Space-saving holographic mean value algorithms for quantum circuits

Nathan Ju
University of Illinois at Urbana-Champaign

The task of approximating the expectation value of a n-qubit observable $O_1 \otimes ... \otimes O_n$ on a shallow quantum circuit U is a vital step in many variational quantum algorithms. Bravyi et al. [arXiv:1909.11485] asked if large-scale instances of this task on a 2D grid can be carried out using limited quantum resources, e.g., small number of qubits, and suggested the use of holographic quantum simulation. We provide a positive answer to this question, quadratically reducing the number of qubits needed for this task, and provide two new contributions on top of this. First, we show how to avoid the gate complexity increase inherent in holographic quantum simulation by adding one extra dimension to the circuit. Second, we extend the holographic simulation technique to achieve qubit savings in 3D quantum circuits. As a result, we demonstrate ways to tradeoff circuit depth and qubit counts, while preserving advantage over classical algorithms: instances of this task on 2D and 3D quantum circuits U can be estimated in less time than known classical algorithms and with fewer than n qubits when the depth d of U is bounded by $\log^{1/2}(n) \leq d \leq n^{1/2}$ and $d \leq n^{1/3}$, respectively, and with no extra gates in the 2D case.

1 Introduction

Given a n-qubit quantum circuit U, along with single-qubit Hermitian observables $O_1, ..., O_n$, the tensor product $O = O_1 \otimes ... \otimes O_n$ is an observable and its mean value with respect to the state $U|0^n\rangle$ is given by

$$\mu = \langle 0^n | U^\dagger O U | 0^n \rangle \tag{1}$$

The estimation of μ is called the mean value problem. The mean value problem can be efficiently solved on a quantum computer by repeatedly applying U to $|0^n\rangle$ and measuring each single-qubit observable. The average of the product of the measured eigenvalues is calculated until the error is diminished.

The mean value problem for tensor-product observables is a central step in many NISQ algorithms. In many variational quantum algorithms, the aim is to minimize a mean value of the form simply in eq. (1). Specifically for VQE [4] and QAOA [5], the key quantum step is to measure $\langle 0^n | U^\dagger H U | 0^n \rangle$ where H is a Hamiltonian that is a linear combination of a polynomial number of mean values taking on the form of eq. (1). In other NISQ optimization tasks such as quantum classifiers [6, 7], the measurement is directly of the form in eq. (1). The mean value, μ, can be viewed as an output probability for measurement outcomes on the state $U|0^n\rangle$. For

Submitted to:
TQC 2020

example, each O_j can be a projector onto some pair of basis states, and the mean value with respect to O_j is the probability of measuring one of the outcomes.

However, many NISQ devices are limited in both qubit count and quantum circuit depth. Additionally, imperfections in gate operations limit the accuracy of the computation. It is natural to ask whether or not certain computational problems can be solved using fewer resources to satisfy these constraints. Naturally, we can consider the qubit count, circuit depth, and gate count as resources to be minimized. It was shown that the mean value problem has an efficient classical solution when the circuit depth, d, is a constant value but scales poorly for super-constant depth [3]. We can also consider constraints on the geometry of the circuit U. In this paper we investigate the tradeoffs between depth, qubit count, and geometry of the holographic simulator solving the mean value problem of super-constant and sub-polynomial depth 2D quantum circuits and sub-polynomial 3D circuits.

We note that we will be concerned with computing an additive approximation $\widetilde{\mu}$ of the mean value μ such that $|\widetilde{\mu} - \mu| < \delta$ for $\delta > 0$. If U is a polynomial depth circuit, then this problem is BQP-complete by definition, and we shouldn't expect an efficient implementation until quantum computer reach large-scale fault-tolerance. If the depth of the circuit is constrained between $\Omega(\log^{1/2}(n))$ and $O(n^{\frac{1}{2}})$, then this becomes a restricted model of computation which we foresee to be an important bridge to realizing BQP computation on real-world devices. We investigate the mean value problem in this region.

Our main tool in computing $\widetilde{\mu}$ with few qubits is a holographic simulation technique introduced in [1, 2]. The author showed how a quasi-1D simulator can simulate a 2D many-body system by reducing one dimension of the 2D grid to time on the quasi-1D simulator with increased gate count and the number of qubits scaling as $O(k\sqrt{n})$ in the simulator where k is the longest correlation length. It was left as an open problem in [3] to extend the holographic simulation techniques to the mean value problem.

2 Contributions and comparison to prior work

In this paper, we consider the computational complexity of the additive-approximation mean value problem as a function of qubit count, quantum depth, and circuit geometry. We show that $\widetilde{\mu}$ can be computed using classical preprocessing and quantum computation in less time than the classical constant-depth algorithm and with fewer qubits than the natural quantum algorithm in the depth, d, region of $\log^{1/2}(n) \leq d \leq n^{1/2}$ for 2D grids and $d \leq n^{1/3}$ for 3D grids.

We do this by holographically converting