda) \geq \alpha I, \quad \alpha = e^{-\mathcal{O}(\beta^c)} \frac{\beta^{c'}}{m}. \tag{12}$$

Here ∇^2 is the Hessian of the function $\log Z_\beta(\lambda)$ with respect to $\lambda \in \mathbb{R}^m$ and I is the $m \times m$ identity matrix. The operator inequality in Eq. (12) means that the matrix $\nabla^2 \log Z_\beta(\lambda) - \alpha I$ is positive semi-definite. Showing that the error in estimating μ is small whenever the statistical error of the marginals is small relies crucially on the strong convexity Eq. (12). Specifically, we show that the optimal solution of Eq. (11) satisfies

$$\|\mu - \hat{\mu}\|_2 \leq \frac{2\beta m^{1/2}}{\alpha} \max_\ell |e_\ell - \hat{e}_\ell|. \tag{13}$$

Proving the strong convexity property Eq. (12) is the main technical contribution of our work. The proof begins by expressing the minimum eigenvalue of the Hessian in terms of the energy variance. For any a vector $v \in \mathbb{R}^m$, we show that

$$\sum_{i,j=1}^{m} v_i v_j \frac{\partial^2}{\partial\mu_i \partial\mu_j} \log Z_\beta(\mu) \geq \beta^2 \mathrm{Var}[\tilde{W}], \tag{14}$$

where $\tilde{W} = \sum_{\ell=1}^m v_\ell \tilde{E}_\ell$, the variance of $\tilde{W}$ is defined with respect to the Gibbs state $\rho_\beta(\mu)$, and the operators $\tilde{E}_\ell$ are defined as

$$\tilde{E}_\ell = \int_{-\infty}^{\infty} f_\beta(t) e^{-iH(\mu)t} E_\ell e^{iH(\mu)t} dt.$$

Here $f_\beta(t)$ is a certain filter function that decays exponentially for large t. The Lieb-Robinson bound [117, 118, 119] asserts that the unitary dynamics generated by a geometrically local Hamiltonian can propagate information only within a lightcone of radius proportional to the evolution time. This implies that $\tilde{E}_\ell$ can be well approximated by a sum of local operators, that is, $\tilde{W}$ is a quasi-local Hamiltonian. We obtain a lower bound on the variance of $\tilde{W}$ using the quantum belief propagation method developed in [120, 121, 122] and a recent result by Arad et al. [123] connecting global and local energy distributions in geometrically local quantum Hamiltonians. It shows that matrix elements of a local observable between eigenstates of a local Hamiltonian which are well-separated in energy are exponentially small.

2.4 Learning quantum phase states

By definition, an n-qubit, degree-d phase state has the form

$$|\psi_f\rangle = 2^{-n/2} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle, \tag{15}$$

where $f : \{0,1\}^n \to \{0,1\}$ is a degree-d polynomial, that is,

$$f(x) = \sum_{\substack{J\subseteq[n]\\|J|\le d}} \alpha_J \prod_{j\in J} x_j \pmod 2, \tag{16}$$

for some coefficients $\alpha_J \in \{0,1\}$. Phase states associated with homogeneous degree-2 polynomials $f(x)$ coincide with graph states that play a prominent role in quantum information theory [124]. Such states can be alternatively represented as

$$|\psi_f\rangle = \prod_{(i,j)\in E} \mathsf{CZ}_{i,j}|+\rangle^{\otimes n},$$

where n qubits live at vertices of a graph, E is the set of graph edges, $\mathsf{CZ}_{i,j}$ is the controlled-Z gate applied to qubits i,j, and $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. It is known that the output state of any Clifford circuit is locally equivalent to a graph state for a suitable graph [125]. Our results imply that graph states can be efficiently learned using only single-qubit gates and measurements. The best previously known protocol for learning graph states [83] requires entangled measurements across two copies of $|\psi_f\rangle$. Other examples of circuits producing these states include measurement-based quantum computing [126] and IQP circuits (Instantaneous Quantum Polynomial-time), which correspond to degree-3 phase states [127]. IQP circuits are prevalent in quantum-advantage experiments [128, 129] and are believed to be hard to simulate classically.

Apart from being natural states to learn, phase states have appeared in several recent works: [130, 131] showed phase states are efficiently preparable and statistically indistinguishable from a Haar random state (for a polynomial-time quantum algorithm), subsequently there have been followup works using phase states for cryptosystems [132]; Irani et al. [133] showed that in order to construct the witness to a QMA complete problem, say the ground state $|\phi\rangle$ to a local-Hamiltonian problem, it suffices to consider a phase state which has a good overlap to $|\phi\rangle$; level-3 phase states are universal for measurement based quantum computing [126, 134].

We also consider generalized degree-d phase states

$$|\psi_f\rangle = 2^{-n/2} \sum_{x\in\{0,1\}^n} e^{\frac{2\pi i}{q}f(x)}\,|x\rangle\,, \tag{17}$$

where $q \ge 2$ is an even integer and $f : \{0,1\}^n \to \mathbb{Z}_q$ is a degree-d polynomial, that is,

$$f(x) = \sum_{\substack{J\subseteq[n]\\|J|\le d}} \alpha_J \prod_{j\in J} x_j \pmod q \tag{18}$$

for coefficients $\alpha_J \in \mathbb{Z}_q = \{0, 1, \ldots, q-1\}$. It is known that the output state of a random n-qubit Clifford circuit is a generalized $q = 4$, degree-2 phase state with a constant probability [135, Appendix D]. It is also known that generalized degree-d phase states with $q = 2^d$ can be prepared from diagonal unitary operators [136] in the d-th level of the Clifford hierarchy [137]. These states have also found applications in cryptography [130, 131], and complexity theory [133].

In this work, we consider learning phase states through two types of tomography protocols based on *separable* and *entangled* measurements. The former can be realized as a sequence of M independent measurements, each performed on a separate copy of $|\psi_f\rangle$. The latter performs a joint measurement on the state $|\psi_f\rangle^{\otimes M}$. Interestingly, our learning protocols based on separable measurements require only single-qubit gates and single-qubit measurements making them well suited for near-term demonstrations. Our goal is to then derive upper and lower bounds on the sample complexity M of learning f, as a function of n and d.

We first introduce some notation before giving an overview of our contributions. For every n and $d \leq n/2$, let $\mathcal{P}(n, d)$ be the set of all degree-d polynomials of the form Eq. (16). Let $\mathcal{P}_q(n, d)$ be the set of all degree-d $\mathbb{Z}_q$-valued polynomials of the form Eq. (17). By definition, $\mathcal{P}_2(n, d) \equiv \mathcal{P}(n, d)$. To avoid confusion, we shall refer to states defined in Eq. (15) as binary phase states and in Eq. (17) as generalized phase states. Our learning protocol takes as input integers n, q, d and M copies of a degree-d phase state $|\psi_f\rangle$ with unknown $f \in \mathcal{P}_q(n, d)$. The protocol outputs a classical description of a polynomial $g \in \mathcal{P}_q(n, d)$ such that $f = g$ with high probability.

Our work reports optimal algorithms for learning phase states, both in the case of separable and entangled measurements. Prior to our work, it was only known how to efficiently learn degree-2 phase states. Here, we show that if allowed separable measurements, the sample complexity of learning binary phase states and generalized phase states is $O(n^d)$. If allowed entangled measurements, we obtain a sample complexity of $O(n^{d-1})$. We further consider settings where the function f we are trying to learn is known to be sparse, has a small Fourier-degree and when given noisy copies of the quantum phase state. In Table 1, we summarize all our main results (except the first two rows, which include the main prior work in this direction[1]).

[1]Montanaro [138] considered learning multilinear polynomials f, assuming we have *query access* to f, which is a stronger learning model than the sample access model that we assume for our learning algorithm.

	Sample complexity	**Measurements**
Binary phase state $\mathbb{F}_2$-degree-1 [139]	$\Theta(1)$	Separable
Binary phase state $\mathbb{F}_2$-degree-2 [83, 140]	$O(n)$	Entangled
Binary phase state $\mathbb{F}_2$-degree-d	$\Theta(n^d)$	Separable
Binary phase state $\mathbb{F}_2$-degree-d	$\Theta(n^{d-1})$	Entangled
Generalized phase states degree-d	$\Theta(n^d)$	Separable
Generalized phase states degree-d	$\Theta(n^{d-1})$	Entangled
Sparse Binary phase state $\mathbb{F}_2$-degree-d, $\mathbb{F}_2$-sparsity s	$O(2^d sn)$	Separable
Binary phase state $\mathbb{F}_2$-degree-2 with global depolarizing noise ε	$n^{1+O(\varepsilon)}$	Entangled
Binary phase state $\mathbb{F}_2$-degree-2 with local depolarizing noise ε	$\Theta((1-\varepsilon)^n)$	Entangled
Binary phase state Fourier-degree-d	$O(2^{2d})$	Entangled

Table 1: Upper and lower bounds on sample complexity for exact learning of n-qubit degree-d phase states.

Next we discuss a couple of motivations for considering the task of learning phase states and corresponding applications.

Quantum complexity. Recently, there has been a few results in quantum cryptography [130, 132, 131] and complexity theory [133] which used the notion of phase states.

Ji et al. [130] introduced the notion of *pseudorandom quantum states* as states of the form

$|\phi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} e^{\frac{2\pi i}{N}F(x)} |x\rangle$ where F is a pseudorandom function.[2] Ji et al. showed that states of the form $|\phi\rangle$ are efficiently preparable and statistically indistinguishable from a Haar random state, which given as input to a polynomial-time quantum algorithm. A subsequent work of Brakerski [131] showed that it suffices to consider $|\phi'\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{F(x)} |x\rangle$ (where F again is a pseudorandom function) and such states are also efficiently preparable and statistically indistinguishable from Haar random states. Subsequently, these states have found applications in proposing many cryptosystems [132]. Although none of these works discuss the degree of the phase function F, our result shows implicitly that when F is low-degree, then $|\phi\rangle$ is exactly learnable and hence distinguishable from Haar random states, implying that they cannot be quantum pseudorandom states.

Learning quantum circuits. Given access to a quantum circuit U, the goal of this learning task is to learn a circuit representation of U. The sample complexity for learning a general n-qubit quantum circuit is known to be $2^{\Theta(n)}$ [141, 142], which is usually impractical.

If we restrict ourselves to particular classes of quantum circuits, there are some known results for efficient learnability. Low [143] showed that an n-qubit Clifford circuit can be learned using $O(n)$ samples. However, this result was only an existential proof and requires access to the conjugate of the circuit. Constructive algorithms were given in Low [143], and Lai and Cheng [144], both of which showed that Clifford circuits can be learned using $O(n^2)$ samples. Both these algorithms require entangled measurements with the former algorithm using pretty-good measurement [145], and the latter using Bell sampling. In this work, we show that Clifford circuits producing degree-2 binary phase states, can be learned in $O(n^2)$ samples, matching their result but only using separable measurements (in fact, using only single-qubit measurements). Moreover, Low [143] also gave an existential proof of algorithms for learning circuits in the d-th level of the Clifford hierarchy, using n^{d-1} samples. In this work, we give constructive algorithms for learning the diagonal elements of the Clifford hierarchy in n^d samples (resp. n^{d-1} samples) using separable measurements (resp. with entangled measurements). A direct result of this is that IQP circuits, which are believed to be hard to simulate [128, 146], are efficiently learnable. Our learning result thus gives an efficient method for verifying IQP circuits that may be part of quantum-advantage experiments [129, 147].

Learning hypergraph states. We finally observe that degree-3 (and higher-degree) phase states have appeared in works [126, 134] on measurement-based quantum computing (MBQC), wherein they refer to these states as *hypergraph states*. These works show that single-qubit measurements in the Pauli X or Z basis performed on a suitable degree-3 hypergraph state are sufficient for universal MBQC. Our learning algorithm gives a procedure for learning these states in polynomial-time and could potentially be used as a subroutine for verifying MBQC.

Our work leaves open a few interesting questions.

Improving runtime. While our algorithms for learning phase states are optimal in terms of the sample complexity, their runtime scales polynomially with the number of qubits only

[2] We do not discuss the details of pseudorandom functions here, we refer the interested reader to [130].

in the case of binary phase states and separable measurements. It remains to be seen whether a polynomial runtime can be achieved in the remaining cases, i.e., learning degree-d binary phase states using $O(n^{d-1})$ samples with entangled measurements and generalized phase states with either separable or entangled measurements.

Quantum advantage. Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called a bent (maximally non-linear) function if there exists a Boolean function $g : \{0,1\}^n \to \{0,1\}$ such that

$$(-1)^{g(y)} = 2^{-n/2} \sum_{x \in \{0,1\}^n} (-1)^{x \cdot y + f(x)}$$

for all $y \in \{0,1\}^n$. The function g is called the dual of f. If f is a bent function as above then n-qubit states $|\psi_f\rangle$ and $H^{\otimes}|\psi_f\rangle = |\psi_g\rangle$ are both binary phase states. Here H is the Hadamard gate. Consider the problem of learning f given a classical description of the dual function g. This problem can be solved efficiently using our algorithm whenever both f and g are low-degree polynomials. Indeed, in this case one can efficiently prepare the state $|\psi_f\rangle$ by first preparing the phase state $|\psi_g$ (using the known classical description of g by a low-degree polynomial) and then applying bit-wise Hadamard operation $H^{\otimes n}$ which maps $|\psi_f\rangle$ to $|\psi_g\rangle$. Now one can learn g using the algorithms reported in our work. We leave as an open question whether the task of learning a bent function f starting with a classical description of the dual bent function g is classically hard. We note that a closely related hidden shift problem for bent function is known to be classically hard in the query complexity sense [148].

Property testing. What is the sample complexity of property testing phase states? Given M copies of an n-qubit state $|\phi\rangle$ with the promise that either $|\phi\rangle$ is a degree-d phase state or ε-far from the set of degree-d phase states, what is an upper and lower bound on M? For $d = 1$, we can learn the entire state using $M = 1$ copy and for $d = 2$, Gross et al. [149] showed that $M = 6$ copies suffice for the property testing. For larger d, understanding the complexity of testing phase states is an intriguing open question left open by our work. In particular, does the sample complexity of testing n-qubit degree-d phase states scale as n^{d-2} (for $d \geq 2$) or does it scale as $\mathsf{poly}(c^d, n)$ for some $c > 1$?

Learning more expressive quantum states. We leave as an open question whether our learning algorithms can be extended to binary phase states with a small *algebraic degree*. Such states have amplitudes proportional to $(-1)^{\mathsf{tr}F(x)}$, where $F(x) = \sum_{i=0}^{d} a_i x^i$ is a degree-d polynomial with coefficients $a_i \in \mathbb{F}_{2^n}$ and $\mathsf{tr} : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is the trace function defined as $\mathsf{tr}(x) = \sum_{j=0}^{n-1} x^{2^j}$. Here all arithmetic operations use the field $\mathbb{F}_{2^n}$. What is the sample complexity of learning n-qubit states produced by circuits containing *non-diagonal* unitaries in the k-th level in the Clifford hierarchy, on the $|+\rangle^n$ input? Similarly, what is the complexity of learning a state which has stabilizer rank k?[3] Similarly can we PAC learn these classes of quantum states in polynomial time?[4]

[3] We know how to learn stabilizer states and stabilizer-rank 2 states in polynomial time, what is the complexity as a function of rank-k?

[4] For stabilizer circuits, we have both positive and negative results in this direction [150, 151] but for more generalized circuits, it remains an open question.

3 Quantum optimization algorithms

Variational quantum optimization (VQO) has recently received significant attention as a candidate application of near-term quantum processors. The basic proposal is appealingly simple: the output state of a parameterized quantum circuit is used as a variational wavefunction to minimize the expected energy of a given Hamiltonian [152]. Depending on the envisioned application, the Hamiltonian may govern electronic interactions in a molecule or material of interest [153], or encode a classical cost function whose minimum is to be approximated [154]. Rotation angles that define individual gates in the state preparation circuit serve as variational parameters. These parameters are adjusted via a classical feedback loop that aims to minimize the expected energy.

The central question common to all VQO proposals is whether the chosen variational class of states is expressive enough to provide a good ground state approximation. Let us point out two factors that can limit the expressive power of VQO.

First, the state preparation quantum circuit must have a small depth to enable reliable implementation on near-term noisy devices lacking error correction. This means that highly entangled ground states may be out of scope for near-term VQO using gate-based devices. For example, preparing ground states of Kitaev's toric code [50] starting from a product state requires a circuit depth growing at least polynomially in the system size using nearest-neighbor gates [155], and logarithmically using non-local gates [156]. Therefore, near-term VQO is unlikely to attain the ground state energy for such Hamiltonians.

Secondly, the number of variational parameters in the state preparation circuit must be small to enable efficient energy minimization. While this is not a serious concern for proof-of-principle experiments with a handful of qubits, it is anticipated that large-scale VQO with an extensive number of variational parameters may give rise to intractable optimization problems, for example due to the barren plateau (vanishing gradient) effect [157, 158]. Algorithmic limitations of this kind may prevent VQO from approximating even very simple unentangled ground states associated with classical cost functions.

Section 3.2, based on Ref. [159], elaborates on the limitations of VQO by establishing formal no-go results for the Quantum Approximate Optimization Algorithm (QAOA) [154]. It is shown that the symmetry and the locality of QAOA variational states severely limit its performance. A surprising consequence of our results is that the classical Goemans-Williamson algorithm [160] outperforms QAOA for certain instances of MaxCut, at any constant level.

Sections 3.3,3.4 based on Refs. [159, 26, 161], show how to overcome the no-go result established above. To this end we augment QAOA by a variable elimination step which results in a new VQO-type algorithm which we call a Recursive QAOA (RQAOA). We benchmark RQAOA for medium-size optimization problems with about 200 variables including MaxCut and graph coloring problems observing that RQAOA is strictly more powerful than the standard QAOA. Surprisingly, for certain problem instances we found that level-1 RQAOA outperforms the Goemans-Williamson algorithm for MaxCut and its generalization to graph coloring problems due to Frieze and Jerrum [162]. To enable efficient numerical simulations, we propose polynomial-time classical algorithms for computing mean values of observables on QAOA states.

VQO algorithms typically use classical processing only for the choice of variational pa-

rameters, for example, following gradient descent. Section 3.5, based on Ref. [27], shows that for certain combinatorial optimization problems, such algorithms can be hybridized further, thus harnessing the power of efficient non-local classical processing to a greater extent. Specifically, we consider combining a quantum variational ansatz with a greedy classical post-processing procedure for the MaxCut-problem on 3-regular graphs. We show that the average cut-size produced by this method can be quantified in terms of the expected energy of a modified problem Hamiltonian on the original variational state. This motivates the consideration of an improved algorithm which variationally optimizes the energy of the modified Hamiltonian over the original class of variational states. We call this a twisted hybrid optimization algorithm since the additional classical processing step is folded into the definition of the problem Hamiltonian. We exemplify the viability of this method using the quantum approximate optimization algorithm (QAOA), giving analytic lower bounds on the expected approximation ratios achieved by twisted QAOA.

Markov Chain Monte Carlo (MCMC) methods provide a general approach to sampling from high-dimensional probability distributions. They have many applications in machine learning for Bayesian inference [163], in theoretical computer science for counting problems [164], volume estimation of convex bodies [165], approximation of the permanent [166], or in statistical physics for estimating thermodynamic properties of systems [167, 168]. It is known that quantum algorithms can provide quadratic speedups over classical algorithms for a variety of tasks related to MCMC, such as amplitude amplification and estimation [169] and spectral gap amplification of Markov Chains [170]. Even though a general quantum speedup for MCMC methods is not known, quantum algorithms for the computation of partition functions using simulated annealing have been previously proposed [171, 172]. Section 3.6 based on Ref. [173], improves upon the prior work and derives the currently best known quantum algorithm for this problem.

Section 3.8, based on Ref. [30], describes quantum subroutines for the simplex method – one of the most impactful algorithms of the past century which is still widely used in a variety of applications. The simplex algorithm solves a linear program $\min_x c^T x$ subject to $Ax = b$, $x \geq 0$, by keeping track of a basis – a set of linearly independent columns of A and repeatedly moves to a different basis that defines a solution with better objective function value. The basis change (called a pivot) is performed by determining a new column that should enter the basis, and removing one column from the current basis. Choosing the new column is called pricing, and it is asymptotically the most expensive step, as it involves a matrix inversion. For problems with a well-conditioned sparse constraint matrix A we give a quantum pricing algorithm that achieves a provable asymptotic speedup over best known classical algorithms. Importantly, this speedup does not depend on the data being available in some "quantum form" such as QRAM; the input of our quantum subroutines is the natural classical description of the problem, and the output is the index of the variables that should leave or enter the basis.

3.1 Quantum Approximate Optimization Algorithm

In the next few sections we consider Quantum Approximate Optimization Algorithm (QAOA) proposed in [22]. This algorithm aims to approximate the maximum of a classical cost function $C(x)$ that depends on n binary variables, $x = (x_1, \ldots, x_n)$. The cost function is encoded

into an n-qubit diagonal Hamiltonian

$$C = \sum_{x\in\{0,1\}^n} C(x)|x\rangle\langle x|.$$

QAOA variationally maximizes the expected energy of C over n-qubit quantum states of the form [154]

$$|\psi(\beta,\gamma)\rangle = \prod_{a=1}^{p} e^{i\beta_a B} e^{i\gamma_a C}|+^n\rangle \tag{19}$$

where β_a, γ_a are variational parameters, $|+^n\rangle$ is the tensor product of n single-qubit states $|+\rangle = (|0\rangle+|1\rangle)/\sqrt{2}$ and $B = \sum_{j=1}^{n} X_j$ is the transverse magnetic field operator. The integer p, called the QAOA *level*, controls the expressive power of the variational ansatz. Finally, QAOA outputs a bit string x obtained by preparing the optimal variational state $|\psi(\beta,\gamma)\rangle$ and measuring each qubit in the standard basis. The expected value of $C(x)$ coincides with the variational energy $\langle\psi(\beta,\gamma)|C|\psi(\beta,\gamma)\rangle$. The performance of QAOA is commonly quantified by an *approximation ratio* defined as the ratio between the maximum variational energy and the maximum value of the cost function $C_{\mathsf{max}} = \max_x C(x)$.

3.2 Limitations of QAOA

A paradigmatic test case for QAOA is the Maximum Cut (MaxCut) problem [154]. Suppose $G = (V, E)$ is a graph with a set of n vertices V labeled by integers $j = 1, \ldots, n$ and a set of edges E. Given an n-bit string x, let $\mathsf{cut}(x)$ be the set of edges $(j, k) \in E$ such that $x_j \neq x_k$. The cost function to be maximized is the cut size, $C(x) = |\mathsf{cut}(x)|$. The corresponding n-qubit Hamiltonian is

$$C = \frac{1}{2}\sum_{(j,k)\in E} (I - Z_j Z_k). \tag{20}$$

Here Z_j is the Pauli Z operator applied to a qubit j and I is the identity. The state preparation circuit defined in Eq. (19) has depth $\approx pD$, where D is the maximum vertex degree of the graph G and p is the QAOA level [174]. To keep the circuit depth and the number of variational parameters small, below we focus on the regime when p and D are constants or slowly growing functions of n.

Our first result is an upper bound on the maximum variational energy attained by level-p states. Namely, we show that for any constant $D \geq 3$ and all large enough n there exist a degree-D graph G with n vertices such that

$$\frac{\langle\psi(\beta,\gamma)|C|\psi(\beta,\gamma)\rangle}{C_{\mathsf{max}}} \leq \frac{5}{6} + \frac{\sqrt{D-1}}{3D} \tag{21}$$

for any $\beta, \gamma \in \mathbb{R}^p$ as long as $p < (1/3\,\log_2 n - 4)(D+1)^{-1}$. This result severely limits the performance of QAOA with any constant level p independent of n. Indeed, the right-hand side of Eq. (21) is approximately $5/6 \approx 0.833$ for large vertex degree D. For comparison, the best-known classical algorithm for MaxCut due to Goemans and Williamson [160] achieves the approximation ratio at least 0.878 on an arbitrary graph. Thus QAOA with a constant

level p cannot outperform classical algorithms. We note that upper bounds on the QAOA approximation ratio were previously known only for $p = 1$ [154]. Followup work [175, 176] established analogous limitations of QAOA applied to MaxCut and Maximum Independent Set for random regular graphs [175, 176]. These results exploit the locality of QAOA and its uniformity: reduced states on isomorphic local subgraphs are identical. In contrast, our focus is on $\mathbb{Z}_2$-symmetry and locality, and our statements also apply to non-uniform local algorithms. We refer to [177] for numerical studies of QAOA applied to MaxCut investigating how the approximation ratio scales with the level p.

Similar concerns about limitations of QAOA have previously been voiced by Hastings [178] who showed analytically that certain local classical algorithms match the performance of level-1 QAOA for Ising-like cost functions with multi-spin interactions. Hastings also gave numerical evidence for the same phenomenon for MaxCut with $p = 1$, and argued that this should extend to $p > 1$ [178].

QAOA states possess a certain symmetry that plays a crucial role in our analysis. Namely the state $|\psi(\beta,\gamma)\rangle$ is invariant under a global spin flip,

$$X^{\otimes n}|\psi(\beta,\gamma)\rangle = |\psi(\beta,\gamma)\rangle. \tag{22}$$

Indeed, the Hamiltonians B, C commute with the symmetry operator $X^{\otimes n}$, while the initial state $|+^n\rangle$ is a $+1$ eigenvector of $X^{\otimes n}$. More generally, let us say that an n-qubit state $|\psi\rangle$ is $\mathbb{Z}_2$-symmetric if it is a $+1$ eigenvector of $X^{\otimes n}$. Our proof of Eq. (21) combines two observations: (i) the symmetry forces good variational states to be highly entangled, and (ii) low-depth circuits are not capable of preparing highly entangled states.

To elaborate on the role of the $\mathbb{Z}_2$-symmetry, suppose $x \in \{0,1\}^n$ is an optimal cut such that $C_{\mathsf{max}} = C(x)$. Let $\overline{x}$ be the bit-wise negation of x. Note that $C(\overline{x}) = C(x)$. Although the state $|x\rangle$ has no entanglement whatsoever, its $\mathbb{Z}_2$-symmetric version $(|x\rangle + |\overline{x}\rangle)/\sqrt{2}$ is a highly entangled state, locally equivalent to the n-qubit GHZ state $(|0^n\rangle + |1^n\rangle)/\sqrt{2}$, which cannot be prepared by a low-depth circuit [155].

The fact that symmetry may prevent one from preparing ground states of certain Hamiltonians by low-depth circuits, even if the Hamiltonian has a ground state with only short-range entanglement, is well-known in the theory of topological quantum order under the name *symmetry protection* [179, 180, 181]. In this language, the bound Eq. (21) asserts that the Hamiltonian C exhibits a particularly strong version of the symmetry protection that extends to all states with energy density above a certain constant threshold.

We shall now argue that for a suitable family of graphs G all good variational states are qualitatively similar to the GHZ state. Specifically, the results of [182, 183, 184] show that for any constant $D \geq 3$ there exists an infinite family of bipartite degree-D graphs G such that

$$C(x) \equiv |\mathsf{cut}(x)| \geq h \min\{|x|, n - |x|\} \tag{23}$$

for any $x \in \{0,1\}^n$, where h is a constant satisfying

$$h \geq \frac{D}{2} - \sqrt{D-1} \tag{24}$$

and $|x|$ is the Hamming weight of x. Such graphs, known as Ramanujan expander graphs, have the peculiar feature that the spectral gap of their adjacency matrix takes the largest

possible value among all degree-D graphs. For example, random D-regular bipartite graphs are known to approach the bound Eq. (24) with high probability [185].

Let G be a bipartite graph as above and $x_{\mathsf{opt}} \in \{0,1\}^n$ be an optimal solution of the MaxCut problem. By definition, the maximum cut of a bipartite graph includes all edges, so that $C_{\mathsf{max}} = C(x_{\mathsf{opt}}) = |E|$. The assumption that G is bipartite also implies that the optimal solution x_{opt} is unique up to the bit-wise negation and

$$C(x) + C(x_{\mathsf{opt}} \oplus x) = |E| \tag{25}$$

for all $x \in \{0,1\}^n$. Here $\oplus$ denotes the bit-wise XOR. Set $\epsilon = h/6$ and consider a level-p QAOA state such that

$$\langle \psi(\beta,\gamma)|C|\psi(\beta,\gamma)\rangle \geq |E| - \epsilon n. \tag{26}$$

Let x be a random n-bit string sampled from the distribution $P(x) = |\langle x|\psi(\beta,\gamma)\rangle|^2$. Markov's inequality and Eq. (26) show that $C(x) \geq |E| - 2\epsilon n$ with probability at least $1/2$. From Eq. (25) one infers that $C(x_{\mathsf{opt}} \oplus x) \leq 2\epsilon n$ with probability at least $1/2$. Let $\mathrm{dist}(x,y) = |x \oplus y|$ be the Hamming distance between bit strings x and y. Combining Eq. (23) and the bound $C(x_{\mathsf{opt}} \oplus x) \leq 2\epsilon n$ one gets

$$\min\{\mathrm{dist}(x_{\mathsf{opt}}, x), \mathrm{dist}(\overline{x_{\mathsf{opt}}}, x)\} \leq \frac{2\epsilon n}{h} = \frac{n}{3} \tag{27}$$

with probability at least $1/2$. This shows that the state $|\psi(\beta,\gamma)\rangle$ has a non-negligible weight on bit strings close to x_{opt} and on those close to $\overline{x_{\mathsf{opt}}}$. Here closeness means being within a Hamming distance of at most $n/3$.

Finally, we employ a fascinating result by Eldar and Harrow stated as Corollary 43 in [186]. It asserts that the output distribution of a low-depth circuit cannot assign a non-negligible weight to subsets of bit strings that are far apart in Hamming distance. Namely, suppose $|\psi\rangle$ is an n-qubit state that can be prepared starting from a product state by a depth-d quantum circuit composed of one- and two-qubit gates. The state $|\psi\rangle$ does not have to be symmetric in any sense. Define a distribution $P(x) = |\langle x|\psi\rangle|^2$. Given a subset $S \subseteq \{0,1\}^n$, let $P(S) = \sum_{x\in S} P(x)$. Ref. [186] showed that for any subsets $S, S' \subseteq \{0,1\}^n$ one has

$$\mathrm{dist}(S,S') \leq \frac{4n^{1/2}2^{3d/2}}{\min\{P(S), P(S')\}}. \tag{28}$$

Here $\mathrm{dist}(S,S') = \min_{x\in S}\min_{y\in S'}\mathrm{dist}(x,y)$ is the minimum number of bit flips required to get from S to S'. Choose S and S' as the sets of n-bit strings x such that $\mathrm{dist}(x_{\mathsf{opt}}, x) \leq n/3$ and $\mathrm{dist}(\overline{x_{\mathsf{opt}}}, x) \leq n/3$ respectively. Then $\mathrm{dist}(S,S') = n/3$. Choose $|\psi\rangle \equiv |\psi(\beta,\gamma)\rangle$. The $\mathbb{Z}_2$-symmetry of QAOA states gives $P(x) = P(\overline{x})$ and thus $P(S) = P(S')$. We have already shown that $P(S \cup S') \geq 1/2$, see Eq. (27), that is, $P(S) = P(S') \geq 1/4$. Combining this and Eq. (28) one arrives at $1 \leq 48n^{-1/2}2^{3d/2}$. This gives a lower bound on the depth d required to approximate the maximum value $C_{\mathsf{max}} = |E|$ within a ratio

$$1 - \frac{\epsilon n}{|E|} = 1 - \frac{h}{3D} \leq \frac{5}{6} + \frac{\sqrt{D-1}}{3D}.$$

Here we recalled that $\epsilon = h/6$, $|E| = Dn/2$, and used Eq. (24). In Appendix A we show that the level-p QAOA circuit has depth $d \leq p(D+1)$ whenever G is a bipartite degree-D graph.

Thus $1 \leq 48n^{-1/2}2^{3d/2}$ implies $p \geq (1/3\log_2 n - 4)(D+1)^{-1}$. This concludes the proof of Eq. (21).

3.3 Recursive QAOA for Ising-type cost functions

Motivated by the limitations established above, we propose a non-local modification of QAOA which we call the recursive quantum approximate optimization algorithm (RQAOA). This is a VQO-type algorithm based on the variational ansatz Eq. (19) with a constant level p. The key new feature of RQAOA is a variable elimination step. The latter transforms a cost function with n variables to one with $n-1$ variables by examining correlations present in the optimal variational state and identifying strongly correlated clusters of variables. To describe this formally, suppose the cost function $C(x)$ describes the Ising model on a graph $G=(V,E)$ with n vertices. The corresponding n-qubit Hamiltonian is

$$C = \sum_{(j,k)\in E} J_{j,k} Z_j Z_k. \tag{29}$$

Here $J_{j,k}$ are arbitrary real coefficients. As before, our goal is to maximize $C(x) = \langle x|C|x\rangle$. The RQAOA consists of the following steps.

First, maximize the expected value $\langle\psi(\beta,\gamma)|C|\psi(\beta,\gamma)\rangle$ over $\beta,\gamma \in \mathbb{R}^p$. For every edge $(j,k)\in E$, compute the mean value $M_{j,k} = \langle\psi(\beta,\gamma)|Z_jZ_k|\psi(\beta,\gamma)\rangle$.

Next, find an edge $(i,j)\in E$ with the largest magnitude of $M_{i,j}$ (breaking ties arbitrarily). The corresponding variables Z_i and Z_j are correlated if $M_{i,j} > 0$ and anti-correlated if $M_{i,j} < 0$. Impose a parity constraint

$$Z_j = \mathrm{sgn}(M_{i,j}) Z_i \tag{30}$$

and substitute it into the cost function C to eliminate the variable j. For example, a term $J_{j,k}Z_jZ_k$ with $k \notin \{i,j\}$ gets mapped to $J_{j,k}\mathrm{sgn}(M_{i,j})Z_iZ_k$. The term $J_{i,j}Z_iZ_j$ gets mapped to a constant energy shift $J_{i,j}\mathrm{sgn}(M_{i,j})$. All other terms remain unchanged. This yields a new Ising cost function C' that depends on $n-1$ variables. The underlying interaction graph G' with $n-1$ vertices is obtained from G by contracting the edge (i,j) [5]. The maximum energy of C' coincides with the maximum energy of C over the subset of assignments satisfying the constraint Eq. (30).

Next, call RQAOA recursively to maximize the cost function C'. Each recursion step eliminates one variable from the cost function. The recursion stops when the number of variables reaches some specified threshold value $n_c \ll n$. The remaining instance of the problem with n_c variables is then solved by a purely classical algorithm (for example, by a brute force method). Thus the value of n_c controls how the workload is distributed between quantum and classical computers.

Finally, assign a value to all eliminated variables Z_j by backtracking the steps of the algorithm and applying the parity constraints Eq. (30) imposed at each step. This results in a tentative solution $x \in \{0,1\}^n$ of the original problem with n variables.

[5] Recall that contraction of an edge (i,j) in a graph G is an operation that removes the edge (i,j) from G and identifies the vertices i,j.

Importantly, the limitations established above for QAOA with a constant level p on bounded degree graphs do not apply to its recursive version. Indeed, each variable elimination step performed by RQAOA results in a contraction of some edge in the graph. The contraction of edges tends to increase the maximum vertex degree, thereby increasing the circuit depth of level-p variational states. In other words, RQAOA overcomes the locality restriction of the standard QAOA without increasing the number of variational parameters that have to be optimized at each step.

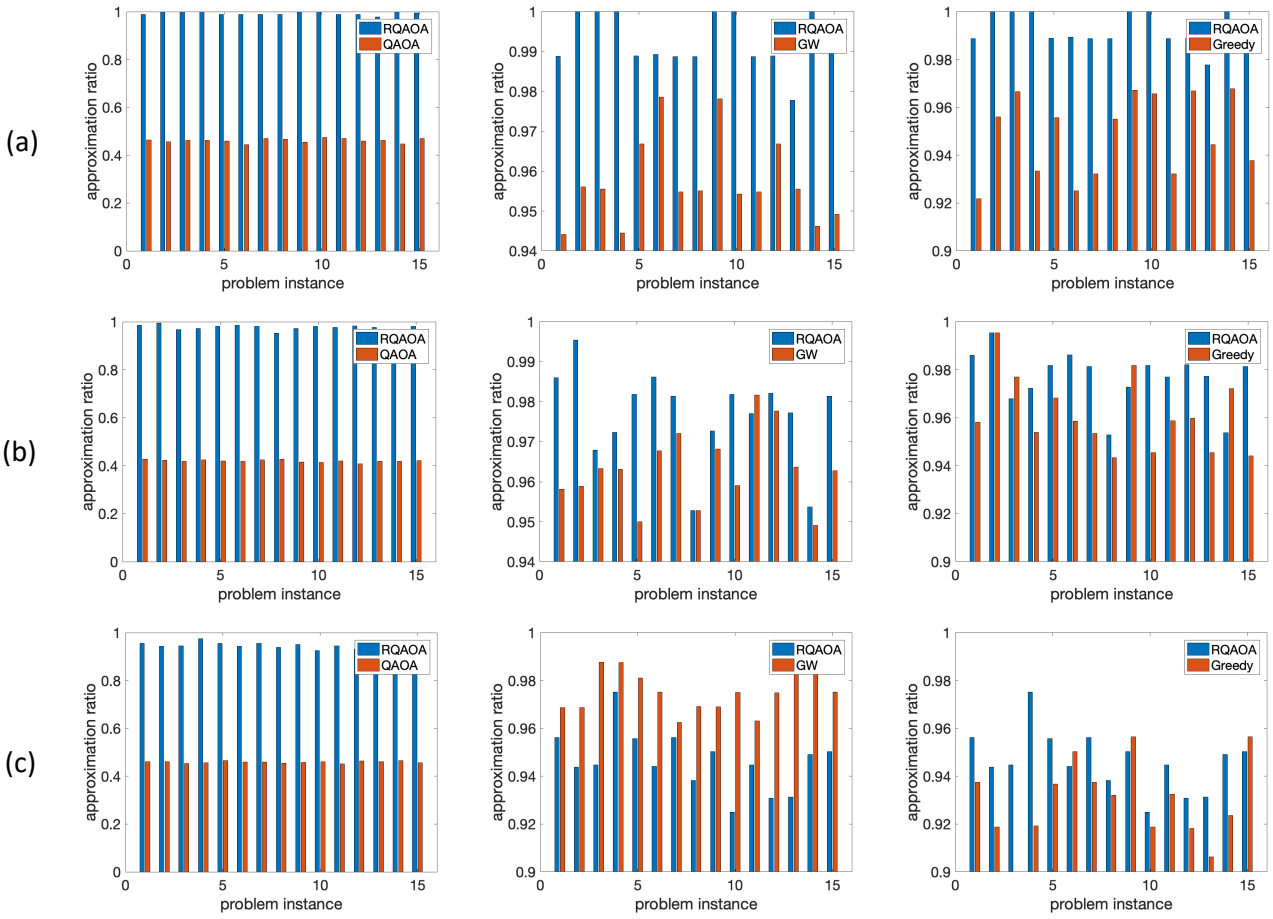


Figure 2: Approximation ratios achieved by the level-1 RQAOA, QAOA, the Goemans-Williamson (GW) algorithm [160], and the greedy algorithm (the classical limit of RQAOA) for 15 instances of the Ising cost function $C = \sum_{(j,k)\in E} J_{j,k} Z_j Z_j$. We consider three choices of the interaction graph: (a,b) the 2D toric grid of size 16×16, (c) random degree-3 graphs. In case (b) the graph has an extra vertex connected to every grid point. This is equivalent to the 2D Ising model with external fields. We choose $J_{j,k} = \pm 1$ at random. The threshold value for RQAOA and the greedy algorithm is $n_c = 20$. Each run of the GW algorithm includes 256 rounding attempts. The exact maximum energy was computed using the integer linear programming.

We report a numerical comparison between the level-1 QAOA, RQAOA, and the Goemans-Williamson algorithm for the Ising cost function Eq. (29) with random coefficients $J_{j,k} = \pm 1$. Three types of interaction graphs are considered:

(a) 2D toric grid of size 16×16

(b) 2D toric grid of size 16×16 with one extra vertex connected to each grid point

(c) random degree-3 graphs with $n = 256$ vertices

Case (b) is equivalent to the 2D Ising model with random ± 1 external fields. The problem of maximizing the energy $C(x)$ admits an efficient algorithm in the 2D case in the absence of external fields and is NP-hard with external fields [187]. To compute the mean values $\langle \psi(\beta,\gamma)|Z_j Z_k|\psi(\beta,\gamma)\rangle$ for medium-size problems with a few hundred variables, we proposed a fast classical algorithm for simulating level-1 QAOA, as detailed in Section **??**. Figure 2 shows approximation ratios achieved by each algorithm for 15 random problem instances. In all three cases we observed that RQAOA is considerably more powerful that the standard QAOA. Surprisingly, in many cases RQAOA also outperforms the Goemans-Williamson algorithm.

A natural question is whether local classical algorithms discussed by Hastings [178] augmented by the variable elimination step can outperform RQAOA. To address this concern, we consider a greedy algorithm obtained from the level-1 RQAOA by taking the limit $\gamma \to 0$ in the variational state. A simple calculation shows that

$$M_{i,j} = \langle \psi(\beta,\gamma)|Z_i Z_j|\psi(\beta,\gamma)\rangle = \gamma \sin(4\beta) J_{i,j} + O(\gamma^2).$$

Thus $M_{i,j} \sim J_{i,j}$ in the limit $\gamma \to 0$. The greedy algorithm follows the same steps as the level-1 RQAOA except that the maximally correlated pair of variables Z_i, Z_j is chosen as the one maximizing $|J_{i,j}|$. We observed that RQAOA tends to outperform the greedy algorithm with the largest performance gain in the case (a), see Figure 2.

Finally, we showed analytically that RQAOA with the level $p = 1$ finds the optimal solution for the so-called ring of disagrees model, see Appendix D of [159]. Meanwhile, the standard level-p QAOA achieves approximation ratio at most $(2p+1)/(2p+2)$ for this model [188]. This proves that in certain cases RQAOA is strictly more powerful than QAOA.

3.4 Recursive QAOA for graph coloring problems

In this section we apply RQAOA to the Maximum k-Cut problem which is an optimization version of the graph k-coloring problem. Suppose $G = (V, E)$ is a graph with $n = |V|$ vertices and $e = |E|$ edges. Given an integer $k \geq 2$, the goal is to find an approximate k-coloring of vertices of G which maximizes the number of edges whose endpoints have different colors. For each vertex $j \in V$ introduce a variable $x_j \in [k]$ which represents a color assigned to j. The k-coloring cost function is defined as

$$C(x) = \sum_{(i,j)\in E} (1 - \delta_{x_i,x_j}), \qquad x \in [k]^n. \tag{31}$$

It can also be viewed as an anti-ferromagnetic k-state Potts model. The standard MaxCut problem corresponds to $k = 2$. Consider first a special case when G is a k-colorable graph. Clearly, a random uniform assignment of colors x achieves an approximation ratio of $1-1/k$ on average. For the case where k is a power of two, Cho, Raje and Sarrafzadeh [189] constructed an $O((e+n)\log k)$-time algorithm which achieves an approximation ratio of $1-1/k(1-1/n)^{\log k}$, improving upon a deterministic $O(enk)$-time algorithm [190] achieving the same ratio $1-1/k$ as random coloring (and obtained by derandomizing the latter). In the general case, when G may not be k-colorable, Frieze and Jerrum [162] gave an algorithm achieving an approximation ratio $1-1/k+2\log(k)/k^2$ for sufficiently large k. This is known to be close to optimal since no polynomial-time algorithm can achieve an approximation ratio better than $1-1/(34k)$ unless $P = NP$ [191]. Frieze and Jerrum's algorithm is based on an SDP relaxation and a randomized rounding scheme inspired by Goemans-Williamson algorithm for MaxCut. It comes with detailed estimates for the approximation ratio $\alpha_k = \mathbb{E}(C(x))/C_{\mathsf{max}}$ for small k, namely

$$\begin{aligned}\alpha_2 &\geq 0.878567\\ \alpha_3 &\geq 0.800217\\ \alpha_4 &\geq 0.850304\\ \alpha_5 &\geq 0.874243\end{aligned}$$

Here the bound on α_2 matches the Goemans-Williamson algorithm for MaxCut. This was further improved by Klerk et al. [192] to

$$\alpha_3 \geq 0.836008 \tag{32}$$

$$\alpha_4 \geq 0.857487\ . \tag{33}$$

We extend RQAOA to Max-k-Cut as follows. Consider a system of n k-dimensional qudits and a cost function Hamiltonian

$$C = \sum_{(i,j)\in E}\ \sum_{b\in\mathbb{Z}_k} J_{i,j}(b)\Pi_{i,j}(b), \tag{34}$$

where $J_{i,j}(b)$ are real coefficients and

$$\Pi(b) = \sum_{a\in\mathbb{Z}_k} |a\rangle\langle a| \otimes |a+b\rangle\langle a+b|$$

is a diagonal projector acting on $\mathbb{C}^k \otimes \mathbb{C}^k$. Here and below the addition of color indices is performed modulo k. The subscripts i, j in $\Pi_{i,j}(b)$ indicate the pair of qudits acted upon by the projector $\Pi(b)$. By definition, $\Pi_{i,j}(b) = \Pi_{j,i}(-b)$. The k-coloring cost function is a special case of Eq. (34) when $J_{i,j}(b) = 1-\delta_{b,0}$. We shall see the family of Hamiltonians defined in Eq. (34) is closed under the variable elimination. The Hamiltonian Eq. (34) commutes with the symmetry operator $X^{\otimes n}$, where

$$X = \sum_{b\in\mathbb{Z}_k} |b+1\rangle\langle b|$$

is the generalized Pauli-X operator. This is analogous to the $\mathbb{Z}_2$-symmetry discussed in Section 3.2.

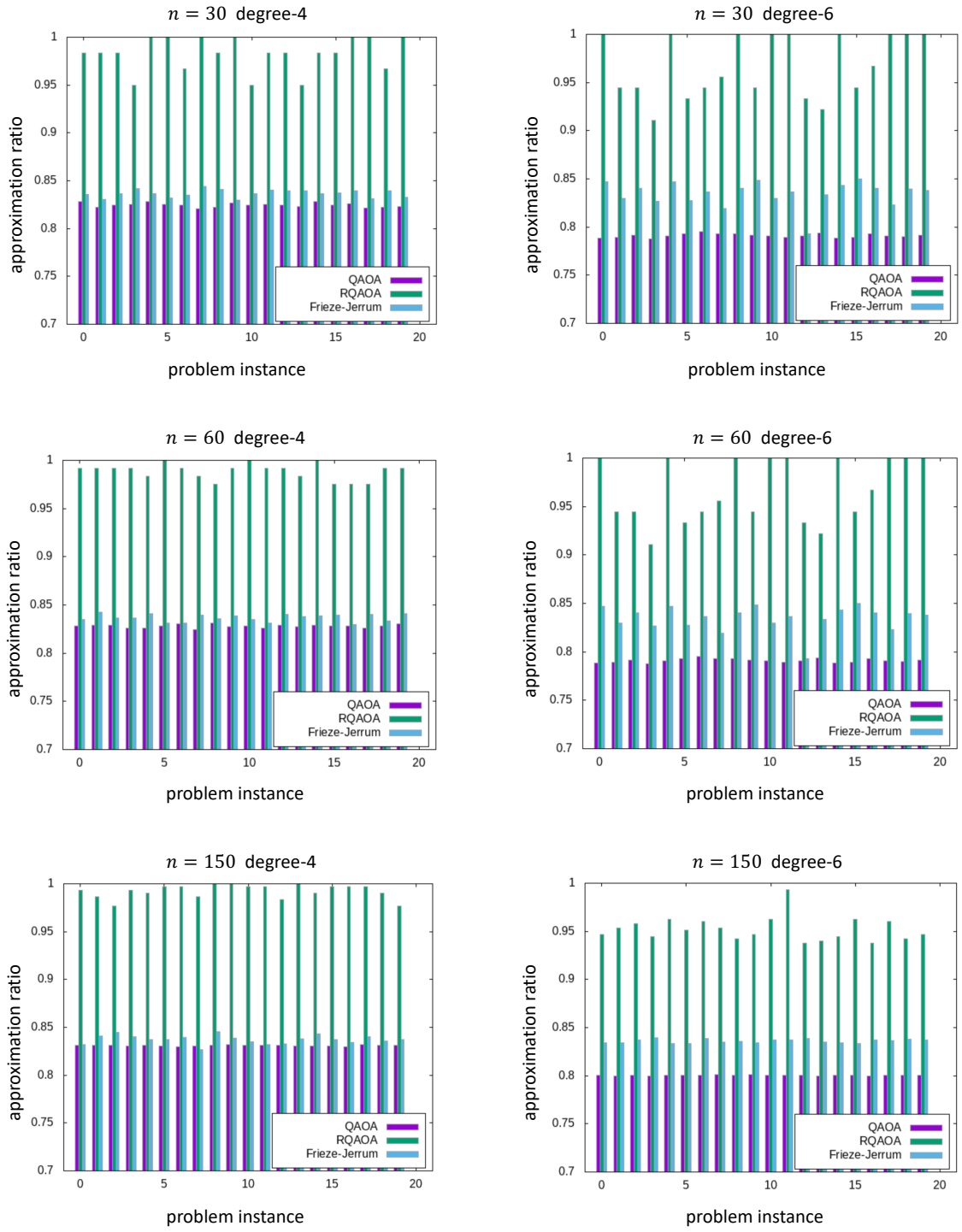


Figure 3: We consider the Max-3-Cut problem on 20 random instances of 3-colorable graphs with $n \in \{30, 60, 150\}$ vertices. Approximation ratios achieved by the level-1 QAOA (red), RQAOA (green), and the SDP relaxation method by Frieze and Jerrum [162] (blue) are shown. *Left:* 4-regular graphs, *Right:* 6-regular graphs.

We shall now generalize the QAOA ansatz to qudits. Let $|+\rangle \in \mathbb{C}^k$ be the $+1$ eigenvector of X, that is, $|+\rangle = k^{-1/2} \sum_{a \in \mathbb{Z}_k} |a\rangle$. Given a real vector $\beta = (\beta^1, \ldots, \beta^k)$, define a unitary $B(\beta) : \mathbb{C}^k \to \mathbb{C}^k$ which is diagonal in the X-basis such that

$$B(\beta) = \sum_{a \in \mathbb{Z}_k} e^{i\beta^a} |\phi_a\rangle\langle\phi_a|, \qquad |\phi_a\rangle \equiv Z^a|+\rangle.$$

The level-p QAOA variational state for k-dimensional qudits is defined as

$$|\psi(\beta,\gamma)\rangle = \prod_{t=1}^{p} B(\beta_t) e^{i\gamma_t C} |+^n\rangle \tag{35}$$

where $\beta_t \in \mathbb{R}^k$ and $\gamma_t \in \mathbb{R}$ are variational parameters. Note that each term $B(\beta_t)$ contains only $k-1$ parameters, since the overall phase of $B(\beta_t)$ is irrelevant. Thus the total number of parameters is $(k-1)p + p = kp$. This reduces to the standard definition Eq. (19) in the case $k = 2$. The RQAOA consists of the following steps.

First, maximize the expected value $\langle\psi|C|\psi\rangle$ with $|\psi\rangle \equiv |\psi(\beta,\gamma)\rangle$ over β and γ. For every edge $(i,j) \in E$ compute the mean value $M_{i,j}(b) = \langle\psi|\Pi_{i,j}(b)|\psi\rangle$. Note that $0 \le M_{i,j}(b) \le 1$ since $\Pi_{i,j}(b)$ is a projector.

Next, find an edge $(i,j) \in E$ and a color $b \in [k]$ with the largest magnitude of $M_{i,j}(b)$ (breaking ties arbitrarily). Impose a constraint

$$x_j = x_i + b \pmod{k} \tag{36}$$

which is analogous to the parity constraint Eq. (30). Note that $|\psi\rangle$ has support on basis states satisfying the constraint Eq. (36) iff $M_{i,j}(b) = 1$. Substitute the constraint Eq. (36) into the cost function Hamiltonian to eliminate the variable x_i. To this end, use the identity

$$\Pi_{i,j}(b)\Pi_{i,h}(a) = \Pi_{i,j}(b)\Pi_{j,h}(a-b)$$

which holds for any $h \notin \{i,j\}$. Thus $\Pi_{i,h}(a) = \Pi_{j,h}(a-b)$ on the subspace satisfying the constraint. Replacing $\Pi_{i,h}(a)$ by $\Pi_{j,h}(a-b)$ in the cost function for all $h \notin \{i,j\}$ one gets a new cost function C' of the form Eq. (34) acting on $n-1$ variables. The maximum energy of C' coincides with the maximum energy of C over the subset of assignments satisfying the constraint Eq. (36).

Finally, call RQAOA recursively, in the same fashion as was done for the Ising cost functions, see Section 3.3.

We benchmarked level-1 RQAOA for random D-regular graphs generated as follows.

1. Pick a partition $V = V_1V_2V_3$ such that $|V_1| = |V_2| = |V_3| = n/3$.

2. For each r, s such that $1 \le r < s \le 3$ join V_r and V_s by a random bipartite $(D/2)$-regular graph with vertex set $V_r \cup V_s$.

3. Check if the obtained graph contains a triangle. If no triangle is found, start over.

Steps 1-2 ensure that the resulting graph is 3-colorable while Step 3 ensures that the graph cannot be colored with fewer than three colors.

Our numerical results for 20 random instances of Max-3-Cut constructed as above are presented on Figure 3. The plots show approximation ratios achieved by QAOA and RQAOA with the level $p = 1$, and the SDP relaxation method by Frieze and Jerrum [162]. It can be seen that RQAOA achieves the best performance among all considered algorithms.

A natural question is how much the performance of RQAOA improves as one goes from the level $p = 1$ to higher levels. A preliminary work shows that expected values of local observables associated with the level-2 QAOA states can be computed classically in time $poly(k, n)$ for any planar graph. Since the variable elimination is equivalent to contraction of edges in the underlying graph, it preserves the planarity. Thus level-2 RQAOA can be simulated classically in time $poly(k, n)$. We are planning to perform a numerical simulation of the level-2 RQAOA for medium-size problems in the future work.

3.5 Twisted QAOA

For the problem of finding (or approximating) the maximum of a combinatorial cost function $C : \{0, 1\}^n \to \mathbb{R}$ (given by polynomially many terms), typical hybrid algorithms proceed by defining the cost function Hamiltonian

$$H_C = \sum_{z \in \{0,1\}^n} C(z)|z\rangle\langle z| \tag{37}$$

and a parametrized family $\{U_G(\theta)\}_{\theta \in \Theta}$ of n-qubit unitary circuits. The later might be parametrized by the underlying graph G of the cost function or in case of hardware-efficient algorithms tailored to the physical device [193]. The parametrized family give rise to variational ansatz states

$$|\Psi(\theta)\rangle = U_G(\theta)\,|0\rangle^{\otimes n} \ . \tag{38}$$

that can be prepared with $U_G(\theta)$ from a product state $|0\rangle^{\otimes n}$. Measuring $\Psi(\theta)$ in the computational basis then provides a sample $z \in \{0, 1\}^n$ from the distribution $p(z) = |\langle z|\Psi(\theta)\rangle|^2$ such that the expectation value of the associated cost function is equal to the energy $\mathbb{E}\,[C(z)] = \langle\Psi(\theta)|\,H_C\,|\Psi(\theta)\rangle$ of the state $\Psi(\theta)$ with respect to H_C. Thus the problem of maximizing C is translated to that of finding a value of the (vector of) parameters θ maximizing the energy of $\Psi(\theta)$. The latter step is envisioned to be performed e.g., by numerical gradient descent or a similar classical procedure prescribing (iteratively) what parameters θ to try. The computation of this prescription (according to obtained measurement results) is the classical processing part of the quantum algorithm leading to the term *hybrid*. We will refer to this form of algorithm as a "bare" hybrid algorithm in the following.

The potential utility of this approach hinges on a number of factors. Of primary importance – beyond questions of convergence or efficiency – is whether the family $\{\Psi(\theta)\}_{\theta \in \Theta}$ of states is sufficiently rich to variationally capture the (classical) correlations of high-energy states of H_C. There is an inherent tension here between the requirement of applicability using near-term devices, and the descriptive power, i.e., required complexity of these states: On the one hand, each unitary $U_G(\theta)$ is supposed to be realized by a low-depth circuit with

local gates (making it amenable to experimental realization on a near-term device), and the dimensionality of the parameter or "search" space Θ should be low to guarantee fast convergence e.g., of gradient descent. On the other hand, states having high energy with respect to H_C and belonging to the considered family of variational states may have intrinsically high circuit complexity, and, correspondingly, may also require a large number of variational parameters to approximate. The unavoidability of this issue has been demonstrated using the MaxCut-problem on degree-d expander graphs with n vertices and the quantum approximate optimization algorithm (QAOA) at level p: Here the parameter space is $\Theta = [0, 2\pi)^{2p}$ and the corresponding circuits $U_G(\theta)$ have depth $O(pd)$. Locality and symmetry of the ansatz imply that achievable expected approximation ratios are upper bounded by a constant (below that achieved by Goemans-Williamson) unless $p = \Omega(\log n)$ [25], see Section 3.2. In fact, the locality of the ansatz alone implies that for smaller values of p, the achieved expected approximation ratio is not better then of a random guessing for random bipartite graphs, as shown in [194].

These fundamental limitations of "standard" hybrid algorithms are tied to the assumption that an increased complexity of the required quantum operations is unacceptable and/or infeasible in the near term. Under these circumstances, the only way forward appears to be to use alternative, possibly more powerful (e.g., non-local) efficient classical processing which could exploit the limited available quantum resources more effectively. One example where a classical post-processing is used is [195], where QAOA is combined with a greedy "pruning" method to produce an independent set of large size. Here post-processing is needed, in particular, to ensure that the output is indeed an independent set. Another proposal in this direction is the idea of "warm-starting" QAOA with a solution provided by the Goemans-Williamson algorithm [196] (see also [197]). The warm-starting approach has the appeal that – by construction – the Goemans-Williamson approximation ratio can be guaranteed in this approach (assuming convergence of the energy optimization). An alternative is the recursive QAOA method [159, 26] which uses QAOA states to iteratively identify variables to eliminate, see Sections 3.3,3.4. This effectively reduces the problem size but increases the connectivity and thus the circuit complexity of the iteratively obtained subproblems. Furthermore, analytical bounds on the expected approximation ratios are unknown except for very special examples [25]. For both warm-starting QAOA as well as RQAOA, one deviates from the original QAOA ansatz, leading to different variational states and corresponding quantum circuits.

Main results. Here we consider arguably more minimal adaptions of hybrid variational algorithms for the MaxCut-problem on 3-regular graphs. For a given bare hybrid algorithm $\mathcal{A}$ involving a family $\{\Psi(\theta)\}_{\theta\in\Theta}$ of variational ansatz states as described above, we show how to construct a modified algorithm $\mathcal{A}^+$ which uses the *same* family of states $\{\Psi(\theta)\}_{\theta\in\Theta}$. The algorithm $\mathcal{A}^+$ will be called *twisted-*$\mathcal{A}$. It requires a set of quantum operations that are comparable (in number and complexity) to that of $\mathcal{A}$. In particular, it involves preparing the states $\{\Psi(\theta)\}_{\theta\in\Theta}$. In addition, $\mathcal{A}^+$ uses extra local measurements because the hybrid optimization step is modified: the energy to be optimized is given by a modified problem Hamiltonian H_G^+ rather than the MaxCut-problem Hamiltonian H_G associated with the considered graph G. The modified Hamiltonian H_G^+ is either a 3- or 4-local Hamiltonian and

(as H_G) diagonal in the computational basis. In particular, this means that measurements of up to 4 qubits at a time in the computational basis are sufficient to determine the (expected) cost function.

By construction, the algorithms $\mathcal{A}$ and $\mathcal{A}^+$ achieve (expected) cut sizes (for any fixed instance G) related by the inequalities

$$\mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}(G)\right)\right] \leq \mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}^+(G)\right)\right] \tag{39}$$

for any (bare) hybrid algorithm $\mathcal{A}$, assuming that the optimal parameters are found in the optimization step. Indeed, (39) follows because, denoting with

$$\theta_* = \arg\max_\theta \langle\Psi(\theta)| H_G |\Psi(\theta)\rangle \tag{40}$$

the optimal parameters for the Hamiltonian H_G, we have by definition of the algorithms that

$$\begin{array}{rcl} \mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}(G)\right)\right] & = & \langle\Psi(\theta_*)| H_G |\Psi(\theta_*)\rangle \\ \mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}^+(G)\right)\right] & = & \max_\theta \langle\Psi(\theta)| H_G^+ |\Psi(\theta)\rangle \ , \end{array} \tag{41}$$

and

$$H_G^+ = H_G + \Delta_G \ , \tag{42}$$

where Δ_G is a sum of non-negative local operators. These considerations apply to any bare hybrid algorithm $\mathcal{A}$.

Our modified algorithms are directly motivated by the work of Feige, Karpinski, and Langberg [198] (referred to as FKL in the following). These authors propose an algorithm for the MaxCut problem on 3-regular graphs which proceeds by solving a semidefinite program relaxation (similar to Goemans and Williamson), and subsequently improving the rounded solution by a simple greedy post-processing technique. We also consider the improvement by Halperin, Livnat, and Zwick [199] (referred to as HLZ below) which involves a more non-local greedy procedure.

Consider a simple motivational example of a greedy post-processing procedure that can improve a given cut. The input will be a 3-regular graph $G = (V, E)$ and a cut. We say that a vertex is unsatisfied when all three of its neighbours lie in the same partition of the cut as it does. The algorithm will repeatedly run through the vertices and check whether some of them are unsatisfied. If it finds an unsatisfied vertex it moves it to the opposite side of the cut and repeats the process with the updated cut until none of the vertices is unsatisfied. Since moving one vertex increases the cut size by 3 and potentially lowers the number of unsatisfied vertices by 4, one can show that this procedure improves the cut size by at least $\frac{3}{4}$ times the number of unsatisfied vertices in the initial cut. Let us apply this greedy procedure to a random cut, which has an expected approximation ratio of 1/2. A vertex will be unsatisfied with probability 2^{-3}. From the linearity of expectation we have that the greedy procedure will improve the cut by at least $\frac{3}{4\cdot 8}|V|$. Since $|V| = \frac{2}{3}|E|$, we achieve approximation ratio at least $\frac{1}{2} + \frac{1}{16} = 0.5625$ in expectation.

We combine these techniques with a hybrid algorithm $\mathcal{A}$ such as level-p QAOA (in the following denoted by QAOA_p), giving a "twisted" hybrid algorithm $\mathcal{A}^+$. The algorithm $\mathcal{A}^+$ proceeds by using the variational family of states defined by the algorithm $\mathcal{A}$

to obtain an approximate cut, but this step is modified or "twisted", as discussed below. The algorithm $\mathcal{A}^+$ then attempts to enlarge the cut size of the obtained cut by applying a classical post-processing procedure: We perform either the post-processing procedure by Feige, Karpinski, and Langberg (obtaining an algorithm FKL-$\mathcal{A}^+$) or the post-processing procedure by Halperin, Livnat, and Zwick (giving an algorithm HLZ-$\mathcal{A}^+$).

Let us now describe the sense in which $\mathcal{A}^+$ is a "twisted" form of $\mathcal{A}$ and not merely a hybrid algorithm augmented by a subsequent classical post-processing step. This terminology stems from the fact that in the quantum subroutine of the algorithm, the variational parameters (angles) are not optimized with respect to the original problem Hamiltonian H_G. Instead, one can express the expected cut size produced by measuring a state $\Psi(\theta)$ and using classical post-processing by the expectation value of a modified Hamiltonian H_G^+ (for both FKL and HLZ) in the variational state $\Psi(\theta)$. The twisted algorithm $\mathcal{A}^+$ thus optimizes the angle θ with respect to the modified Hamiltonian H_G^+. Importantly, this does not change the ansatz/variational family of states used. This allows us to make a fair comparison (in terms of quantum resources and, especially, the number of variational parameters) to the original algorithm $\mathcal{A}$.

We specialize our considerations to QAOA_p and establish lower bounds on the approximation ratio for bare and twisted QAOA, i.e., we consider the algorithms QAOA_p and QAOA_p^+. Specifically, we consider low values of p for 3-regular graphs, triangle-free 3-regular graphs and high girth 3-regular graphs. We denote the expected approximation ratio achieved by an algorithm $\mathcal{A}$ on a graph G with maximum cut size $\mathrm{MaxCut}(G)$ by

$$\alpha_G(\mathcal{A}) := \frac{\mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}(G)\right)\right]}{\mathrm{MaxCut}(G)}. \tag{43}$$

In the following, we will refer to the expected approximation ratio achieved by an algorithm $\mathcal{A}$ simply as the approximation of $\mathcal{A}$ (omitting the term "expected") unless specified otherwise. In the case of $\mathcal{A} = \mathrm{QAOA}_p$, $\mathbb{E}\left[\mathsf{cutsize}\left(\mathcal{A}(G)\right)\right]$ is defined as in (41), but with the level-p QAOA trial function $\Psi_G(\beta,\gamma)$, $\beta,\gamma \in [0,2\pi)^p$ instead of $\Psi(\Theta)$.

Our results are summarized in Figure 4, which gives our lower bounds on the approximation ratio for each of these methods. For comparison, we also state the following known bounds on bare QAOA for any 3-regular graph G,

$$\begin{array}{lcll} \alpha_G(\mathrm{QAOA}_1) & \geq & 0.6924 & \text{established in [200]} \\ \alpha_G(\mathrm{QAOA}_2) & \geq & 0.7559 & \text{conjectured in [200], established in [201]} \\ \alpha_G(\mathrm{QAOA}_3) & \geq & 0.79239 & \text{conjectured in [201].} \end{array} \tag{44}$$

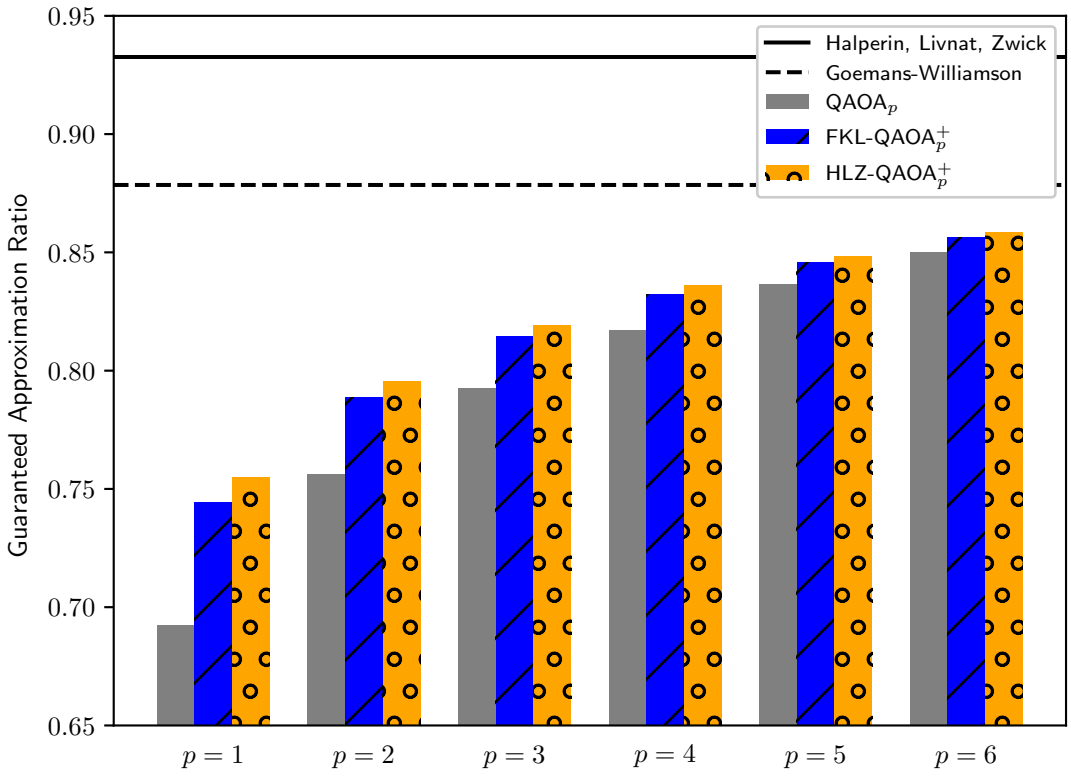
Also shown in Figure 4 are the guaranteed approximation ratios of the best-known classical algorithms: This includes the Goemans-Williamson algorithm (GW) for general graphs (which is optimal when assuming the unique games conjecture [202]) which achieves

$$\alpha_G(\mathrm{GW}) \geq 0.8785 \qquad \text{for any graph } G \qquad \text{(see [203])}\ . \tag{45}$$

For 3-regular graphs, the best efficient classical algorithms are the algorithm by Feige, Karpinski, and Langberg [198] which relies on a semidefinite program whose solution is then improved by a simple greedy post-processing technique, and a refinement of this technique

by Halperin, Livnat, and Zwick [199]. They achieve

$$\begin{array}{rcl} \alpha_G\,(\mathrm{FKL}) & \geq & 0.924 \\ \alpha_G\,(\mathrm{HLZ}) & \geq & 0.9326 \end{array} \qquad \text{for any 3-regular graph } G \qquad \begin{array}{l} \text{see [198]} \\ \text{see [199].} \end{array} \tag{46}$$



Method	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
Bare QAOA$_p$	**0.6924** [200]	**0.7559** [201]	0.7923 [201]	0.8168	0.8363	0.8498
FKL-QAOA$_p^+$	**0.7443**	0.7887	0.8146	0.8323	0.8457	0.8564
HLZ-QAOA$_p^+$	0.7548	0.7954	0.8191	0.8358	0.8482	0.8582

Figure 4: The main results of this work. We compare the provably guaranteed approximation ratios of bare QAOA$_p$, FKL-QAOA$_p^+$ and HLZ-QAOA$_p^+$ for the MaxCut problem specialized to 3-regular graphs with girth greater than $2p+2$. Numbers written in boldface also apply to general 3-regular graphs. All quantities are rounded down to four decimals. Guaranteed approximation ratios which have been established in other work are indicated with citations.

We find that going from the original QAOA to its twisted version leads to a significant improvement, roughly saving one level p: We approximately have

$$\alpha_G\left(\mathrm{QAOA}_{p-1}^+\right) \gtrsim \alpha_G\left(\mathrm{QAOA}_p\right) \qquad \text{for} \qquad p=2,\ldots,6\ . \tag{47}$$

Open questions. One potential avenue to obtaining improved approximation ratios with hybrid algorithms is to use a different variational family of ansatz states. Here our work gives clear guidance when this is combined with classical post-processing: For a graph G, the energy of a modified cost function Hamiltonian $H_G^+ = H_G + \Delta_G$ should be optimized instead of that of H_G. In particular, since Δ_G is a sum of 3-local terms in the case of FKL and a sum of 4-local terms in the case of HLZ, this motivates introducing new terms (e.g., proportional to these terms) in the QAOA ansatz such that entangling layers $\exp[-i\gamma_j H_G]$ in the variational circuit are replaced by $\exp[-i\gamma_j H_G^+]$. Such a modification of the algorithm is superficially related to the fact that the classical (randomized rounding-based) algorithms of [198, 199] also use additional (3-variable) constraints in the SDP compared to the Goemans-Williamson algorithm. We note, however, that using different variational ansatz states will require a different accounting of resources (e.g., circuit depth). In contrast, our twisted algorithms use the same circuits to prepare ansatz states as their bare version.

Another promising approach may be to combine warm-starting-type ideas with classical post-processing. Here one could consider algorithms that first solve the SDP underlying the classical algorithms [198, 199], and subsequently prepare a corresponding quantum state. One may hope that – similar to [196] – suitably designed approaches give a guaranteed approximation ratio matching that of these classical algorithms.

Moving beyond combinatorial optimization problems, it is natural to ask if variational quantum algorithms for many-body quantum Hamiltonian problems (e.g., quantum analogues of MaxCut as considered in [204]) can be improved by similar greedy (quantum) post-processing procedures.

3.6 Quantum algorithms for classical partition functions

Recall that the Gibbs distribution of a classical Hamiltonian $H : \Omega \to \mathbb{R}$, where Ω is a finite set, at inverse temperature β is defined as

$$\mu_\beta(x) := (1/Z(\beta)) \exp(-\beta H(x))$$

for $x \in \Omega$. Below we assume that

$$0 \leq H(x) \leq n$$

for all $x \in \Omega$. The normalization factor $Z(\beta) := \sum_{x\in\Omega} e^{-\beta H(x)}$ is the partition function. We study algorithms for computing multiplicative approximations to $Z(\beta)$: given an inverse temperature β and a precision parameter $\epsilon > 0$, our algorithms produce an estimate $\hat{Z}$ such that

$$(1-\epsilon)Z(\beta) \leq \hat{Z} \leq (1+\epsilon)Z(\beta)$$

using as few samples from the Gibbs distribution as possible. In the quantum setting we assume to have access to qsamples, i.e., coherent encodings of the Gibbs distribution

$$|\mu_\beta\rangle = \sum_{x\in\Omega} \sqrt{\mu_\beta(x)}|x\rangle.$$

Simulated annealing is a MCMC algorithm that approaches this problem by expanding the target partition function into a telescoping product:

$$Z(\beta) = Z(0) \cdot (Z(b_1)/Z(0)) \cdot (Z(b_2)/Z(b_1)) \cdots (Z(\beta)/Z(b_\ell)),$$

where $0 \leq b_1 \leq b_2 \leq \cdots \leq \beta$ is a sequence of inverse temperatures called the *cooling schedule*. If each term in the telescoping product can be approximated sufficiently well, then we obtain $\hat{Z}$ to the required precision. The key idea of the simulated annealing algorithm is to estimate each term by random (or quantum) sampling. We propose classical and quantum simulated annealing algorithms and achieve the following:

1. We give a quantum algorithm based on the simulated annealing scheme of [164] (*Algorithm Q1*). This is a major improvement compared to the previously best known approach of Montanaro [172]: Montanaro relies on the classical cooling schedule computation procedure of [164], while our work gives a quantum algorithm for the same task. Our key technical ingredient is a non-destructive version of the amplitude estimation algorithm of [169] which may be of independent interest.

2. We simplify the classical algorithm of [164], improving their sample complexity to almost match that of the current classical state of the art. The key ingredient of our method is the paired-product estimator of Huber [205]. We subsequently quantize this algorithm and obtain the best known quantum algorithm for computing partition functions of classical Hamiltonians (*Algorithm Q2*).

Our quantum algorithms achieve two improvements compared to classical ones: a quadratic speedup both in the spectral gap of the Markov chain and the precision, as well as a shorter cooling schedule. *Algorithm Q2* is the best known algorithm for approximating Gibbs partition functions in terms of sample complexity. Our results are summarized below.

Algorithm	**Type**	**Sample Complexity**	**Cooling schedule length**
SVV [164]	Classical	$O(\ln \lvert\Omega\rvert \cdot (\ln\ln \lvert\Omega\rvert + \ln n)^5 \cdot \varepsilon^{-2})$	$O(\sqrt{\ln \lvert\Omega\rvert} \cdot \ln n \cdot \ln\ln \lvert\Omega\rvert)$
Huber [205, 206]	Classical	$O(\ln \lvert\Omega\rvert \cdot \ln n \cdot \varepsilon^{-2})$	$O(\ln \lvert\Omega\rvert)$
Montanaro [172]	Hybrid	$O(\ln \lvert\Omega\rvert \cdot (\ln\ln \lvert\Omega\rvert + \ln n)^{5/2} \cdot \varepsilon^{-1})^\star$	$O(\sqrt{\ln \lvert\Omega\rvert} \cdot \ln n \cdot \ln\ln \lvert\Omega\rvert)$
Algorithm Q1	Quantum	$O(\ln \lvert\Omega\rvert \cdot (\ln\ln \lvert\Omega\rvert + \ln n)^{5/2} \cdot \varepsilon^{-1})$	$O(\sqrt{\ln \lvert\Omega\rvert} \cdot \ln n \cdot \ln\ln \lvert\Omega\rvert)$
Algorithm Q2	Quantum	$O(\ln \lvert\Omega\rvert \cdot \ln n \cdot \varepsilon^{-1})$	$O(\sqrt{\ln \lvert\Omega\rvert \cdot \ln n})$

Table 2: Sample complexity and cooling schedule length of our algorithms and selected prior work. Montanaro's algorithm [172] uses the classical SVV algorithm [164] to generate a cooling schedule. As a consequence, its time complexity scales as $1/\delta$, where δ is the minimum spectral gap of the Markov chains. In contrast, *Algorithms Q1 and Q2* compute the cooling schedule quantumly and achieve the time complexity $1/\sqrt{\delta}$.

Algorithm Q1 is a quantum simulated annealing algorithm for approximating partition functions with quadratically better dependence on the Markov chain spectral gap than any classical algorithm. We build on two previous algorithmic results. The first is the Quantum Simulated Annealing (QSA) algorithm of [171] that shows that, given a series of ℓ Markov chains such that the first Markov chain is easy to qsample, all the spectral gaps are lower bounded by δ, and the stationary states have constant overlap, qsampling from the last Markov chain can be performed using $\tilde{O}(\ell/\sqrt{\delta})$ Markov chain steps. As was noted in [164], a better sequence of Markov chains can be found if we choose them *adaptively*; based on information extracted from samples as we run the algorithm. Using this observation, they

gave a classical algorithm for finding sequences of Markov chains with an almost quadratic improvement in the sequence length compared to the best non-adaptive solution. As the early quantum simulated annealing algorithms [171, 207, 208] were discovered before the invention of adaptive simulated annealing, it is natural to ask whether this observation can be used to design better quantum simulated annealing algorithms. At first, adaptive algorithms appear difficult to quantize since extracting information from qsamples – say in order to determine the adaptive sequence – will generally damage the states, seemingly increasing the sample complexity. Indeed, the only quantum algorithm that previously considered adaptive schedules was Montanaro's [172] algorithm that computed the cooling sequence classically. Our work circumvents this deficiency by fully quantizing the cooling schedule generation part of [164]. This way, we obtain a schedule matching the length from SVV but also retain the scaling with $1/\sqrt{\delta}$ from previous QSA algorithms [207, 208, 171]. In doing so we show that amplitude estimation [169] can be made nondestructive using a state restoration scheme inspired by [209]. We believe that this result will be useful in its own right.

Algorithm Q2 is based on the *paired-product estimator* of [205] that allows us to use shorter cooling schedules. A major advantage of this improvement is that it clearly decouples the cooling schedule generation from the final partition function approximation. In [164], the final sample complexity is a product of two main terms – the complexity of the schedule generation algorithm, independent of ε and the complexity of estimating the ratios in the telescoping product. In our algorithm, these terms are added. As it is reasonable to assume that ε is "reasonably small" in practice, the complexity is dominated by the term which depends only on the length of the cooling schedule and ε: the time to generate the schedule becomes asymptotically negligible. We quantize this improved classical algorithm to obtain the fastest known algorithm for partition function approximation. First, we show that there is a simple procedure based on binary search for obtaining a cooling schedule of length $O(\sqrt{\ln|\Omega| \cdot \ln n})$ for any inverse temperature β. This is a $\sqrt{\ln n}$ factor shorter than the classical cooling sequence and matches the $\Theta(\sqrt{\ln|\Omega| \cdot \ln n})$ complexity conjectured to be optimal in [164]. To estimate the ratios in the telescopic product, we give a simple quantum algorithm for the estimation of the expected value of random variables with bounded relative variance. In particular, for a distribution D, and a random variable V sampled from D satisfying $\mathbb{E}[V^2]/\mathbb{E}[V]^2 \leq B$, we describe a quantum algorithm that uses $O(B)$ copies of a qsample $|\psi_D\rangle = \sum_x \sqrt{D(x)}\,|x\rangle$, and $\tilde{O}\big(\sqrt{B}/\varepsilon\big)$ reflections about $|\psi_D\rangle$, to obtain with high probability an ε-relative approximation of $\mathbb{E}[V]$; additionally, the algorithm restores one copy of $|\psi_D\rangle$. Our algorithm is based on the work of Montanaro [172] and improves upon its scaling from $O(B/\varepsilon)$ to $O(\sqrt{B}/\varepsilon)$.

3.7 Space-efficient quantization of reversible Markov chains

In a seminal paper [170] Szegedy showed how to construct a quantum walk $W(P)$ for any reversible Markov chain P such that its eigenvector with eigenphase 0 is a quantum sample of the limiting distribution of the random walk and its eigenphase gap is quadratically larger than the spectral gap of P. This quadratic gap amplification is at the heart of the quantum speed-ups of many classical random walk based algorithms. The standard construction of Szegedy's quantum walk requires an ancilla register of Hilbert space dimension equal to the

size of the state space of the Markov chain. The team showed [210] that it is possible to avoid this doubling of state space for certain Markov chains that arise naturally in many applications. For such Markov chains, we presented a quantization method which requires an ancilla register of dimension equal to only the number of different energy values, which is often significantly smaller than the size of the state space. This could help realize quantum walk based algorithms on smaller quantum hardware. To reduce the required quantum memory, the team developed a novel technique for block encoding Hadamard products of matrices which may be of wider interest for designing quantum algorithms.

3.8 Quantum subroutines for the simplex algorithm

The simplex method solves the following linear optimization problem: minimize $c^\top x$ over $x \in \mathbb{R}^n$ subject to $Ax = b$ and $x \geq 0$, where $A \in \mathbb{R}^{m\times n}$, $c \in \mathbb{R}^n, b \in \mathbb{R}^m$. Let us first introduce some notations and terminology. We assume that $m \leq n$ and A has rank m. A basis is a set of m linearly independent columns of A. Given a basis B, assume that it is an ordered set and let $B(j)$ be the j-th element of the set. A variable x_k is called basic if $k \in B$. A feasible solution x is called basic if $x_k = 0$ for $k \notin B$. We write A_k for the k-th column of A. The set $N := \{1, \ldots, n\} \setminus B$ is called the set of nonbasic variables. We denote by A_B the square invertible submatrix of A corresponding to columns in B, and A_N the remaining submatrix. The term "basis" may refer to B or A_B, depending on context. It is well-known that an optimal solution can always be chosen as $x = A_B^{-1}b$ for some basis B such that $x \geq 0$. The simplex method can be described compactly as follows; see, e.g., [211] for a more detailed treatment.

- Start with any basic feasible solution x. Let B be the current basis, N the nonbasic variables, $x = A_B^{-1}b$ the current solution.
- Repeat the following steps:
  1. For each nonbasic variable $k \in N$ compute its reduced cost $\bar{c}_k$. The reduced cost vector is defined as $\bar{c}_N^\top = c_N^\top - c_B^\top A_B^{-1} A_N$. If $\bar{c}_N \geq 0$ the basis is optimal and the algorithm terminates returning the current solution x. Otherwise, choose $k \in N$ such that $\bar{c}_k < 0$. This step is called *pricing*.
  2. Compute $u = A_B^{-1}A_k$. If $u \leq 0$, the optimal cost is unbounded from below and the algorithm terminates.
  3. If at least one component of u is positive, compute

$$r^* = \min_{j=1,\ldots,m:u_j>0} \frac{x_{B(j)}}{u_j}. \tag{48}$$

     This step is called *ratio test*.
  4. Let ℓ be such that $r^* = \frac{x_{B(\ell)}}{u_\ell}$. Form a new basis replacing $B(\ell)$ with $B(k)$. Update the feasible solution according to $x \leftarrow x - r^*u$. This step is called a *pivot*.

To perform the pricing step, we compute an LU factorization of A_B; this requires time $O(d_c^{0.7}m^{1.9} + m^{2+o(1)})$ using fast sparse matrix multiplication techniques [212], where d_c is

the maximum number of nonzeros per column of A. (In practice, the traditional $O(m^3)$ Gaussian elimination is used instead, but the factorization is not computed from scratch at every iteration.) Then, we can compute the vector $c_B^\top A_B^{-1}$ and finally perform the $O(n)$ calculations $c_k^\top - c_B^\top A_B^{-1} A_k$ for all $k \in N$; this requires an additional $O(d_c n)$ time, bringing the total time to $O(d_c^{0.7} m^{1.9} + m^{2+o(1)} + d_c n)$. To perform the ratio test, we need the vector $u = A_B^{-1} A_k$, which takes time $O(m^2)$ assuming the LU factorization of A_B is available from pricing. Finally, since the calculations are performed with finite precision, we use an optimality tolerance ϵ and the optimality criterion becomes $\bar{c}_N \geq -\epsilon$.

Quantum subroutines. Similar to other papers in the quantum optimization literature, we use a classical algorithm (the simplex method) and accelerate the subroutines executed at each iteration of the simplex. Specifically, we aim to obtain a quantum advantage for the pricing step. However, our asymptotic speedup does not depend on the availability of qRAM or of the data in "quantum form". The key insight to obtain an asymptotic speedup even with classical input and output is to interpret the simplex method as a collection of subroutines that output only binary or integer scalars, avoiding the cost of extracting real vectors from the quantum computer. Indeed, the simplex method does not require explicit knowledge of the full solution vector $A_B^{-1} A_k$ associated with a basis, or of the full simplex tableau $A_B^{-1} A_N$, provided that we are able to:

- Identify if the current basis is optimal or unbounded;
- Identify a pivot, i.e., the index of a column with negative reduced cost that enters the basis, and the index of a column leaving the basis.

While subroutines to perform these tasks require access to $A_B^{-1} b$ and/or $A_B^{-1} A_N$, we will show that we can get an asymptotic speedup by never computing a classical description of A_B^{-1}, $A_B^{-1} b$ or $A_B^{-1} A_N$. This is because extracting vectors of real numbers from a quantum computer is much more expensive than obtaining integer or binary outputs, as these can be encoded directly as basis states and read from a single measurement with high probability if the amplitudes are set correctly.

Throughout this section we assume that the LP data is properly normalized. The normalization can be carried out as a classical preprocessing step, and the running time of this step is negligible compared to the remaining subroutines.

Our first objective is to implement a quantum oracle that determines if a column has negative reduced cost, so that we can apply Grover's search to this oracle. To reach this goal we rely on the Quantum Linear System Algorithm (QLSA) [29]. Using that algorithm, with straightforward data preparation we can construct an oracle that, given a column index k in a certain register, outputs $\left|A_B^{-1} A_k\right\rangle$ in another register. We still need to get around three obstacles: (i) the output of the QLSA is a renormalization of the solution, rather than the (unscaled) vector $A_B^{-1} A_k$; (ii) we want to compute $c_k - c_B^\top A_B^{-1} A_k$, while so far we only have access to $\left|A_B^{-1} A_k\right\rangle$; (iii) the output has to be a binary yes/no condition (i.e., not just an amplitude) so that Grover search can be applied to it, and we are not allowed to perform measurements. We overcome the first two obstacles by: extending and properly scaling the linear system so that c_k is suitably encoded in the QLSA output; and using the inverse of the unitary that maps $\left|0^{\lceil \log m+1 \rceil}\right\rangle$ to $|(-c_B, 1)\rangle$ to encode $c_k - c_B^\top A_B^{-1} A_k$ in the amplitude of one of the basis states. To determine the sign of such amplitude, we rely on interference to

create a basis state with amplitude α such that $|\alpha| \geq \frac{1}{2}$ if and only if $c_k - c_B^\top A_B^{-1} A_k \geq 0$. At this point, we can apply amplitude estimation [169] to determine the magnitude of α up to precision ϵ. This requires $O(1/\epsilon)$ iterations of the amplitude estimation algorithm. We therefore obtain a unitary operation that overcomes the three obstacles. A similar scheme can be used to determine if the basis is optimal, i.e., no column with negative reduced cost exists.

Runtime scaling. Runtimes quoted below depend on the following parameters:

- d_c: maximum number of nonzero entries in any column of A.
- d_r: maximum number of nonzero entries in any row of A_B.
- $d := \max\{d_c, d_r\}$: sparsity of A_B.
- κ: ratio of largest to smallest nonzero singular value of A_B.
- ϵ: error tolerance. The pricing subroutine returns a column with the reduced cost $\leq \epsilon$. The optimality is declared when all reduced costs are at least $-\epsilon$.

Our first result is

Theorem 2. *There exist quantum subroutines to identify if a basis is optimal, or determine a column with negative reduced cost, with running time $\tilde{O}(\frac{1}{\epsilon}\sqrt{n}(\kappa d_c n + \kappa^2 d^2 m))$. The runtime can be reduced to $\tilde{O}(\frac{1}{\epsilon}\kappa^{1.5} d\sqrt{d_c} n\sqrt{m})$ if the ratio n/m is larger than $2\frac{\kappa d^2}{d_c}$.*

Recall that the classical cost of the pricing step is $O(d_c^{0.7} m^{1.9} + m^{2+o(1)} + d_c n)$. Thus the quantum subroutine offers a polynomial speedup. For example, if $m \sim n$ and all other parameters have a poly-logarithmic dependence on n, the classical runtime is $O(n^{2+o(1)})$ while the quantum runtime is $\tilde{O}(n^{3/2})$. The regime with large n/m is interesting because it includes many natural LP formulations; e.g., the LP relaxations of cutting stock problems, vehicle routing problems, or any other formulation that is generally solved by column generation [213]. Note that the running time of the quantum subroutines depends explicitly on the condition number of the basis and the precision of reduced costs ϵ is fixed, while classically κ is not explicit when using Gaussian elimination, but the ϵ obtained would depend on it (because the basis inverse could be inaccurate).

If pricing is performed via our quantum subroutine, we obtain the index of a column that has negative reduced cost with arbitrarily high probability. To determine which column should leave the basis, we have to perform the *ratio test*. Using techniques similar to those used for the pricing step, we can identify the column that leaves the basis in time $\tilde{O}(\frac{t}{\delta}\kappa^2 d^2 m^{1.5})$, where δ and t are suitable precision parameters of this step, see [30] for details. Classically, the ratio test requires time $O(m^2)$ in the worst case, because the basis inverse could be dense even if the basis is sparse (although it is unlikely in practice). We summarize this result below.

Theorem 3. *There exists a quantum subroutine to perform the ratio test in time $\tilde{O}(\frac{t}{\delta}\kappa^2 d^2 m^{1.5})$. There also exists a quantum subroutine to identify if a nonbasic column proves unboundedness of the LP in time $\tilde{O}(\frac{1}{\delta}\kappa^2 d^2 m^{1.5})$.*

It is known that for most practical LPs the maximum number of nonzeroes in a column is essentially constant; for example, on the entire benchmark set MIPLIB2010, less than 1% of the columns have more than 200 nonzeroes (and less than 5% have more than 50 nonzeroes). Similarly, the number of nonzeroes per row of the basis is small: on MIPLIB2010, looking at the optimal bases of the LP relaxations, less than 0.01% of the rows have more than 50 nonzeroes. As m, n increase, so typically does the sparsity. For example, the largest problem available in the benchmark set MIPLIB2017 has $m \approx 7.1 \times 10^6, n \approx 3.9 \times 10^7$, and 99.999% of the columns have less than 30 nonzero elements; for the second largest problem, which has $m \approx 2.0 \times 10^7, n \approx 2.1 \times 10^7$, 99.998% of the columns have this property. Hence, we expect many bases to be extremely sparse, and it is interesting to look at the scaling of the running time under the assumption that the sparsity parameters are constant, or at most polylogarithmic in m and n. In this case, the running time of the oracle for the reduced costs in the gate model without qRAM is $\tilde{O}(\frac{1}{\epsilon}(\kappa n + \kappa^2 m))$, giving a total running time for pricing of $\tilde{O}(\frac{1}{\epsilon}\sqrt{n}(\kappa n + \kappa^2 m))$. Hence, for a well-conditioned basis and under the assumption (often verified in practice) that $d = O(\log mn)$, we obtain running time $\tilde{O}(\frac{1}{\epsilon}\sqrt{n}(n+m))$ for the quantum pricing subroutine, which can be reduced to $\tilde{O}(\frac{1}{\epsilon}n\sqrt{m})$ if the ratio n/m is large; and running time $\tilde{O}(\frac{t}{\delta}m^{1.5})$ for the quantum ratio test subroutine.

4 Quantum simulation algorithms

Here we report several results pertaining to simulation of quantum many-body systems and quantum Hamiltonian complexity theory. We show how to overcome limitations of near-term variational algorithms by employing embedding algorithms and adaptive circuits. On the complexity-theoretic side, we study the problem of approximating the partition function of a local Hamiltonian system and relate this problem to a classical combinatorial problem — approximating the Kronecker coefficients of the symmetric group.

Section 4.1, based on Ref. [40], introduces a scheme we call *classically forged entanglement*, which represents a $2N$-qubit wavefunction as multiple N-qubit states embedded in a classical computation. Beyond the reduction in requisite qubit number, offloading entanglement synthesis to classical processing permits the constituent N-qubit quantum circuits to be shallower, relaxing requirements on gate error and connectivity, at the cost of an increased number of circuit executions.

Section 4.2, based on Ref. [51], focuses on the ground state preparation problem for quantum many-body systems with non-abelian topological quantum order. Such states are notoriously hard to prepare by regular quantum circuits composed of two-qubit unitary gates. In the case of 2D topologically ordered systems considered here, the circuit depth required for the state preparation is known to scale at least linearly with the lattice size [155]. A natural alternative is to consider *adaptive* quantum circuits that include intermediate measurements and classical feedback such that each operation may depend on the outcomes of earlier measurements. It is well-known that certain families of deep unitary quantum circuits can be implemented by a constant-depth adaptive circuit using ancillary qubits. Notable examples include Clifford circuits [214], CNOT+T circuits [215, 216], and controlled gates with multiple control qubits [217]. This motivates the question of whether adaptiveness can facilitate simulation of exotic phases of matter with a non-abelian particle statistics. Our

work [51] answers this question in affirmative by presenting constant-depth adaptive circuits for the ground state preparation and manipulation of anyons in a large class of topologically ordered 2D systems known as quantum double models. Our results apply to the quantum double of any finite *solvable* group G. Constant-depth adaptive circuits are well suited for implementation on a noisy hardware since it may be possible to execute the entire circuit within the qubit coherence time. Thus our results enable an experimental study of systems with non-abelian anyons, see [53] for a recent experimental demonstration. We also show that adaptiveness is essential for our circuit construction. Namely, braiding of anyons cannot be realized by non-adaptive constant-depth local circuits for any non-abelian group G. This is in a sharp contrast with abelian anyons which can be created and moved over an arbitrary distance by a depth-1 circuit composed of generalized Pauli gates.

Predicting properties of a quantum many-body system that emerge from its microscopic description in terms of constituent particles and interactions among them is a fundamental problem in physics. Many properties of a system in thermal equilibrium are determined by the partition function

$$\mathcal{Z} = \mathrm{Tr}[e^{-\beta H}],$$

where β is the inverse temperature and H is the Hamiltonian describing the system. The partition function appears as a normalization factor in the Gibbs state $\rho = e^{-\beta H}/\mathcal{Z}$ and determines the Helmholtz free energy

$$F = -(1/\beta)\log \mathcal{Z}. \tag{49}$$

The ability to calculate the free energy and its derivatives with respect to the temperature and Hamiltonian parameters such as external fields is instrumental for mapping out the phase diagram of the system and predicting physical properties of each phase. Accordingly, the problem of estimating the free energy has been extensively studied, both in the physics and computer science communities. Section 4.3, based on Ref. [54], investigates this problem for k-local Hamiltonians [218] that describe a system of n qubits with interactions among subsets of at most k qubits. Despite its fundamental significance, little is known about the computational complexity of estimating the free energy of a local Hamiltonian to a given additive error (or equivalently the partition function to a given relative error). We establish polynomial-time equivalence between the problem of approximating the free energy of local Hamiltonians and several other natural tasks ubiquitous in condensed-matter physics and quantum computing, such as the problem of approximating the number of input states accepted by a polynomial-size quantum circuit. These results suggest that the simulation of quantum many-body systems in thermal equilibrium may precisely capture the complexity of a broad family of computational problems that have yet to be defined or characterized in terms of known complexity classes. Finally, we summarize state-of-the-art classical and quantum algorithms for approximating the free energy and show how to improve their runtime and memory footprint.

Section 4.4, based on Ref. [55], describes a surprising connection between quantum partition functions and a purely classical combinatorial problem — computing Kronecker coefficients of the symmetric group S_n. Recall, that Kronecker coefficients $g_{\mu,\nu,\lambda}$ associated with a triple of irreducible representations of S_n is defined as the multiplicity of the trivial representation of S_n in the tensor product $\mu \otimes \nu \otimes \lambda$. By definition, $g_{\mu,\nu,\lambda}$ is a non-negative integer. A

longstanding open problem in algebraic combinatorics stated as Problem 10 on Stanley's list [219] is to find a combinatorial formula for $g_{\mu,\nu,\lambda}$. Such a formula would express $g_{\mu,\nu,\lambda}$ as the cardinality of some natural set of combinatorial objects that admits an efficient membership test. In this way we are led to the complexity-theoretic question [220, 221, 222, 223, 224]: does Kronecker coefficient $g_{\mu\nu\lambda}$ admit a #P formula, i.e., does it count the number of accepting witnesses of an NP verifier? While this question remains largely open, our work shows that Kronecker coefficient $g_{\mu,\nu,\lambda}$ can be represented as the number of accepting witnesses for a quantum (QMA) verifier. Accordingly, the problem of approximating $g_{\mu,\nu,\lambda}$ falls into the same complexity class as the one of approximating quantum partition function of a local Hamiltonian. We also give an efficient quantum algorithm that computes Kronecker coefficient $g_{\mu,\nu,\lambda}$ exactly, assuming that at least one of the representations μ, ν, λ has dimension at most $poly(n)$. We are not aware of an efficient classical algorithm for this task.

Section 4.5, based on Ref. [225], addresses the problem of simulating open quantum systems whose dynamics is governed by the Lindblad master equation. It shows how to generalize the Szegedy walk unitary [170] from classical stochastic matrices to quantum channels (trace preserving completely positive maps) satisfying a certain detailed balance condition.

4.1 Entanglement forging

Here we sketch main ideas behind the entanglement forging method, see Ref. [40] for details. We begin with Schmidt decomposition, a standard application of singular value decomposition that allows one to write any state $|\psi\rangle$ of a bipartite $N+N$ qubit system as

$$|\psi\rangle = (U \otimes V) \sum_{n=1}^{2^N} \lambda_n |b_n\rangle \otimes |b_n\rangle \,. \tag{50}$$

Here $|b_n\rangle$ are the N-qubit bitstring states, also known as computational basis states, U, V are unitary operators respective to the two subsystems, and the Schmidt coefficients λ_n may be taken to be non-negative. The flatter the distribution of Schmidt coefficients, the stronger the entanglement; a uniform distribution $\lambda_n = 1/\sqrt{2^N}$ indicates the two halves of the system are maximally entangled, while only one nonzero coefficient corresponds to no entanglement.

Mixing classical and quantum information, entanglement-forging is more naturally expressed using density operators rather than wavefunctions, even for pure states such as $|\psi\rangle$. As shown in Ref. [40], one can write the density operator as

$$|\psi\rangle\langle\psi| = (U \otimes V) \sum_{n=1}^{2^N} \Big(\lambda_n^2 |b_n\rangle\langle b_n|^{\otimes 2} + \sum_{m=1}^{n-1} \lambda_n \lambda_m \sum_{p \in \mathbb{Z}_4} (-1)^p |\phi^p_{b_n b_m}\rangle\langle \phi^p_{b_n b_m}|^{\otimes 2} \Big) (U^\dagger \otimes V^\dagger), \tag{51}$$

where we have used the definition $\left|\phi^p_{xy}\right\rangle = \big(|x\rangle + i^p |y\rangle \big)/\sqrt{2}$ with $p \in \{0,1,2,3\} = \mathbb{Z}_4$. For example, in the minimal case of two qubits (Fig. 5B), $|\phi^p_{01}\rangle$ correspond to four equatorial points on Bloch sphere, rewriting a quantum superposition of product states in terms of classical products of superposition states. Eq. (51) generalizes methods proposed in [226, 227, 31], and is connected to tensor network representations of quantum circuits [228, 38],

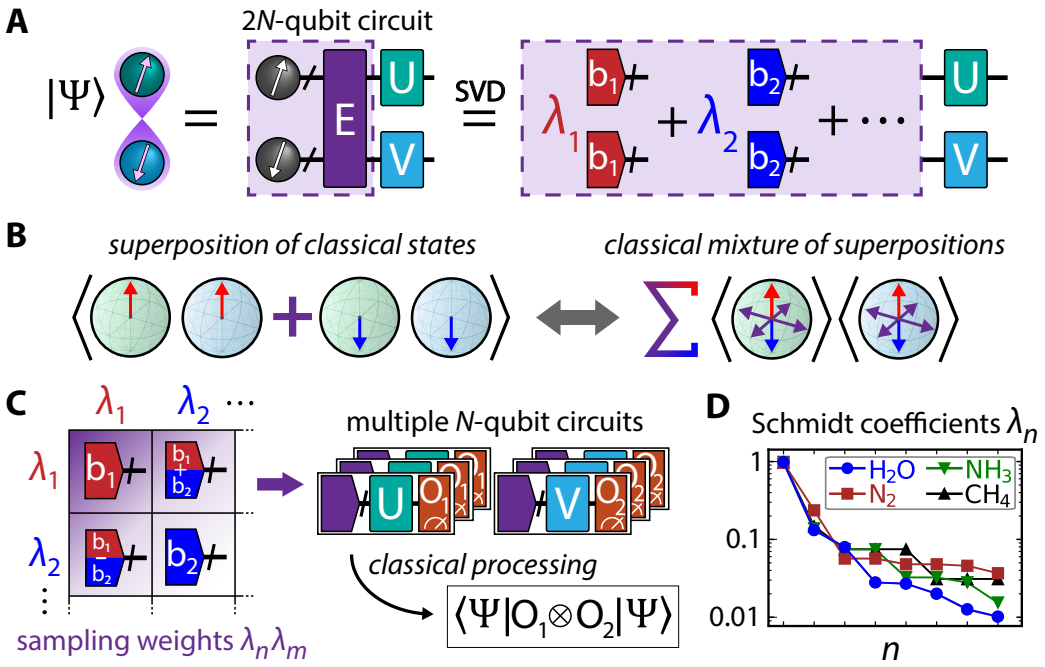


Figure 5: **Schematic overview of the entanglement forging protocol.** **A**, A state $|\Psi\rangle$ of a bipartite quantum system, here labeled with arrows alluding to spin polarization, can be defined by gates E, U, and V, where E outputs a combination of bitstring states $|b_n\rangle\,|b_n\rangle$. **B**, A two-qubit entangled state can be rewritten using one-qubit superposition states. Changing labels $0, 1 \to b_n, b_m$ gives a transformation acting on components of the $2N$-qubit state. **C**, $|\Psi\rangle$ can be reconstructed from N-qubit circuits initialized as bitstrings and pairwise superpositions thereof. Circuits associated with small $\lambda_n\lambda_m$ can be estimated adequately from few samples. **D**, Rapid (slow) decay of the leading Schmidt coefficients in the decomposition of a molecular ground state signals weak (strong) entanglement between spin-up and spin-down particles.

variational simulation of open quantum systems [229], and the encoding of open-shell singlet and triplet states [230]. The expectation of a $2N$-qubit operator $O = O_1 \otimes O_2$ is now

$$\begin{aligned}\langle O\rangle = \sum_{n=1}^{2^N} \Big(&\lambda_n^2 \langle b_n|\tilde{O}_1|b_b\rangle \cdot \langle b_n|\tilde{O}_2|b_n\rangle \\ &+ \sum_{m=1}^{n-1} \lambda_n\lambda_m \sum_{p\in\mathbb{Z}_4} (-1)^p \langle \phi^p_{b_n b_m}|\tilde{O}_1|\phi^p_{b_n b_m}\rangle\langle \phi^p_{b_n b_m}|\tilde{O}_2|\phi^p_{b_n b_m}\rangle\Big),\end{aligned} \tag{52}$$

where $\tilde{O}_1 = U^\dagger O_1 U$ and $\tilde{O}_2 = V^\dagger O_2 V$, and each constituent requires only N qubits to evaluate.

The resulting summation for $\langle O\rangle$ is not obviously scalable, involving as many as $2^{N+1}(2^{N+1}-1)$ distinct N-qubit quantum circuits. Nonetheless, if one restricts to simulations of sufficiently weak entanglement, $\langle O\rangle$ can be efficiently estimated by sampling each circuit in proportion to the associated coefficients $\lambda_n\lambda_m$ in (52), with a total number of samples for

target precision ϵ scaling as

$$S \sim \left(\frac{1}{\epsilon}\sum_{n,m}|\lambda_n\lambda_m|\right)^2 = \frac{\|\vec{\lambda}\|_1^4}{\epsilon^2}\ ,\ \|\vec{\lambda}\|_1 = \sum_n |\lambda_n|\ . \tag{53}$$

Executing a quantum circuit once provides one sample of the corresponding expectation value, such that total runtime scales linearly with S. Since the one-norm decreases toward 1 in the limit of weak entanglement, the overhead cost of entanglement forging is smaller for simulations of states divisible into weakly-entangled halves, such as the spin-up and spin-down components of certain molecular ground states [40] and scales efficiently when the one-norm is at most polynomial in the problem size. For example, in some statically correlated ground states, S can be independent of the number of basis orbitals. Outside of the domain of scalability, entanglement forging still enables useful heuristic simulations beyond the standard capacity of given quantum hardware, which may be realized with precision by truncating the list of bitstring states retained in the Schmidt decomposition.

Alternatively, this overhead may be reduced to a constant factor independent of qubit number via a complementary scheme [40] simulating quantum correlations between subsystems using those within a subsystem, rather than using classical correlations as above. This method can be seen as an application of forging in the Heisenberg picture, reinterpreting an observable acting on $N+N$ qubits as a classical mixture of operators describing the forward and backward time evolution of N-qubits, at a cost of deeper circuits. Provided certain sampling assumptions, this method is not limited to weakly entangled states, so it may be applicable to a wider range of systems, as recently demonstrated in [231].

In Ref. [40] we use entanglement forging for a variational simulation of the water molecule in the minimal STO-6G basis. We freeze the core oxygen $1s$ and the out-of-plane oxygen $2p$ orbitals, leaving an active space of 10 spin-orbitals. The Jordan Wigner mapping encodes either the spin-up or spin-down orbitals onto each five-qubit quantum circuit, and we simplify the problem structure by asserting $V = U$ per the known symmetry between spin-polarizations in the closed-shell singlet ground state.

4.2 Adaptive circuits for manipulating non-abelian anyons

Kitaev's quantum double model [232] is a paradigmatic example of an exactly solvable model exhibiting topological order. For a finite group G, the construction provides a spin-lattice model whose low-energy physics is described by the non-chiral topological quantum field theory associated with Drinfeld's quantum double $D(G)$. Its most well-known manifestation – the toric [232] or surface code [233] based on the abelian group $G = \mathbb{Z}_2$ – is currently among primary contenders when it comes to experimental realizations, see e.g., [234]. The preeminence of this abelian model can be attributed, in part, to the fact that it falls into a well-studied family of quantum error-correcting codes: It is a CSS-stabilizer code with local generators. As such, it permits the application of a plethora well-studied techniques for fault-tolerant quantum computing, including, in particular, efficient syndrome extraction and decoding as well as magic state distillation. Futhermore, its constituent physical degrees of freedom are qubits since $|\mathbb{Z}_2| = 2$, making this model especially amenable to experimental efforts.

Despite the promise of the surface codes for topological quantum computation, models with non-abelian anyons may still offer distinct advantages. In particular, unlike with abelian anyons, universal computation can be achieved by purely topological means when the associated braid group representation is sufficiently rich. The required operations are

(i) the creation/initialization of a ground (vacuum) state having no excitations

(ii) the braiding or exchange of two particles

(iii) the fusion of two particles: this amounts to bringing two anyons to the same site and performing a measurement of the joint topological charge

(iv) the creation of specific particle-antiparticle pairs.

Mochon [235] has shown that for anyons described by $D(G)$ with a solvable but not nilpotent group G, these operations suffice to realize universal computation[6]. This applies, in particular, to the group S_3. The power of $D(S_3)$-anyons to realize universal computation has prompted the construction of explicit protocols implementing these operations, including experimental proposals for their realization with atoms trapped in an optical lattice, see [236, 237]. Unlike the corresponding operations for the surface code, however, the protocols proposed in [236, 237] require quantum circuits whose depth scales with the system size, respectively the spatial separation between anyons. It is natural to ask if this increased complexity is in fact necessary when dealing with non-abelian anyons.

To see how non-abelian and abelian anyons may differ in their (circuit) complexity of their associated operations, consider the problem (i) of initial state preparation. There is no significant difference here when considering unitary circuits: A circuit consisting of nearest-neighbor two-qubit unitary gates (a unitary local circuit) preparing the ground state of the surface code from a product state has been proposed in [238]. To prepare a ground state on an $L \times L$-lattice, the circuit has depth of order $O(L^2)$. An analogous result applies to the preparation of a ground state of any quantum double model, see [237, 239]. An $O(L)$-depth circuit for the surface code was given in [240], and linear depth circuits for the surface code and certain Levin-Wen models were also obtained in [241]. In [155], it is shown that a circuit depth of order $\Omega(L)$ is necessary for any local unitary circuit. This lower bound applies to any topologically ordered system, irrespective of whether or not the corresponding anyons are abelian. With non-local two-qudit gates, the circuit depth can be lowered to $O(\log L)$ [242] for any quantum double model; this matches the lower bound of [243] for $G = \mathbb{Z}_2$.

A more efficient way of preparing states with $D(\mathbb{Z}_2)$-topological order involves the use of measurements. Indeed, as with any stabilizer code, a code state (i.e., a ground state in the case of the surface code) is obtained from an arbitrary initial state by

(a) measuring a complete set of commuting stabilizer generators, obtaining a syndrome s

(b) applying an associated unitary correction operation $C(s)$ to the post-measurement state.

[6]Mochon additionally assumes that a supply of calibrated electric/magnetic charge ancillas is available. This is implied by operation (iv), where in contrast to [235], we assume that the specific pair created is not random.

We note that such an approach could in principle also be taken for any quantum double model. In particular, the required measurements in step (a) can also be chosen to be local as in any commuting local projector code. The distinguishing feature of the surface code is the simplicity of step (b): The required correction $C(s)$ is a Pauli operator which can be determined from s by an efficient classical decoding algorithm. In particular, application of $C(s)$ is easily accomplished by a depth-1 unitary circuit consisting of single-qubit Pauli gates only. In terms of the excitations (anyons) of the surface code, this process can be summarized as follows. The syndrome s reveals the locations of the excitations (violations of the stabilizer constraints). The applied correction $C(s)$ (determined, e.g., by minimal matching as in [238]) is a product of ribbon-operators where each ribbon connects two excitations. In the surface code, such ribbon operators are tensor products of Pauli-X, Pauli-Y or Pauli-Z-operators along the ribbon, see [232] or [238]. The simplicity of this preparation scheme is a consequence of the simple form of these ribbon operators for the case $G = \mathbb{Z}_2$.

For a general finite group G, each anyon is labeled by an irreducible representation ρ of the Drinfeld double $D(G)$ of G. For each (open) ribbon ξ on the lattice, there are associated ribbon operators $\{F_\xi^{\rho;\alpha}\}_\alpha$ which create a corresponding particle/antiparticle pair at the two endpoints of ξ. Here α controls degrees of freedom localized at these endpoints. We note that a ribbon operator $F_\xi^{\rho;\alpha}$ can also be understood as realizing a process where a pair is created locally at one endpoint, and one of the anyons is moved to the other endpoint of the ribbon ξ. In particular, for non-abelian models, such operations can have a non-trivial logical action on encoded quantum information in the presence of other (preexisting) anyons because of non-trivial braiding relations. The idea that – since the quantum double model is an error-correcting code – such logical operations require a circuit depth which scales with the length of the ribbon is folklore, see e.g., [244, Section 8] for a related argument. Our work gives a rigorous proof of this fact for non-abelian quantum double models. It gives the following statement, where we use the term *extensive* to refer to a scaling linear in the system size (or, more precisely, the code distance as given, e.g., by the separation of holes in the surface).

Corollary 1. *For any non-abelian group G, there are ribbon operators $F_\xi^{\rho;\alpha}$ whose implementation by a local unitary circuit requires an extensive circuit depth.*

Corollary 1 is a no-go result ruling out the possibility of applying ribbon operators using constant-depth circuits in non-abelian anyon models. This is a novel operational separation between non-abelian and abelian topological order.

It is tempting to think that the conclusion of Corollary 1 extends to the problem (i), preventing initialization by constant-depth quantum circuits because of the difficulty of implementing the ribbon operators involved in the correction step (b). In fact, step (b) involves the additional challenges of classical decoding the corresponding code (i.e., finding the right product of ribbon operators to apply), a problem which is non-trivial [245]. Furthermore, in a non-abelian model, a single application of a ribbon operator is typically not sufficient to remove a pair of excitations since anyon pairs do not need to fuse to the vacuum. Nevertheless, we find that for certain non-abelian quantum double models (including the case $G = S_3$), efficient initialization is possible by using an approach different from (a)–(b). In fact, all operations (i)–(iv) above can be realized by what we call *constant-depth local adaptive circuits*.

This notion captures the operations involved in the procedures used for the surface code: We allow a constant number of the alternating layers of

(A) constant-depth quantum circuits that may use auxiliary qubits and consist of local unitaries and single-qubit measurements and

(B) efficient, possibly non-local classical computations (based on measurement results).

Adaptivity refers to the fact that quantum operations may be classically controlled by measurement outcomes and computational results obtained in previous layers. In other words, all involved quantum operations are local and realized by constant-depth circuits; this is supplemented by efficient non-local classical processing.

In more detail, our result applies to any solvable group G. Recall that a group G is solvable if it can be mapped to the trivial group by iteratively factoring out normal abelian subgroups. We show the following:

Theorem 4. *(informal) Suppose G is a solvable group. Then there is a constant-depth local adaptive circuit for each of the following tasks:*

1. *Creation of a ground state from a product initial state.*
2. *For any anyon label ρ and local action specified by α, application of the ribbon operators $F_\xi^{\rho;\alpha}$ associated with an open ribbon ξ of possibly extensive length. Here the circuit has support (i.e., acts non-trivially) only on qudits along the ribbon ξ.*
3. *Execution of the topological charge measurement of any region encircled by a closed ribbon σ (possibly of extensive length). This is given by a POVM $\{K_\sigma^\rho\}_\rho$ whose outcomes are anyon labels ρ, and whose POVM elements K_σ^ρ have support on the ribbon σ. The circuit realizing this measurement has support only on qudits along σ.*

The operations considered in Theorem 4 can be used to realize braiding/movement of anyons by (repeated) pair creation and topological charge measurement.

Prior work. The power of measurements and adaptive operations for reducing the depth of quantum circuits has been recognized in the seminal work by Høyer and Spalek [246] – some of our constructions are motivated by their fan-out gate. This work has lead to a discovery of constant-depth adaptive circuits realizing any unitary in the Clifford group [247, 248], multiple control Toffoli gate and integer arithmetic circuits [246, 249], and Quantum Fourier Transform [246] (the latter can only be realized approximately). These are important subroutines employed by numerous quantum algorithms. Moreover, it is known that the entire quantum part of Shor's factoring algorithm [250] can be parallelized to a constant depth using adaptive circuits [251]. Measurements and adaptive operations also play a central role in synthesizing quantum circuits over fault-tolerant gate sets using gate teleportation [247] and repeat-until-success techniques [252]. In a recent breakthrough work Liu and Gheorghiu [253] established an efficiently verifiable quantum advantage for certain classically hard interactive tasks that can be solved by constant depth adaptive quantum circuits. The apparent power of low-depth adaptive circuits has led Jozsa to conjecture [248] that, in fact, any

polynomial depth quantum computation can be efficiently simulated by logarithmic depth adaptive quantum circuits [248]. However, a more recent work [254] cast doubt on Jozsa's conjecture by demonstrating that the Welded Tree Problem [255] is provably hard for the logarithmic depth adaptive circuits (as measured by query complexity) even though this problem admits an efficient quantum algorithm based on quantum walks. Likewise, it is not known whether the depth of quantum algorithms for simulating unitary time evolution of many-body Hamiltonians can be significantly reduced using measurements and adaptive operations.

In a recent work Piroli et al [256] examined constant-depth quantum circuits assisted by LOCC (local operations and classical communication) and equivalence classes of many-body quantum states convertible to each other by such circuits. It was demonstrated that certain highly entangled states exhibiting topological quantum order become trivial in the LOCC-assisted classification framework. We note however that constant depth local adaptive circuits considered in the present work are strictly weaker than LOCC-assisted circuits of [256] since we only allow $O(1)$ rounds of mid-circuits measurements whereas [256] allows a linear number of rounds. Closer to the topic of this paper, existing non-unitary protocols for preparing states with non-abelian topological order include those of [237, 239] for creating quantum double states associated with S_3, as well as the work [241] which provides procedures for preparing ground states of the Levin-Wen model [257]. In contrast to these procedures, which require extensive circuit depth, Verresen et al. [258] reported adaptive constant-depth circuits preparing the ground state of the quantum double $D(S_3)$ and $D(D_4)$, where D_4 is the dihedral group. Subsequently, it was observed in [259] that this generalizes to any solvable group. Here we give explicit adaptive constant-depth circuits for ground state preparation for arbitrary solvable groups and additionally discuss the creation and manipulation of excited states. We note that constant-depth circuits also figure prominently in [260], where braiding with non-abelian anyons in the Levin-Wen model is achieved by constant-depth circuits by certain dynamic lattice deformations. Here we follow a different approach and ask for implementations of "standard" braiding operations etc. without changing the underlying lattice.

Open questions. Perhaps the most intriguing one is whether the operations realized here for a solvable group G may also be realized with local operations in constant adaptive depth in the case of a non-solvable group such as S_5. It is conceivable that there are fundamental complexity-theoretic obstructions to this similar to those found in [261]) related to Barrington's theorem [262]. Note however that the smallest non-solvable group has size $|A_5| = 60$. Thus realizing the quantum double Hamiltonian based on A_5 would involve few-body interactions between 60-dimensional qudits posing a challenge for experimental demonstrations. Thus our work covers all cases of the quantum double model that can be plausibly realized in the experiment in the near term. More generally, one may ask for more general classes of topologically ordered systems where the considered operations (preparation, anyon creation, braiding, and topological charge measurements) can be realized in adaptive constant depth. For the Levin-Wen model, this amounts to identifying relevant properties of the underlying tensor category. A preliminary work shows that the ground state of the double semion model [257] can be prepared in constant adaptive depth by measuring syndromes of suitable

augmented stabilizer generators composed from the elementary plaquette and vertex stabilizers. Finally, it remains to be seen to what extent these procedures can be made robust to noise: Without suitable error correction mechanisms, the non-local operations realized here will likely be highly susceptible to errors. Because of their low complexity, they may nevertheless be useful, e.g., for proof-of-principle demonstrations in the near term.

4.3 Complexity of quantum partition functions

We consider k-local Hamiltonians

$$H = \sum_{S \subseteq \{1,\ldots,n\}} H_S \,,$$

where each term H_S acts non-trivially only on the subset of qubits S, and $H_S = 0$ unless $|S| \leq k$ where k is a constant independent of n. Below we assume that $\|H_S\| \leq \mathrm{poly}(n)$ for all S. We focus on approximating the partition function $\mathcal{Z} = \mathrm{Tr}(e^{-\beta H})$ for a given inverse temperature $\beta \leq \mathrm{poly}(n)$ within a relative error $\delta \geq \mathrm{poly}(1/n)$. Let us refer to this task as the Quantum Partition Function (QPF) problem.

It is crucial that here we consider *relative error* estimation. Previous works have shown that the less challenging problem of additively approximating the (normalized) partition function [263, 264, 265] admits an efficient quantum algorithm and, for Hamiltonians with locality $k = O(\log(n))$, is complete for the complexity class DQC_1 of problems that can be solved in polynomial time with only "one clean qubit" [266]. However, it is unclear whether such less stringent approximation is sufficient to characterize thermal equilibrium properties of a quantum system.

In contrast to the ground energy, the partition function $\mathcal{Z}$ depends on all eigenvalues of the Hamiltonian as well as their *degeneracy*. Thus one can view the QPF problem as a quantum analogue of approximate counting — a fundamental task in classical computational complexity. In particular, it is known that the problem of approximating the partition function of a classical local Hamiltonian with a small relative error is contained in the complexity class $\mathsf{BPP}^{\mathsf{NP}}$, essentially due to a fundamental result of Stockmeyer [267]. We can interpret this result as stating that the classical version of the QPF problem is not much harder than NP, which is surprising because computing the partition function exactly is #P-hard. However, if we allow β to be complex, the task of approximating $\mathcal{Z}$ with a constant relative error is #P-hard [268], even for classical Hamiltonians. The discrepancy between the problem complexity for real and complex β's suggests that non-negativity of the Gibbs state should play an important role in the complexity analysis.

The fact that Stockmeyer's approximate counting method does not generalize to the quantum case has been observed previously [269]. This may explain why the complexity of the QPF problem – a basic and natural question from a physics perspective – remains largely open. A recent result by Cubitt et al. [270] introduced the notion of universal quantum Hamiltonians and showed (among other things) that the QPF problem for k-local Hamiltonian has the same complexity for any constant $k \geq 2$. The same universality result also applies to specialized models such as the 2-local anti-ferromagnetic Heisenberg model on a two-dimensional lattice of qubits and 1D qudit systems [270, 271], and extends even to fermionic Hamiltonians relevant for quantum chemistry and material science [270]. This

suggests that the QPF problem is complete for some complexity class that is "universal" in the sense that it can be defined only in terms of a suitable computational model and does not depend on details of the considered Hamiltonians, such as the locality parameter k (as long as $k \geq 2$) or whether the underlying system consists of qubits or fermions.

Given this state of affairs, and the apparently challenging nature of characterizing the complexity of the QPF problem in terms of known complexity classes, here we use a standard computer science dodge – we look for other computational problems which are equivalent to QPF under polynomial time reductions with the hope that these other problems may ultimately be easier to understand.

Main results. Our work shows that the QPF problem is equivalent (under polynomial-time reductions) to approximating the following quantities:

(1) The expected value of a local observable in the Gibbs state of a k-local Hamiltonian.

(2) The number of eigenvalues of a k-local Hamiltonian in a given energy interval, i.e., the density of states.

(3) The number of witness states accepted by a quantum circuit.

We require an approximation within a small additive error in case (1) and a small relative error in cases (2,3). Problems (1,2) are ubiquitous in condensed matter physics since expected values of local observables and the density of states provide important insights into properties of a quantum material. Problem (3) is defined in terms of quantum circuits rather than Hamiltonians. We consider a polynomial-size quantum circuit followed by a measurement of some designated output qubit. The circuit takes as input a witness state and possibly ancilla qubits initialized in $|0\rangle$. A witness state is accepted if the probability of the measurement outcome '1' is above a specified threshold. The problem is to approximate the number of linearly independent witness states accepted by the circuit. It can be viewed as a counting analogue of the QMA-complete circuit satisfiability problem [218] where the goal is to decide whether a quantum circuit as above accepts at least one witness state.

We also reproduce the result of Cubitt et al. [270] showing the equivalence between QPF problems for 2-local and k-local Hamiltonians with any constant k. Our proof is slightly more direct than the one of [270] as we do not use perturbation theory gadgets; our technique appears to share some features of later works which used Kitaev's circuit-to-Hamiltonian mapping [271, 272, 273]. Note however that reducing k-local partition functions to 2-local partition functions is simpler than showing that such Hamiltonians are universal in the sense considered in Refs. [270, 271, 272, 273].

To state our results let us define formal versions of the considered problems.

Problem 1 (Quantum Partition Function). *Given a k-local Hamiltonian H acting on n qubits, inverse temperature $\beta \leq poly(n)$, and a precision parameter $\delta \geq poly(1/n)$. Compute an estimate ξ such that*

$$(1-\delta)\mathrm{Tr}(e^{-\beta H}) \leq \xi \leq (1+\delta)\mathrm{Tr}(e^{-\beta H}).$$

Below we refer to this problem as k-QPF. Next, define the problem of Quantum Approximate Counting (QAC). Consider a *verifier circuit* of the following form. It takes as input an n-qubit state, adjoins $n_a \leq \text{poly}(n)$ ancilla qubits in the $|0\rangle$ state, and then applies a quantum circuit U of size $\text{poly}(n)$ followed by measurement of a single output qubit. Such a circuit implements an n-qubit two-outcome POVM $\{A, I-A\}$ where the operator A corresponds to measurement outcome '1' and satisfies $0 \leq A \leq I$. Formally we have

$$A = \left(I \otimes \langle 0^{n_a}|\right) U^\dagger |1\rangle\langle 1|_{\text{out}} U \left(I \otimes |0^{n_a}\rangle\right).$$

For any n-qubit input state ψ, the probability that it is accepted by the verifier circuit is given by $\langle\psi|A|\psi\rangle$. For any $\lambda \in [0,1]$ we write $\mathcal{L}_\lambda$ for the linear subspace spanned by all eigenstates of A with eigenvalues greater than or equal to λ, and Π_λ for the projector onto this subspace. The dimension of this subspace is denoted

$$N_\lambda \equiv \dim(\mathcal{L}_\lambda) = \text{Tr}(\Pi_\lambda). \tag{54}$$

Informally, this is the number of witnesses accepted with probability at least λ.

Problem 2 (**Quantum approximate counting).** *We are given a verifier circuit with n input qubits and size poly(n), a precision parameter $\delta \geq$ poly$(1/n)$, and two thresholds a, b such that $0 < b < a \leq 1$ and $a - b \geq$ poly$(1/n)$. Compute an estimate ξ satisfying*

$$(1-\delta)N_a \leq \xi \leq (1+\delta)N_b. \tag{55}$$

Below we use an acronym QXC which stands for Quantum approXimate Counting problem[7]. Our main result is as follows.

Theorem 5. *For any $k \geq 2$, k-QPF is polynomial time equivalent to QXC.*

We also prove polynomial-time equivalence between QXC and two other important problems: estimating the quantum density of states for a k-local Hamiltonian and estimating the quantum mean value of a Pauli observable for the thermal Gibbs state of a k-local Hamiltonian.

Next, we investigate exponential-time classical and quantum algorithms which solve the QPF problem for general k-local Hamiltonians. We show that a *Clifford compression* technique [274, 275] based on the unitary 2-design property of the Clifford group [276, 277] can be used to improve the runtime of state-of-the art classical algorithms for QPF. We then use the same technique to almost halve the memory footprint of state-of-the-art quantum algorithms for QPF without compromising their runtime.

We give a classical algorithm that solves the QPF problem for any n-qubit k-local Hamiltonian H with $\beta\|H\| \leq b$ in time

$$O\left((b + \log(1/\delta))\ell 2^n \delta^{-1} + n^2 2^n \delta^{-1} + \delta^{-4}\right), \tag{56}$$

where $\ell \leq O(n^k)$ is the number of non-zero k-local terms that appear in H. For example, if H is a geometrically local Hamiltonian on a regular lattice with bounded strength interactions

[7]The original work [54] used an acronym QAC instead of QXC used here. Both acronyms refer to the same computational problem.

then $\|H\| = O(n)$ and $\ell = O(n)$. For such Hamiltonians the runtime in Eq. (56) becomes $O((1+\beta)n^2 2^n/\delta)$, ignoring terms logarithmic in $1/\delta$ and assuming $2^n \geq 1/\delta^3$.

To obtain the runtime Eq. (56) we use a stochastic trace estimator due to Hutchinson [278] and its modern version known as Hutch++ [279]. The latter algorithm approximates the trace of a positive-semidefinite matrix A of size $d \times d$ within a relative error δ by performing only $O(1/\delta)$ matrix-vector multiplications for the matrix A and suitable d-dimensional vectors. However, the overall runtime of Hutch++ has an additional term scaling as d/δ^2 which has a quadratically worse dependence on δ. Here we show how to improve this scaling to $d\log^2(d)/\delta$, assuming that $d \geq 1/\delta^3$. To this end we apply the original Hutch++ algorithm to a smaller matrix obtained from A using the Clifford compression. Our construction shares some features of the methods used in Refs. [274, 275] to obtain compressed classical descriptions of quantum states. Specializing the improved Hutch++ to the matrix exponential $A = e^{-\beta H}$ with a k-local Hamiltonian H gives a classical algorithm for the QPF problem with runtime Eq. (56).

Secondly, we give a memory-efficient quantum algorithm for the QPF problem which requires only $O(\log(n) + \log(1/\delta))$ ancilla qubits. This is an improvement in space requirements over an earlier algorithm due to Poulin and Wocjan [280] which needed an $\Omega(n)$-sized ancilla register. The running-time of our algorithm is $\tilde{O}(\sqrt{2^n/\mathcal{Z}} \cdot \beta/\delta)$, slightly better than that reported in Ref. [280]. But note that the improved running-time follows essentially from new techniques for Gibbs-state preparation [281, 282, 283].

The algorithm of Ref. [280] constructs a quantum circuit that prepares a purification of a finite-temperature Gibbs state using Hamiltonian simulation and quantum phase estimation as an intermediate step. The overhead of $\Omega(n)$ ancilla qubits in this approach is a consequence of purifying an n-qubit mixed state. We use Clifford compression to get around this bottleneck: a similar purification step in our algorithm requires only $O(\log(1/\delta))$ ancilla qubits.

4.4 Quantum complexity of the Kronecker coefficients

The Kronecker coefficients appear as the analogues of the Clebsch-Gordan coefficients for the symmetric group S_n. Recall that a *unitary* representation ρ of S_n is a homomorphism from S_n to the unitary matrices. Any such representation can be block-diagonalized into a set of *irreducible representations (irreps)* as $\rho \simeq \bigoplus_\lambda m_\lambda \rho_\lambda$, where the $\simeq$ stands for isomorphism, $m_\lambda \geq 0$ is a non-negative integer *multiplicity* of an *irrep* ρ_λ and the direct sum runs over all inequivalent irreducible representations of S_n.[8] The S_n irreps can be labelled by partitions of n.

A tensor product of two irreps can also be block-diagonalized: if we let ρ_μ and ρ_ν to be irreps of S_n labeled by partitions $\mu, \nu \vdash n$, the Kronecker coefficient $g_{\mu\nu\lambda}$ is the multiplicity of the irrep ρ_λ in the representation $\rho_\mu \otimes \rho_\nu$:

$$\rho_\mu \otimes \rho_\nu \simeq \bigoplus_\lambda g_{\mu\nu\lambda} \rho_\lambda. \tag{57}$$

[8] As per the usual convention, we mean $\bigoplus_\lambda m_\lambda \rho_\lambda \equiv \bigoplus_\lambda \rho_\lambda^{\otimes m_\lambda}$ and not a multiplication by an integer.

The definition Eq. (57) ensures that the Kronecker coefficients are nonnegative integers[9], but it is a longstanding open problem in algebraic combinatorics—Problem 10 on Stanley's list [219]—to find a combinatorial formula for them. That is, *do the Kronecker coefficients count some natural set of combinatorial objects?* Let us interpret "natural" as any set of objects where the description size is polynomial in n and such that membership in the set can be verified efficiently using a classical computer.

We define the problem of exactly computing the Kronecker coefficients as follows; compute $g_{\mu\nu\lambda}$ for input partitions $\mu, \nu, \lambda \in S_n$ given in unary, so that the input size is $O(n)$.

An (old) new quantum algorithm. We showed [55] that there is an efficient quantum algorithm that gives an approximation of $g_{\mu\nu\lambda}$ with additive error $\epsilon \frac{d_\mu d_\nu}{d_\lambda}$ in $O(\mathrm{poly}(n)/\epsilon^2)$ time, where d_μ, d_ν, d_λ are the dimensions of the respective irreducible representations. This algorithm is a consequence of the sieving algorithm for the non-Abelian Hidden Subgroup problem that was discussed in Ref. [284]. The approximation obtained this way is good enough to determine $g_{\mu\nu\lambda}$ exactly if $\min(d_\mu, d_\nu, d_\lambda) \leq \mathrm{poly}(n)$. We do not know if such approximation can be obtained classically and, to our best knowledge, the best polynomial-time classical algorithms for the same problem are known only in the special case when the partitions μ, ν, λ have constant, or nearly constant number of parts [221].

Connection to quantum complexity. There is little doubt that computing the Kronecker coefficients in the general case is an extremely hard computational problem as it is known that any problem in #P can be reduced in polynomial time to it [222][10], and that it is contained in the class GapP [220] of functions that can be expressed as the difference $f - g$ of two functions f and g in #P [285]. In this sense, the Kronecker coefficients are as hard as #P, but not much harder. Pinning down their complexity is therefore entirely concerned with the sliver of daylight between #P and GapP. Why should we care?

Aside from attacking Stanley's problem #10 from a new direction, it has to do with the complexity of approximating Kronecker coefficients to within a given relative error.

Problem 3 (ApproxKron)**.** *Given $\mu, \nu, \lambda \vdash n$ and $\epsilon = \Omega(1/\mathsf{poly}(n))$, compute an estimate $\tilde{g}$, such that:*

$$(1 - \epsilon) g_{\mu\nu\lambda} \leq \tilde{g} \leq (1 + \epsilon) g_{\mu\nu\lambda}.$$

This problem is of course no harder than exact computation of $g_{\mu\nu\lambda}$; but can it be much easier? It turns out that the seemingly small difference between #P and GapP can lead to drastic difference in the complexity of the corresponding *approximation* problems. It is known that some problems in GapP are just as hard to approximate as they are to compute exactly (with respect to polynomial-time reductions). On the other hand, Stockmeyer has shown that, assuming a standard conjecture in complexity theory, the hardest functions in #P are *vastly easier to approximate*[267]: any #P function can be approximated to a given relative error using a polynomial-time randomized algorithm that has access to an NP oracle, i.e., in the class $\mathsf{FBPP}^{\mathsf{NP}}$. Computing a multiplicative approximation to a #P problem

[9] A negative $g_{\mu\nu\lambda}$ would mean that there is a negative number of blocks in a block-diagonalized matrix $\rho_\mu \otimes \rho_\nu$

[10] See Remark 14 in Ref. [223]

	Problem in...	Approximation problem upper bound...
Classical counting	#P	$\mathsf{FBPP}^{\mathsf{NP}}$
Quantum counting	#BQP	QXC
Gap counting	GapP	#P

Table 3: Three classes of counting problems that are polynomial-time reducible to one another. The associated approximation problems are very unlikely to be polynomial-time equivalent. Classical counting problems in #P can be approximated within the third level of the polynomial hierarchy via Stockmeyer's approximate counting algorithm. In contrast, some problems in GapP can be #P-hard to approximate. Quantum approximate counting problems QXC lie somewhere between these two extremes.

thus cannot be #P-hard unless the polynomial hierarchy collapses, which is believed to be unlikely.

A central message of our work [55] and Ref. [54] is that in between these two extremes there is a rich class of approximation tasks associated with a third type of *quantum approximate counting* problem. In analogy to the definition of #P, Refs. [286, 287] defined the class #BQP of functions that can be thought of, informally, as determining the dimension of the subspace spanned by accepting witness states of a QMA verifier. This can be thought of as the quantum analogue of counting the number of accepting witnesses. The corresponding class QXC of *quantum approximate counting problems* [54] is then the approximation version of #BQP, see Section 4.3. As noted in this section, the complexity of quantum approximate counting is largely an open question; all we can say is that quantum approximate counting is at least QMA-hard and at most #P-hard. Nevertheless, we conjecture that quantum approximate counting is not #P-hard—i.e., that it defines an intermediate class that lies somewhere between classical approximate counting and #P. Table 3 summarizes the three types of counting problems discussed above, and their approximation versions.

We show the that ApproxKron is contained in QXC by giving an efficient quantum circuit that measures a projector with rank $d_\mu d_\nu d_\lambda g_{\mu\nu\lambda}$. Since the dimension d_ω of any irreducible representation ω of S_n can be computed efficiently, this shows that $g_{\mu\nu\lambda}$ is equal to an efficiently computable dimensional factor $(d_\mu d_\nu d_\lambda)^{-1}$ times a #BQP function. From this, we conclude that the problem of approximating Kronecker coefficients to within a given relative error is not harder than QXC. This furnishes a natural computational problems that is polynomial-time reducible to quantum approximate counting (and not known to be easier) but whose definition does not involve quantum many-body systems[11]. Our proof technique relies on Beals' efficient quantum Fourier transform over the symmetric group [289] and its use in the *generalized phase estimation algorithm* (GPE) [290] . We use GPE to construct projectors onto irreducible representations contained in representations that can be efficiently implemented as quantum circuits. Specifically, we give quantum circuits that project from the regular representation into its irreducibles and another circuit that projects a three-fold tensor copy of regular representations onto its irreps. We show that the subspace projected

[11]The only other problem of this kind that we are aware of involves approximating Betti numbers of a cochain complex [288] and particular, the approximation version of the problem described in Theorem 5 of Reference [288]

on by these mutually commuting measurements has dimension $d_\mu d_\nu d_\lambda g_{\mu\nu\lambda}$, which implies the containment in QXC. We also show that deciding positivity of Kronecker coefficients is in QMA, with perfect soundness and completeness.

4.5 Simulation of open quantum systems

Many classical algorithms based on random walks can be quantized using a generic method developed by Szegedy [170]. An important contribution of [170] was the quantization of any reversible (also called detailed balanced) random walk in the following sense. Let $P = (p_{yx})_{x,y\in\Omega}$ denote the stochastic matrix representing the reversible random walk on state space Ω with limiting distribution $\pi = (\pi_x)_{x\in\Omega}$. Szegedy showed how to construct a corresponding quantum walk unitary $W(P)$ such that its unique eigenvector with eigenvalue 1 (or equivalently with eigenphase 0) is the quantum sample

$$|\pi\rangle = \sum_{x\in\Omega} \sqrt{\pi_x}|x\rangle \tag{58}$$

that is, a coherent encoding of π. The reason for the many quantum speed-ups is that the phase gap of walk unitary $W(P)$ is $\sqrt{\Delta}$, where Δ denotes the spectral gap of the stochastic matrix P. Szegedy's construction can now be understood using the quantum singular value transformation [291] that provides a unifying approach to many quantum algorithms and methods.

Our work [225] gives a generalization of the Szegedy walk unitary from classical stochastic matrices to quantum channels (trace preserving completely positive maps) satisfying a certain detailed balance condition. Suppose $\mathcal{T}$ is a quantum map with fixed point σ. We construct a unitary $W(\mathcal{T})$ such that its eigenvector with eigenphase 0 is the purification

$$|\sigma^{1/2}\rangle = (\sigma^{1/2} \otimes I)|\Omega\rangle \tag{59}$$

of σ, where $|\Omega\rangle = \sum_{x\in\Omega} |x\rangle \otimes |x\rangle$ denotes the (unnormalized) maximally entangled state. We prove that the eigenphase of $W(\mathcal{T})$ is quadratically amplified compared to the spectral gap of the quantum map $\mathcal{T}$. We also present quantum circuits that efficiently implement $W(\mathcal{T})$. To accomplish these goals we proceed as follows.

First, we develop a generic framework for quantizing continuous time purely irreversible detailed balanced quantum maps. More precisely, we show how to quantize quantum Markov semigroups. Let $\mathcal{L}$ denote the Lindbladian, that is, the generator of a quantum Markov semigroup with fixed point σ. Assuming that $\mathcal{L}$ satisfies the detailed balanced condition with respect to σ, we show how to construct a unitary $W(\mathcal{L})$ such that its eigenvector with eigenphase 0 is a purification of σ having the above form Eq. (59). We leverage the fact that detailed balanced Lindbladians can be expressed in a certain canonical form showing that they essentially have the structure of Davies generators [292]. We relate the unitary $W(\mathcal{L})$ to a quantum discriminate $\mathcal{Q}$, which arises from $\mathcal{L}$ through a similarity transformation defined with respect to the fixed point σ. Analogously to the classical setting the quantum discriminate $\mathcal{Q}$ is shown to be independent of the fixed point σ and even the particular form in which the detailed balance condition is stated (due to the non-commutative nature of quantum mechanics, there are several natural notions of detailed balanced). Centrally, we

prove that the phase gap of $W(\mathcal{L})$ is quadratically amplified compared to the spectral gap of $\mathcal{L}$. We also discuss how the problem of quantizing detailed balanced quantum channels reduces to the problem of quantizing detailed balanced Lindbladians.

Second, we examine for which quantum maps the above quantization method can be realized efficiently on a quantum computer. We focus on the Davies weak coupling limit [292] describing a a Markovian dynamics driving a system with a Hamiltonian H towards the equilibrium thermal Gibbs states $e^{-\beta H}/\mathrm{Tr}\big(e^{-\beta H}\big)$. We show how to efficiently quantize Davies generators provided that the energies of the Hamiltonian H can be resolved, for instance, when we have access to a block-encoding of the Hamiltonian H whose energies satisfy a rounding promise and we use the energy estimation method in [293].

The quadratic gap amplification is at the heart the quantum speed-ups of many classical random walk based algorithms [208, 294, 295, 296]. This polynomial speed-up is measured relative to the mixing time of the classical walk. For quantum maps, recent work [297, 298, 299, 300, 301, 302] has investigated the mixing behavior of the corresponding Markov process. Tools have been developed to bound the mixing time and to analyze the spectral gaps of detailed balanced quantum maps. For several explicit examples [303, 304, 305, 306, 307, 308, 309] spectral gap bounds could be obtained.

A special case of a detailed balanced quantum map was introduced in the context of the quantum Metropolis algorithm [209] to prepare the Gibbs state of a quantum Hamiltonian with a time evolution that can be simulated efficiently. The Gibbs state can be prepared efficiently if the corresponding quantum map is rapidly mixing. A Szegedy walk unitary has been constructed for the classical Metropolis-Hastings algorithm [208]. The Metropolis-Hastings random walk prepares the Gibbs state for a Hamiltonian that is diagonal in the computational basis. An extension of this walk algorithm to non-diagonal quantum Hamiltonian subject to specific assumptions has been constructed in [310]. This work assumes access to a projective measurement that enables one to distinguish between some arbitrary but fixed eigenvector basis of the system Hamiltonian. This assumption essentially makes it possible to reduce the problem to a classical Metropolis random walk on the fixed eigenbasis. We emphasize that when the system Hamiltonian has degenerate spectra such eigenbasis measurement cannot be realized by measuring the energies of the Hamiltonian – even if one can perfectly resolve the energies. This is because one cannot distinguish between some arbitrary vectors within a degenerate eigenspace without any additional assumptions. Furthermore, in the generic situation of energy estimation with finite resources, any quantum Hamiltonian on an exponentially large Hilbert space with polynomial bounded operator norm will exhibit degeneracies. The present work enables the direct "quantization" of thermalizing quantum maps such as Davies generators and therefore does not need to make any assumptions on the identifiabilty of some eigenbases.

5 Publications, preprints, and patent applications

- Sergey Bravyi, Alexander Kliesch, Robert Koenig, Eugene Tang, *"Obstacles to Variational Quantum Optimization from Symmetry Protection"*, Physical Review Letters 125, 260505 (2020)
- Srinivasan Arunachalam, Reevu Maity *"Quantum Boosting"*, In International Conference on Machine Learning (pp. 377-387). PMLR
- Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, Mehdi Soleimanifar, *"Sample-efficient learning of interacting quantum systems"*, Nature Physics 931–935 (2021)
- Yunchao Liu, Srinivasan Arunachalam, Kristan Temme, *"A rigorous and robust quantum speed-up in supervised machine learning"*, Nature Physics 17, pp 1013–1017 (2021)
- Jennifer Glick, Tanvi Gujarati, Antonio Corcoles, Youngseok Kim, Abhinav Kandala, Jay M. Gambetta, Kristan Temme, *"Covariant quantum kernels for data with group structure"*, arXiv:2105.03406
- Sergey Bravyi, Anirban Chowdhury, David Gosset, Pawel Wocjan *"Quantum Hamiltonian complexity in thermal equilibrium"*, Nature Physics 18, pp. 1367–1370 (2022)
- Andrew Eddins, Mario Motta, Tanvi Gujarati, Sergey Bravyi, Antonio Mezzacapo, Charles Hadfield, and Sarah Sheldon *"Doubling the size of quantum simulators by entanglement forging"*, PRX Quantum 3, 010309 (2022)
- Libor Caha, Alexander Kliesch, Robert Koenig *"Twisted hybrid algorithms for combinatorial optimization"*, Quantum Science and Technology, 7 045013 (2022)
- Sergey Bravyi, Alexander Kliesch, Robert Koenig, Eugene Tang *"Hybrid quantum-classical algorithms for approximate graph coloring"*, Quantum 6, 678 (2022)
- Srinivasan Arunachalam, Alex B. Grilo, Tom Gur, Igor C. Oliveira, Aarthi Sundaram *"Quantum learning algorithms imply circuit lower bounds"*, Proceedings of 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)
- Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, Pawel Wocjan *"Simpler (classical) and faster (quantum) algorithms for Gibbs partition functions"* Proceedings of 2021 IEEE International Conference on Quantum Computing and Engineering (QCE), pp. 112-122
- Srinivasan Arunachalam, Yihui Quek, John Smolin *"Private learning implies quantum stability"* Proceedings of 35th Conference on Neural Information Processing Systems (NeurIPS 2021)
- Giacomo Nannicini, *"Fast quantum subroutines for the simplex method"*, Integer Programming and Combinatorial Optimization: 22nd International Conference, IPCO 2021, Atlanta, GA, USA, May 19–21, 2021, Proceedings 22, pages 311-325

- Sergey Bravyi, David Gosset, Daniel Grier, *"Classical algorithms for Forrelation"*, arXiv:2102.06963
- Brandon Augustino, Giacomo Nannicini, Tamás Terlaky, Luis F. Zuluaga *"Quantum interior point methods for semidefinite optimization"*, arXiv:2112.06025
- Sergey Bravyi, Isaac Kim, Alexander Kliesch, Robert Koenig *"Adaptive constant-depth circuits for manipulating non-abelian anyons"*, arXiv:2205.01933
- Pawel Wocjan, Kristan Temme *"Szegedy Walk Unitaries for Quantum Maps"*, arXiv:2107.07365
- Srinivasan Arunachalam, Sergey Bravyi, Arkopal Dutt, Theodore J. Yoder *"Optimal algorithms for learning quantum phase states"* arXiv:2208.07851
- Chen-Fu Chiang, Anirban Chowdhury, Pawel Wocjan *"Space-efficient Quantization Method for Reversible Markov Chains"*, arXiv:2206.06886
- Sergey Bravyi, Anirban Chowdhury, David Gosset, Vojtech Havlicek, Guanyu Zhu *"Quantum complexity of the Kronecker coefficients"*, arXiv:2302.11454
- Sergey Bravyi, Sarah Sheldon, Mario Motta, Antonio Mezzacapo, Tanvi Gujarati, Andrew Eddins *"Entanglement forging for quantum simulations"*, Patent application US202/0292381A1

References

[1] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum random access memory. *Physical Review Letters*, 100(16):160501, 2008.

[2] Scott Aaronson. Read the fine print. *Nature Physics*, 11(4):291–293, Apr 2015.

[3] Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

[4] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, pages 1–5, 2021.

[5] Jennifer R Glick, Tanvi P Gujarati, Antonio D Corcoles, Youngseok Kim, Abhinav Kandala, Jay M Gambetta, and Kristan Temme. Covariant quantum kernels for data with group structure. *arXiv preprint arXiv:2105.03406*, 2021.

[6] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V Albert, and John Preskill. Provably efficient machine learning for quantum many-body problems. *Science*, 377(6613):eabk3333, 2022.

[7] Eyal Bairey, Itai Arad, and Netanel H Lindner. Learning a local Hamiltonian from local measurements. *Physical Review Letters*, 122(2):020504, 2019.

[8] Eyal Bairey, Chu Guo, Dario Poletti, Netanel H Lindner, and Itai Arad. Learning the dynamics of open quantum systems from their steady states. *New Journal of Physics*, 22(3):032001, 2020.

[9] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David G Cory. Hamiltonian learning and certification using quantum resources. *Physical Review Letters*, 112(19):190501, 2014.

[10] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David Cory. Quantum Hamiltonian learning using imperfect quantum resources. *Physical Review A*, 89(4):042314, 2014.

[11] Tim J Evans, Robin Harper, and Steven T Flammia. Scalable bayesian Hamiltonian learning. *arXiv preprint arXiv:1912.07636*, 2019.

[12] Jianwei Wang, Stefano Paesani, Raffaele Santagati, Sebastian Knauer, Antonio A Gentile, Nathan Wiebe, Maurangelo Petruzzella, Jeremy L O'Brien, John G Rarity, Anthony Laing, et al. Experimental quantum Hamiltonian learning. *Nature Physics*, 13(6):551–555, 2017.

[13] Guy Bresler. Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 771–782, 2015.

[14] Adam Klivans and Raghu Meka. Learning graphical models using multiplicative weights. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 343–354. IEEE, 2017.

[15] Marc Vuffray, Sidhant Misra, Andrey Lokhov, and Michael Chertkov. Interaction screening: Efficient and sample-optimal learning of Ising models. In *Advances in Neural Information Processing Systems*, pages 2595–2603, 2016.

[16] Linus Hamilton, Frederic Koehler, and Ankur Moitra. Information theoretic properties of Markov random fields, and their algorithmic applications. In *Advances in Neural Information Processing Systems*, pages 2463–2472, 2017.

[17] Pradeep Ravikumar, Martin J Wainwright, John D Lafferty, et al. High-dimensional Ising model selection using L1-regularized logistic regression. *The Annals of Statistics*, 38(3):1287–1319, 2010.

[18] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17(8):931–935, 2021.

[19] Jeongwan Haah, Robin Kothari, and Ewin Tang. Optimal learning of quantum hamiltonians from high-temperature gibbs states. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 135–146. IEEE, 2022.

[20] Leslie G Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984.

[21] Srinivasan Arunachalam, Sergey Bravyi, Arkopal Dutt, and Theodore J Yoder. Optimal algorithms for learning quantum phase states. *arXiv preprint arXiv:2208.07851*, 2022.

[22] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[23] Matthew B Hastings. Classical and quantum bounded depth approximation algorithms. *arXiv preprint arXiv:1905.07047*, 2019.

[24] Zhang Jiang, Eleanor G Rieffel, and Zhihui Wang. Near-optimal quantum circuit for Grover's unstructured search using a transverse field. *Physical Review A*, 95(6):062317, 2017.

[25] Sergey Bravyi, Alexander Kliesch, Robert Koenig, and Eugene Tang. Obstacles to Variational Quantum Optimization from Symmetry Protection. *Phys. Rev. Lett.*, 125:260505, Dec 2020.

[26] Sergey Bravyi, Alexander Kliesch, Robert Koenig, and Eugene Tang. Hybrid quantum-classical algorithms for approximate graph coloring. *Quantum*, 6:678, 2022.

[27] Libor Caha, Alexander Kliesch, and Robert Koenig. Twisted hybrid algorithms for combinatorial optimization. *Quantum Science and Technology*, 7, 2022.

[28] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15):150502, 2009.

[29] Andrew M Childs, Robin Kothari, and Rolando D Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, 2017.

[30] Giacomo Nannicini. Fast quantum subroutines for the simplex method. *Operations Research*, 2022.

[31] Sergey Bravyi, Graeme Smith, and John A Smolin. Trading classical and quantum computational resources. *Physical Review X*, 6(2):021043, 2016.

[32] Tianyi Peng, Aram W Harrow, Maris Ozols, and Xiaodi Wu. Simulating large quantum circuits on a small quantum computer. *Physical Review Letters*, 125(15):150504, 2020.

[33] Takeshi Yamazaki, Shunji Matsuura, Ali Narimani, Anushervon Saidmuradov, and Arman Zaribafiyan. Towards the Practical Application of Near-Term Quantum Computers in Quantum Chemistry Simulations: A Problem Decomposition Approach. *arXiv:1806.01305*, 2018.

[34] Juha M Kreula, Laura García-Álvarez, Lucas Lamata, Stephen R Clark, Enrique Solano, and Dieter Jaksch. Few-Qubit Quantum-Classical Simulation of Strongly Correlated Lattice Fermions. *EPJ Quantum Technol.*, 3(1):11–30, 2016.

[35] Nicholas C Rubin. A hybrid Classical/Quantum Approach for Large-Scale Studies of Quantum Systems with Density Matrix Embedding Theory. *arXiv:1610.06910*, 2016.

[36] Bela Bauer, Dave Wecker, Andrew J Millis, Matthew B Hastings, and Matthias Troyer. Hybrid quantum-classical approach to correlated materials. *Physical Review X*, 6(3):031045, 2016.

[37] Sergey Bravyi and David Gosset. Complexity of Quantum Impurity Problems. *Comm. Math. Phys.*, 356(2):451–500, 2017.

[38] Xiao Yuan, Jinzhao Sun, Junyu Liu, Qi Zhao, and You Zhou. Quantum simulation with hybrid tensor networks, 2020.

[39] Yukio Kawashima, Marc P. Coons, Yunseong Nam, Erika Lloyd, Shunji Matsuura, Alejandro J. Garza, Sonika Johri, Lee Huntington, Valentin Senicourt, Andrii O. Maksymov, Jason H. V. Nguyen, Jungsang Kim, Nima Alidoust, Arman Zaribafiyan, and Takeshi Yamazaki. Efficient and accurate electronic structure simulation demonstrated on a trapped-ion quantum computer, 2021.

[40] Andrew Eddins, Mario Motta, Tanvi P Gujarati, Sergey Bravyi, Antonio Mezzacapo, Charles Hadfield, and Sarah Sheldon. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum*, 3(1):010309, 2022.

[41] Sam McArdle, Suguru Endo, Alan Aspuru-Guzik, Simon C Benjamin, and Xiao Yuan. Quantum computational chemistry. *Reviews of Modern Physics*, 92(1):015003, 2020.

[42] Jens Eisert, Marcus Cramer, and Martin B Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277, 2010.

[43] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of physics*, 326(1):96–192, 2011.

[44] Jin-Guo Liu, Yi-Hong Zhang, Yuan Wan, and Lei Wang. Variational quantum eigensolver with fewer qubits. *Physical Review Research*, 1(2):023025, 2019.

[45] Gabriel Kotliar, Sergej Y Savrasov, Kristjan Haule, Viktor S Oudovenko, O Parcollet, and CA Marianetti. Electronic structure calculations with dynamical mean-field theory. *Reviews of Modern Physics*, 78(3):865–951, 2006.

[46] Qiming Sun and Garnet Kin-Lic Chan. Quantum Embedding Theories. *Acc. Chem. Res.*, 49(12):2705–2712, 2016.

[47] Laimutis Bytautas, Thomas M Henderson, Carlos A Jiménez-Hoyos, Jason K Ellis, and Gustavo E Scuseria. Seniority and orbital symmetry as tools for establishing a full configuration interaction hierarchy. *The Journal of chemical physics*, 135(4):044119, 2011.

[48] Vincent E Elfving, Marta Millaruelo, José A Gámez, and Christian Gogolin. Simulating quantum chemistry in the seniority-zero space on qubit-based quantum computers. *arXiv preprint arXiv:2002.00035*, 2020.

[49] Klaas Gunst, Dimitri Van Neck, Peter Andreas Limacher, and Stijn De Baerdemacker. The seniority quantum number in tensor network states. *arXiv preprint arXiv:2008.00733*, 2020.

[50] A Yu Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003.

[51] Sergey Bravyi, Isaac Kim, Alexander Kliesch, and Robert Koenig. Adaptive constant-depth circuits for manipulating non-abelian anyons. *arXiv preprint arXiv:2205.01933*, 2022.

[52] Sergey Bravyi, Oliver Dial, Jay M Gambetta, Dario Gil, and Zaira Nazario. The future of quantum computing with superconducting qubits. *Journal of Applied Physics*, 132(16):160902, 2022.

[53] Mohsin Iqbal, Nathanan Tantivasadakarn, Ruben Verresen, Sara L Campbell, Joan M Dreiling, Caroline Figgatt, John P Gaebler, Jacob Johansen, Michael Mills, Steven A Moses, et al. Creation of non-abelian topological order and anyons on a trapped-ion processor. *arXiv preprint arXiv:2305.03766*, 2023.

[54] Sergey Bravyi, Anirban Chowdhury, David Gosset, and Pawel Wocjan. Quantum hamiltonian complexity in thermal equilibrium. *Nature Physics*, 18(11):1367–1370, 2022.

[55] Sergey Bravyi, Anirban Chowdhury, David Gosset, Vojtech Havlicek, and Guanyu Zhu. Quantum complexity of the kronecker coefficients. *arXiv preprint arXiv:2302.11454*, 2023.

[56] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, Sep 2017.

[57] Srinivasan Arunachalam and Ronald de Wolf. Guest column: A survey of quantum learning theory. *SIGACT News*, 48(2):41–67, June 2017.

[58] Vedran Dunjko and Hans J Briegel. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Reports on Progress in Physics*, 81(7):074001, jun 2018.

[59] Carlo Ciliberto, Mark Herbster, Alessandro Davide Ialongo, Massimiliano Pontil, Andrea Rocchetto, Simone Severini, and Leonard Wossnig. Quantum machine learning: a classical perspective. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2209):20170551, 2018.

[60] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91:045002, Dec 2019.

[61] Nathan Wiebe, Daniel Braun, and Seth Lloyd. Quantum algorithm for data fitting. *Phys. Rev. Lett.*, 109:050505, Aug 2012.

[62] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum algorithms for supervised and unsupervised machine learning, 2013.

[63] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, Sep 2014.

[64] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Phys. Rev. Lett.*, 113:130503, Sep 2014.

[65] Seth Lloyd, Silvano Garnerone, and Paolo Zanardi. Quantum algorithms for topological and geometric analysis of data. *Nature communications*, 7(1):1–7, 2016.

[66] Iris Cong and Luming Duan. Quantum discriminant analysis for dimensionality reduction and classification. *New Journal of Physics*, 18(7):073011, jul 2016.

[67] Iordanis Kerenidis and Anupam Prakash. Quantum recommendation systems. In *8th Innovations in Theoretical Computer Science Conference, ITCS*, volume 67 of *LIPIcs*, pages 49:1–49:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017.

[68] Fernando G. S. L. Brandão, Amir Kalev, Tongyang Li, Cedric Yen-Yu Lin, Krysta M. Svore, and Xiaodi Wu. Quantum SDP Solvers: Large Speed-Ups, Optimality, and Applications to Quantum Learning. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:14, 2019.

[69] Patrick Rebentrost, Adrian Steffens, Iman Marvian, and Seth Lloyd. Quantum singular-value decomposition of nonsparse low-rank matrices. *Phys. Rev. A*, 97:012327, Jan 2018.

[70] Zhikuan Zhao, Jack K. Fitzsimons, and Joseph F. Fitzsimons. Quantum-assisted gaussian process regression. *Phys. Rev. A*, 99:052331, May 2019.

[71] Ewin Tang. A quantum-inspired classical algorithm for recommendation systems. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC, page 217–228, 2019.

[72] Ewin Tang. Quantum-inspired classical algorithms for principal component analysis and supervised clustering, 2018.

[73] András Gilyén, Seth Lloyd, and Ewin Tang. Quantum-inspired low-rank stochastic regression with logarithmic dependence on the dimension, 2018.

[74] Nai-Hui Chia, Han-Hsuan Lin, and Chunhao Wang. Quantum-inspired sublinear classical algorithms for solving low-rank linear systems, 2018.

[75] Chen Ding, Tian-Yi Bao, and He-Liang Huang. Quantum-inspired support vector machine, 2019.

[76] Nai-Hui Chia, András Gilyén, Tongyang Li, Han-Hsuan Lin, Ewin Tang, and Chunhao Wang. Sampling-based sublinear low-rank matrix arithmetic framework for dequantizing quantum machine learning. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC, page 387–400, 2020.

[77] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Phys. Rev. A*, 98:032309, Sep 2018.

[78] Edward Farhi and Hartmut Neven. Classification with quantum neural networks on near term processors, 2018.

[79] Edward Grant, Marcello Benedetti, Shuxiang Cao, Andrew Hallam, Joshua Lockhart, Vid Stojevic, Andrew G Green, and Simone Severini. Hierarchical quantum classifiers. *NPJ Quantum Information*, 4(1):1–8, 2018.

[80] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Phys. Rev. A*, 101:032308, Mar 2020.

[81] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, nov 2019.

[82] Marcus Cramer, Martin B Plenio, Steven T Flammia, Rolando Somma, David Gross, Stephen D Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. Efficient quantum state tomography. *Nature Communications*, 1(1):1–7, 2010.

[83] Ashley Montanaro. Learning stabilizer states by Bell sampling. *arXiv preprint arXiv:1707.04012*, 2017.

[84] Géza Tóth, Witlef Wieczorek, David Gross, Roland Krischek, Christian Schwemmer, and Harald Weinfurter. Permutationally invariant quantum tomography. *Physical review letters*, 105(25):250403, 2010.

[85] Anurag Anshu and Srinivasan Arunachalam. A survey on the complexity of learning quantum states. *arXiv preprint arXiv:2305.20069*, 2023.

[86] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, Mar 2019.

[87] Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Phys. Rev. Lett.*, 122:040504, Feb 2019.

[88] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT, page 144–152, 1992.

[89] Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.

[90] Martin Anthony and Peter L. Bartlett. Function learning from interpolation. *Combinatorics, Probability and Computing*, 9(3):213–225, 2000.

[91] J. Shawe-Taylor, P. L. Bartlett, R. C. Williamson, and M. Anthony. Structural risk minimization over data-dependent hierarchies. *IEEE Transactions on Information Theory*, 44(5):1926–1940, 1998.

[92] Peter Bartlett and John Shawe-Taylor. *Generalization Performance of Support Vector Machines and Other Pattern Classifiers*, page 43–54. MIT Press, Cambridge, MA, USA, 1999.

[93] J. Shawe-Taylor and N. Cristianini. On the generalization of soft margin algorithms. *IEEE Transactions on Information Theory*, 48(10):2721–2735, 2002.

[94] Michael J Kearns. *The computational complexity of machine learning*. MIT press, 1990.

[95] Rocco A. Servedio and Steven J. Gortler. Equivalences and separations between quantum and classical learnability. *SIAM J. Comput.*, 33(5):1067–1092, May 2004.

[96] Ryan Sweke, Jean-Pierre Seifert, Dominik Hangleiter, and Jens Eisert. On the quantum versus classical learnability of discrete distributions, 2020.

[97] X. Gao, Z.-Y. Zhang, and L.-M. Duan. A quantum machine learning algorithm based on generative models. *Science Advances*, 4(12), 2018.

[98] John Shawe-Taylor and Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

[99] Michael J Kearns and Umesh V Vazirani. *An introduction to computational learning theory*. MIT press, 1994.

[100] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[101] Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13(4):850–864, November 1984.

[102] Persi Diaconis. Group representations in probability and statistics. *Lecture notes-monograph series*, 11:i–192, 1988.

[103] Risi Kondor and Marconi S Barbosa. Ranking with kernels in fourier space. In *COLT*, pages 451–463, 2010.

[104] Imre Risi Kondor. *Group theoretical methods in machine learning*. Columbia University, 2008.

[105] Yunlong Jiao and Jean-Philippe Vert. The kendall and mallows kernels for permutations. In *International Conference on Machine Learning*, pages 1935–1944. PMLR, 2015.

[106] AS Holevo. Covariant measurements and uncertainty relations. *Reports on mathematical physics*, 16(3):385–400, 1979.

[107] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. Quantum embeddings for machine learning. *arXiv preprint arXiv:2001.03622*, 2020.

[108] Matthew Otten, Imène R Goumiri, Benjamin W Priest, George F Chapline, and Michael D Schneider. Quantum machine learning using gaussian processes with performant quantum kernels. *arXiv preprint arXiv:2004.11280*, 2020.

[109] Nello Cristianini, John Shawe-Taylor, Andre Elisseeff, and Jaz Kandola. On kernel-target alignment. *Advances in neural information processing systems*, 14, 2001.

[110] Corinna Cortes, Mehryar Mohri, and Afshin Rostamizadeh. Learning non-linear combinations of kernels. *Advances in neural information processing systems*, 22, 2009.

[111] David H Ackley, Geoffrey E Hinton, and Terrence J Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive science*, 9(1):147–169, 1985.

[112] Geoffrey E Hinton, Terrence J Sejnowski, et al. Learning and relearning in Boltzmann machines. *Parallel distributed processing: Explorations in the microstructure of cognition*, 1(282-317):2, 1986.

[113] Andrea Montanari et al. Computational implications of reducing data to sufficient statistics. *Electronic Journal of Statistics*, 9(2):2370–2390, 2015.

[114] Sergey Bravyi and David Gosset. Polynomial-time classical simulation of quantum ferromagnets. *Physical Review Letters*, 119(10):100503, 2017.

[115] Aram W Harrow, Saeed Mehraban, and Mehdi Soleimanifar. Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 378–386, 2020.

[116] Tomotaka Kuwahara, Kohtaro Kato, and Fernando GSL Brandão. Clustering of conditional mutual information for quantum Gibbs states above a threshold temperature. *Physical Review Letters*, 124(22):220601, 2020.

[117] Elliott H Lieb and Derek W Robinson. The finite group velocity of quantum spin systems. In *Statistical mechanics*, pages 425–431. Springer, 1972.

[118] Matthew B Hastings and Tohru Koma. Spectral gap and exponential decay of correlations. *Communications in Mathematical Physics*, 265(3):781–804, 2006.

[119] Sergey Bravyi, Matthew B Hastings, and Frank Verstraete. Lieb-Robinson bounds and the generation of correlations and topological quantum order. *Physical Review Letters*, 97(5):050401, 2006.

[120] Matthew B Hastings. Quantum belief propagation: An algorithm for thermal quantum systems. *Physical Review B*, 76(20):201102, 2007.

[121] Isaac H Kim. Markovian matrix product density operators: Efficient computation of global entropy. *arXiv preprint arXiv:1709.07828*, 2017.

[122] Kohtaro Kato and Fernando GSL Brandao. Quantum approximate Markov chains are thermal. *Communications in Mathematical Physics*, 370(1):117–149, 2019.

[123] Itai Arad, Tomotaka Kuwahara, and Zeph Landau. Connecting global and local energy distributions in quantum spin models on a lattice. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(3):033301, 2016.

[124] Marc Hein, Jens Eisert, and Hans J Briegel. Multiparty entanglement in graph states. *Physical Review A*, 69(6):062311, 2004.

[125] Dirk Schlingemann. Stabilizer codes can be realized as graph codes. *Quantum Info. Comput.*, 2(4):307–323, Jun 2002.

[126] Matteo Rossi, Marcus Huber, Dagmar Bruß, and Chiara Macchiavello. Quantum hypergraph states. *New Journal of Physics*, 15(11):113022, 2013.

[127] Ashley Montanaro. Quantum circuits and low-degree polynomials over $\mathbb{F}_2$. *Journal of Physics A: Mathematical and Theoretical*, 50(8):084002, Jan 2017.

[128] Michael J Bremner, Richard Jozsa, and Dan J Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, 2011.

[129] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, 2017.

[130] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In *Annual International Cryptology Conference*, pages 126–152. Springer, 2018.

[131] Zvika Brakerski and Omri Shmueli. (Pseudo) Random Quantum States with Binary Phase. In *Theory of Cryptography Conference*, pages 229–250. Springer, 2019.

[132] Prabhanjan Ananth, Luowen Qian, and Henry Yuen. Cryptography from pseudorandom quantum states. *arXiv preprint arXiv:2112.10020*, 2021.

[133] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum search-to-decision reductions and the state synthesis problem. *arXiv preprint arXiv:2111.02999*, 2021.

[134] Yuki Takeuchi, Tomoyuki Morimae, and Masahito Hayashi. Quantum computational universality of hypergraph states with Pauli-X and Z basis measurements. *Scientific Reports*, 9(1):1–14, 2019.

[135] Sergey Bravyi and David Gosset. Improved classical simulation of quantum circuits dominated by clifford gates. *Physical Review Letters*, 116(25):250501, 2016.

[136] Shawn X Cui, Daniel Gottesman, and Anirudh Krishna. Diagonal gates in the Clifford hierarchy. *Physical Review A*, 95(1):012329, 2017.

[137] Daniel Gottesman and Isaac L Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, 1999.

[138] Ashley Montanaro. The quantum query complexity of learning multilinear polynomials. *Information Processing Letters*, 112(11):438–442, 2012.

[139] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997.

[140] Martin Rötteler. Quantum algorithms to solve the hidden shift problem for quadratics and for functions of large Gowers norm. In *International Symposium on Mathematical Foundations of Computer Science*, pages 663–674. Springer, 2009.

[141] Isaac L Chuang and Michael A Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[142] Masoud Mohseni, Ali T Rezakhani, and Daniel A Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3):032322, Mar 2008.

[143] Richard A Low. Learning and testing algorithms for the Clifford group. *Physical Review A*, 80(5):052314, 2009.

[144] Ching-Yi Lai and Hao-Chung Cheng. Learning quantum circuits of some T gates. *IEEE Transactions on Information Theory*, 68(6):3951–3964, 2022.

[145] Aram W Harrow and Andreas Winter. How many copies are needed for state discrimination? *IEEE Transactions on Information Theory*, 58(1):1–2, 2012.

[146] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Average-case complexity versus approximate simulation of commuting quantum computations. *Physical review letters*, 117(8):080501, 2016.

[147] Leonardo Novo, Juani Bermejo-Vega, and Raúl García-Patrón. Quantum advantage from energy measurements of many-body quantum systems. *Quantum*, 5:465, 2021.

[148] Martin Rötteler. Quantum algorithms for highly non-linear Boolean functions. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete algorithms*, pages 448–457. SIAM, 2010.

[149] David Gross, Sepehr Nezami, and Michael Walter. Schur–Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations. *Communications in Mathematical Physics*, 385(3):1325–1393, 2021.

[150] Andrea Rocchetto. Stabiliser states are efficiently PAC-learnable. *arXiv preprint arXiv:1705.00345*, 2017.

[151] Daniel Liang. Clifford Circuits can be Properly PAC Learned if and only if $\mathsf{RP} = \mathsf{NP}$. *arXiv:2204.06638*, 2022.

[152] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5:4213, 2014.

[153] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M Chow, and Jay M Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242, 2017.

[154] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[155] S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-Robinson bounds and the generation of correlations and topological quantum order. *Physical Review Letters*, 97:050401, Jul 2006.

[156] Dorit Aharonov and Yonathan Touati. Quantum Circuit Depth Lower Bounds For Homological Codes. *arXiv preprint arXiv:1810.03912*, 2018.

[157] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):4812, 2018.

[158] Samson Wang, Enrico Fontana, M Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J Coles. Noise-induced barren plateaus in variational quantum algorithms. *arXiv preprint arXiv:2007.14384*, 2020.

[159] Sergey Bravyi, Alexander Kliesch, Robert Koenig, and Eugene Tang. Obstacles to variational quantum optimization from symmetry protection. *Physical Review Letters*, 125(26):260505, 2020.

[160] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, November 1995.

[161] Sergey Bravyi, David Gosset, and Daniel Grier. Classical algorithms for Forrelation. *arXiv preprint arXiv:2102.06963*, 2021.

[162] Alan Frieze and M. Jerrum. Improved approximation algorithms for Max k-Cut and Max Bisection. *Algorithmica*, 18(1):67–81, 05 1997.

[163] Christophe Andrieu, Nando De Freitas, Arnaud Doucet, and Michael I Jordan. An introduction to MCMC for machine learning. *Machine learning*, 50(1):5–43, 2003.

[164] Daniel Štefankovič, Santosh Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *Journal of the ACM (JACM)*, 56(3):1–36, 2009.

[165] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1):1–17, 1991.

[166] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM journal on computing*, 18(6):1149–1178, 1989.

[167] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM Journal on computing*, 22(5):1087–1116, 1993.

[168] Kurt Binder, Dieter Heermann, Lyle Roelofs, A John Mallinckrodt, and Susan McKay. Monte Carlo simulation in statistical physics. *Computers in Physics*, 7(2):156–157, 1993.

[169] Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 305:53–74, 2002.

[170] Mario Szegedy. Quantum speed-up of markov chain based algorithms. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 32–41, 2004.

[171] Pawel Wocjan and Anura Abeyesinghe. Speedup via quantum sampling. *Physical Review A*, 78(4):042336, 2008.

[172] Ashley Montanaro. Quantum speedup of monte carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301, 2015.

[173] Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, and Pawel Wocjan. Simpler (classical) and faster (quantum) algorithms for gibbs partition functions. *Quantum*, 6:789, 2022.

[174] Edward Farhi and Aram W Harrow. Quantum Supremacy through the Quantum Approximate Optimization Algorithm. *arXiv preprint arXiv:1602.07674*, 2016.

[175] Edward Farhi, David Gamarnik, and Sam Gutmann. The quantum approximate optimization algorithm needs to see the whole graph: A typical case. *arXiv preprint arXiv:2004.09002*, 2020.

[176] Edward Farhi, David Gamarnik, and Sam Gutmann. The quantum approximate optimization algorithm needs to see the whole graph: Worst case examples. *arXiv preprint arXiv:2005.08747*, 2020.

[177] Ruslan Shaydulin and Yuri Alexeev. Evaluating quantum approximate optimization algorithm: A case study. In *2019 Tenth International Green and Sustainable Computing Conference (IGSC)*, pages 1–6. IEEE, 2019.

[178] M. B. Hastings. Classical and Quantum Bounded Depth Approximation Algorithms. *arXiv preprint arXiv:1905.07047*, 2019.

[179] Zheng-Cheng Gu and Xiao-Gang Wen. Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order. *Physical Review B*, 80:155131, Oct 2009.

[180] Norbert Schuch, David Pérez-García, and Ignacio Cirac. Classifying quantum phases using matrix product states and projected entangled pair states. *Physical Review B*, 84:165139, Oct 2011.

[181] Frank Pollmann, Ari M. Turner, Erez Berg, and Masaki Oshikawa. Entanglement spectrum of a topological phase in one dimension. *Physical Revew B*, 81:064439, Feb 2010.

[182] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families iv: Bipartite Ramanujan graphs of all sizes. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1358–1377, Oct 2015.

[183] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families i: Bipartite Ramanujan graphs of all degrees. *Annals of Mathematics*, 182(1):307–325, 2015.

[184] M. Morgenstern. Existence and explicit constructions of q + 1 regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory, Series B*, 62(1):44 – 62, 1994.

[185] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *arXiv preprint cs/0405020*, 2004.

[186] L. Eldar and A. W. Harrow. Local Hamiltonians whose ground states are hard to approximate. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 427–438, Oct 2017.

[187] Francisco Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241, 1982.

[188] G.B. Mbeng, R. Fazio, and G. Santoro. Quantum Annealing: a journey through Digitalization, Control, and hybrid Quantum Variational schemes. *arXiv preprint arXiv:1906.08948*, 2019.

[189] Jun-Dong Cho, S. Raje, and M. Sarrafzadeh. Fast approximation algorithms on Max-Cut, k-coloring, and k-color ordering for VLSI applications. *IEEE Transactions on Computers*, 47(11):1253–1266, 1998.

[190] Paul M.B. Vitányi. How well can a graph be n-colored? *Discrete Mathematics*, 34(1):69 – 80, 1981.

[191] Viggo Kann, Sanjeev Khanna, Jens Lagergren, and Alessandro Panconesi. On the hardness of approximating max k-cut and its dual. Technical report, 1997.

[192] Etienne Klerk, Dmitrii Pasechnik, and J.P. Warners. On Approximate Graph Colouring and MAX-k-CUT algorithms based on the ϑ-function. *Journal of Combinatorial Optimization*, 8:267–294, 09 2004.

[193] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, September 2017.

[194] Edward Farhi, David Gamarnik, and Sam Gutmann. The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: Worst Case Examples, 2020. arXiv:2005.08747.

[195] Edward Farhi, David Gamarnik, and Sam Gutmann. The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: A Typical Case, 2020. arXiv:2005.09002.

[196] Daniel J. Egger, Jakub Mareček, and Stefan Woerner. Warm-starting quantum optimization. *Quantum*, 5:479, June 2021.

[197] Reuben Tate, Majid Farhadi, Creston Herold, Greg Mohler, and Swati Gupta. Bridging Classical and Quantum with SDP initialized warm-starts for QAOA, 2021. arXiv:2010.14021.

[198] Uriel Feige, Marek Karpinski, and Michael Langberg. Improved Approximation of Max-Cut on Graphs of Bounded Degree. *J. Algorithms*, 43(2):201–219, May 2002.

[199] Eran Halperin, Dror Livnat, and Uri Zwick. MAX CUT in cubic graphs. *Journal of Algorithms*, 53(2):169–185, 2004.

[200] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm, 2014. arXiv:1411.4028.

[201] Jonathan Wurtz and Peter Love. MaxCut quantum approximate optimization algorithm performance guarantees for $p > 1$. *Phys. Rev. A*, 103:042612, Apr 2021.

[202] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM J. Comput.*, 37(1):319–357, Apr 2007.

[203] Michel X. Goemans and David P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *J. ACM*, 42(6):1115–1145, Nov 1995.

[204] Anurag Anshu, David Gosset, and Karen Morenz. Beyond Product State Approximations for a Quantum Analogue of Max Cut. *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)*, 158:7:1–7:15, 2020.

[205] Mark Huber. Approximation algorithms for the normalizing constant of gibbs distributions. *The Annals of Applied Probability*, 25(2):974–985, 2015.

[206] Vladimir Kolmogorov. A faster approximation algorithm for the Gibbs partition function. In *Conference On Learning Theory*, pages 228–249. PMLR, 2018.

[207] Rolando Somma, Sergio Boixo, and Howard Barnum. Quantum simulated annealing. *arXiv preprint arXiv:0712.1008*, 2007.

[208] Rolando D. Somma, Sergio Boixo, Howard Barnum, and Emanuel Knill. Quantum simulations of classical annealing processes. *Physical review letters*, 101(13):130504, 2008.

[209] Kristan Temme, Tobias J Osborne, Karl G Vollbrecht, David Poulin, and Frank Verstraete. Quantum metropolis sampling. *Nature*, 471(7336):87–90, 2011.

[210] Chen-Fu Chiang, Anirban Chowdhury, and Pawel Wocjan. Space-efficient quantization method for reversible markov chains. *arXiv preprint arXiv:2206.06886*, 2022.

[211] Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

[212] Raphael Yuster and Uri Zwick. Fast sparse matrix multiplication. *ACM Transactions On Algorithms (TALG)*, 1(1):2–13, 2005.

[213] Marco E Lübbecke and Jacques Desrosiers. Selected topics in column generation. *Operations research*, 53(6):1007–1023, 2005.

[214] D. Gottesman and I. L. Chuang. Quantum teleportation is a universal computational primitive. *Nature*, 402:390–393, 1999.

[215] P. Selinger. Quantum circuits of T-depth one. *Physical Review A*, 87(4):042302, 2013.

[216] M. Amy, D. Maslov, and M. Mosca. Polynomial-time T-depth optimization of Clifford+T circuits via matroid partitioning. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 33(10):1476–1489, 2014.

[217] E. T. Campbell and M. Howard. Unifying gate synthesis and magic state distillation. *Physical review letters*, 118(6):060501, 2017.

[218] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, USA, 2002.

[219] Richard P. Stanley. Positivity problems and conjectures in algebraic combinatorics. In *in Mathematics: Frontiers and Perspectives*, pages 295–319. American Mathematical Society, 1999.

[220] Peter Bürgisser and Christian Ikenmeyer. The complexity of computing kronecker coefficients. In *Discrete Mathematics and Theoretical Computer Science*, pages 357–368. Discrete Mathematics and Theoretical Computer Science, 2008.

[221] Igor Pak and Greta Panova. On the complexity of computing Kronecker coefficients. *arXiv e-prints*, page arXiv:1404.0653, April 2014.

[222] Christian Ikenmeyer, Ketan D Mulmuley, and Michael Walter. On vanishing of kronecker coefficients. *computational complexity*, 26(4):949–992, 2017.

[223] Igor Pak and Greta Panova. Breaking down the reduced kronecker coefficients. *Comptes Rendus. Mathématique*, 358(4):463–468, 2020.

[224] Christian Ikenmeyer and Igor Pak. What is in #P and what is not? *arXiv e-prints*, page arXiv:2204.13149, April 2022.

[225] Pawel Wocjan and Kristan Temme. Szegedy walk unitaries for quantum maps. *arXiv preprint arXiv:2107.07365*, 2021.

[226] Héctor J García, Igor L Markov, and Andrew W Cross. On the geometry of stabilizer states. *Quantum Information & Computation*, 14(7&8):683–720, 2014.

[227] Robert M Parrish, Edward G Hohenstein, Peter L McMahon, and Todd J Martínez. Quantum computation of electronic transitions using a variational quantum eigensolver. *Physical review letters*, 122(23):230401, 2019.

[228] Igor L Markov and Yaoyun Shi. Simulating quantum computation by contracting tensor networks. *SIAM Journal on Computing*, 38(3):963–981, 2008.

[229] Suguru Endo, Jinzhao Sun, Ying Li, Simon C Benjamin, and Xiao Yuan. Variational quantum simulation of general processes. *Physical Review Letters*, 125(1):010501, 2020.

[230] Gabriel Greene-Diniz and David Muñoz Ramo. Generalized unitary coupled cluster excitations for multireference molecular states optimized by the variational quantum eigensolver. *International Journal of Quantum Chemistry*, 121(4):e26352, 2021.

[231] Patrick Huembeli, Giuseppe Carleo, and Antonio Mezzacapo. Entanglement forging with generative neural network models. *arXiv preprint arXiv:2205.00933*, 2022.

[232] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003.

[233] S. B. Bravyi and A. Yu. Kitaev. Quantum codes on a lattice with boundary. *arXiv preprint arXiv:quant-ph/9811052*, 1998.

[234] K. J. Satzinger, Y.-J Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth, C. Gidney, I. Aleiner, F. Arute, K. Arya, J. Atalaya, R. Babbush, J. C. Bardin, R. Barends, J. Basso, A. Bengtsson, A. Bilmes, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, B. Chiaro,

R. Collins, W. Courtney, S. Demura, A. R. Derk, D. Eppens, C. Erickson, L. Faoro, E. Farhi, A. G. Fowler, B. Foxen, M. Giustina, A. Greene, J. A. Gross, M. P. Harrigan, S. D. Harrington, J. Hilton, S. Hong, T. Huang, W. J. Huggins, L. B. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, T. Khattar, S. Kim, P. V. Klimov, A. N. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, A. Locharla, E. Lucero, O. Martin, J. R. McClean, M. McEwen, K. C. Miao, M. Mohseni, S. Montazeri, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, T. E. O'Brien, A. Opremcak, B. Pató, A. Petukhov, N. C. Rubin, D. Sank, V. Shvarts, D. Strain, M. Szalay, B. Villalonga, T. C. White, Z. Yao, P. Yeh, J. Yoo, A. Zalcman, H. Neven, S. Boixo, A. Megrant, Y. Chen, J. Kelly, V. Smelyanskiy, A. Kitaev, M. Knap, F. Pollmann, and P. Roushan. Realizing topologically ordered states on a quantum processor. *Science*, 374(6572):1237–1241, 2021.

[235] Carlos Mochon. Anyon computers with smaller groups. *Physical Review A*, 69(3), Mar 2004.

[236] M. Aguado, G. K. Brennen, F. Verstraete, and J. I. Cirac. Creation, Manipulation, and Detection of Abelian and Non-Abelian Anyons in Optical Lattices. *Phys. Rev. Lett.*, 101:260501, Dec 2008.

[237] G K Brennen, M Aguado, and J I Cirac. Simulations of quantum double models. *New Journal of Physics*, 11(5):053009, May 2009.

[238] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, 2002.

[239] Miguel Aguado. From entanglement renormalisation to the disentanglement of quantum double models. *Annals of Physics*, 326(9):2444–2473, 2011.

[240] Oscar Higgott, Matthew Wilson, James Hefford, James Dborin, Farhan Hanif, Simon Burton, and Dan E. Browne. Optimal local unitary encoding circuits for the surface code. *Quantum*, 5:517, August 2021.

[241] Yu-Jie Liu, Kirill Shtengel, Adam Smith, and Frank Pollmann. Methods for simulating string-net states and anyons on a digital quantum computer. *arXiv preprint arXiv:2110.02020*, 2021.

[242] Miguel Aguado and Guifré Vidal. Entanglement renormalization and topological order. *Phys. Rev. Lett.*, 100:070404, Feb 2008.

[243] Dorit Aharonov and Yonathan Touati. Quantum Circuit Depth Lower Bounds For Homological Codes. *arXiv preprint arXiv:1810.03912*, 2018.

[244] Jeongwan Haah and John Preskill. Logical-operator tradeoff for local quantum codes. *Phys. Rev. A*, 86:032308, Sep 2012.

[245] James R. Wootton, Jan Burri, Sofyan Iblisdir, and Daniel Loss. Error correction for non-abelian topological quantum computation. *Phys. Rev. X*, 4:011051, Mar 2014.

[246] Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005.

[247] Daniel Gottesman and Isaac Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402:390 – 393, 1999.

[248] Richard Jozsa. An introduction to measurement based quantum computation. *NATO Science Series, III: Computer and Systems Sciences. Quantum Information Processing-From Theory to Experiment*, 199:137–158, 2006.

[249] Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. *Computational Complexity*, 25(4):849–881, 2016.

[250] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[251] Dan Browne, Elham Kashefi, and Simon Perdrix. Computational Depth Complexity of Measurement-Based Quantum Computation. In *Conference on Quantum Computation, Communication, and Cryptography*, pages 35–46. Springer, 2010.

[252] Adam Paetznick and Krysta M Svore. Repeat-until-success: Non-deterministic decomposition of single-qubit unitaries. *arXiv preprint arXiv:1311.1074*, 2013.

[253] Zhenning Liu and Alexandru Gheorghiu. Depth-efficient proofs of quantumness. *arXiv preprint arXiv:2107.02163*, 2021.

[254] Matthew Coudron and Sanketh Menda. Computations with greater quantum depth are strictly more powerful (relative to an oracle). In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 889–901, 2020.

[255] Andrew M Childs, Richard Cleve, Enrico Deotto, Edward Farhi, Sam Gutmann, and Daniel A Spielman. Exponential algorithmic speedup by a quantum walk. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 59–68, 2003.

[256] Lorenzo Piroli, Georgios Styliaris, and J. Ignacio Cirac. Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter. *Phys. Rev. Lett.*, 127:220503, Nov 2021.

[257] Michael A Levin and Xiao-Gang Wen. String-net condensation: A physical mechanism for topological phases. *Physical Review B*, 71(4):045110, 2005.

[258] Ruben Verresen, Nathanan Tantivasadakarn, and Ashvin Vishwanath. Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices. *arXiv preprint arXiv:2112.03061*, 2022.

[259] Nathanan Tantivasadakarn, Ryan Thorngren, Ashvin Vishwanath, and Ruben Verresen. Long-range entanglement from measuring symmetry-protected topological phases. *arXiv preprint arXiv:2112.01519*, 2021.

[260] Guanyu Zhu, Ali Lavasani, and Maissam Barkeshli. Universal logical gates on topologically encoded qubits via constant-depth unitary circuits. *Phys. Rev. Lett.*, 125:050502, Jul 2020.

[261] Daniel Grier and Luke Schaeffer. *Interactive Shallow Clifford Circuits: Quantum Advantage against NC^1 and Beyond*, page 875–888. Association for Computing Machinery, New York, NY, USA, 2020.

[262] David A. Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in NC1. *Journal of Computer and System Sciences*, 38(1):150–164, 1989.

[263] Fernando G. S. L. Brandão. *Entanglement Theory and the Quantum Simulation of Many-Body Physics*. PhD thesis, Imperial College of Science, Technology and Medicine, 2008.

[264] Chris Cade and Ashley Montanaro. The quantum complexity of computing schatten p-norms. *arXiv preprint arXiv:1706.09279*, 2017.

[265] Anirban N. Chowdhury, Rolando D. Somma, and Yigit Subasi. Computing partition functions in the one-clean-qubit model. *Phys. Rev. A*, 103:032422, Mar 2021.

[266] Emanuel Knill and Raymond Laflamme. Power of one bit of quantum information. *Physical Review Letters*, 81(25):5672, 1998.

[267] Larry Stockmeyer. The complexity of approximate counting. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC '83, page 118–126, New York, NY, USA, 1983. Association for Computing Machinery.

[268] Leslie Ann Goldberg and Heng Guo. The complexity of approximating complex-valued Ising and Tutte partition functions. *computational complexity*, 26(4):765–833, 2017.

[269] D. Aharonov, M. Ben-Or, F. GSL Brandao, and O. Sattath. The pursuit for uniqueness: Extending Valiant-Vazirani theorem to the probabilistic and quantum settings. *arXiv preprint arXiv:0810.4840*, 2008.

[270] Toby S Cubitt, Ashley Montanaro, and Stephen Piddock. Universal quantum Hamiltonians. *Proceedings of the National Academy of Sciences*, 115(38):9497–9502, 2018.

[271] Leo Zhou and Dorit Aharonov. Strongly universal hamiltonian simulators. *arXiv preprint arXiv:2102.02991*, 2021.

[272] Tamara Kohler, Stephen Piddock, Johannes Bausch, and Toby Cubitt. Translationally-invariant universal quantum hamiltonians in 1d. *arXiv preprint arXiv:2003.13753*, 2020.

[273] Tamara Kohler, Stephen Piddock, Johannes Bausch, and Toby Cubitt. General conditions for universality of quantum hamiltonians. *arXiv preprint arXiv:2101.12319*, 2021.

[274] David Gosset and John Smolin. A compressed classical description of quantum states. In *14th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[275] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

[276] Richard Cleve, Debbie Leung, Li Liu, and Chunhao Wang. Near-linear constructions of exact unitary 2-designs. *arXiv preprint arXiv:1501.04592*, 2015.

[277] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80(1):012304, 2009.

[278] Michael F Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, 18(3):1059–1076, 1989.

[279] Raphael A Meyer, Cameron Musco, Christopher Musco, and David P Woodruff. Hutch++: Optimal stochastic trace estimation. In *Symposium on Simplicity in Algorithms (SOSA)*, pages 142–155. SIAM, 2021.

[280] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Phys. Rev. Lett.*, 103:220502, 2009.

[281] Anirban Narayan Chowdhury and Rolando D. Somma. Quantum algorithms for gibbs sampling and hitting-time estimation. *Quant. Inf. Comp.*, 17(1–2):41–64, 2017.

[282] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *ACM Symposium on Theory of Computing 2019*, page 193–204, New York, NY, USA, 2019. Association for Computing Machinery.

[283] Joran van Apeldoorn, András Gilyén, Sander Gribling, and Ronald de Wolf. Quantum SDP-Solvers: Better upper and lower bounds. *Quantum*, 4:230, February 2020.

[284] Cristopher Moore, Alexander Russell, and Piotr Sniady. On the impossibility of a quantum sieve algorithm for graph isomorphism. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing - STOC '07*. ACM Press, 2007.

[285] Stephen A. Fenner, Lance J. Fortnow, and Stuart A. Kurtz. Gap-definable counting classes. *Journal of Computer and System Sciences*, 48(1):116–148, 1994.

[286] Yaoyun Shi and Shengyu Zhang. Note on quantum counting classes. *Available from Shengyu Zhang's homepage: http://www. cse. cuhk. edu. hk/syzhang/papers/SharpBQP. pdf*, 2009.

[287] Brielin Brown, Steven T Flammia, and Norbert Schuch. Computational difficulty of computing the density of states. *Physical review letters*, 107(4):040501, 2011.

[288] Marcos Crichigno and Tamara Kohler. Clique homology is QMA1-hard. *arXiv preprint arXiv:2209.11793*, 2022.

[289] Robert Beals. Quantum computation of fourier transforms over symmetric groups. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 48–53, 1997.

[290] Aram W. Harrow. Applications of coherent classical communication and the Schur transform to quantum information theory. *arXiv e-prints*, pages quant–ph/0512255, December 2005.

[291] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 193–204, New York, NY, USA, 2019. Association for Computing Machinery.

[292] E Brian Davies. Markovian master equations. *Communications in mathematical Physics*, 39(2):91–110, 1974.

[293] Patrick Rall. Faster coherent quantum algorithms for phase, energy, and amplitude estimation. arXiv:2103.09717, 2021.

[294] David Poulin and Pawel Wocjan. Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer. *Phys. Rev. Lett.*, 103:220502, Nov 2009.

[295] Aram W. Harrow and Annie Y. Wei. Adaptive quantum simulated annealing for bayesian inference and estimating partition functions. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 193–212. SIAM, 2020.

[296] Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, and Pawel Wocjan. Simpler (classical) and faster (quantum) algorithms for Gibbs partition functions. *arXiv preprint arXiv:2009.11270*, 2020.

[297] Robert Olkiewicz and Bogusław Zegarlinski. Hypercontractivity in noncommutative Lp spaces. *Journal of functional analysis*, 161(1):246–285, 1999.

[298] Kristan Temme, Michael. J. Kastoryano, Mary B. Ruskai, Michael. M. Wolf, and Frank Verstraete. The χ-divergence and mixing times of quantum markov processes. *Journal of Mathematical Physics*, 51(12):122201, 2010.

[299] Michael J. Kastoryano and Kristan Temme. Quantum logarithmic sobolev inequalities and rapid mixing. *Journal of Mathematical Physics*, 54(5):052202, 2013.

[300] Ivan Bardet. Estimating the decoherence time using non-commutative functional inequalities. *arXiv preprint arXiv:1710.01039*, 2017.

[301] Ivan Bardet and Cambyse Rouzé. Hypercontractivity and logarithmic sobolev inequality for non-primitive quantum markov semigroups and estimation of decoherence rates. *arXiv preprint arXiv:1803.05379*, 2018.

[302] Nilanjana Datta and Cambyse Rouzé. Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality. In *Annales Henri Poincaré*, volume 21, pages 2115–2150. Springer, 2020.

[303] Kristan Temme. Lower bounds to the spectral gap of Davies generators. *Journal of Mathematical Physics*, 54(12):122110, 2013.

[304] Michael J. Kastoryano and Fernando G. S. L. Brandao. Quantum gibbs samplers: The commuting case. *Communications in Mathematical Physics*, 344(3):915–957, 2016.

[305] Kristan Temme. Thermalization time bounds for Pauli stabilizer Hamiltonians. *Communications in Mathematical Physics*, 350(2):603–637, 2017.

[306] Angela Capel, Cambyse Rouzé, and Daniel Stilck França. The modified logarithmic Sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions. *arXiv preprint arXiv:2009.11817*, 2020.

[307] Ivan Bardet, Marius Junge, Nicholas LaRacuente, Cambyse Rouzé, and Daniel Stilck França. Group transference techniques for the estimation of the decoherence times and capacities of quantum markov semigroups. *IEEE Transactions on Information Theory*, 67(5):2878–2909, 2021.

[308] Ivan Bardet, Angela Capel, Angelo Lucia, David Pérez-García, and Cambyse Rouzé. On the modified logarithmic Sobolev inequality for the heat-bath dynamics for 1D systems. *Journal of Mathematical Physics*, 62(6):061901, 2021.

[309] Angelo Lucia, David Pérez-García, and Antonio Pérez-Hernández. Thermalization in Kitaev's quantum double models via tensor network techniques, 2021.

[310] Man-Hong Yung and Alán Aspuru-Guzik. A quantum–quantum Metropolis algorithm. *Proceedings of the National Academy of Sciences*, 109(3):754–759, 2012.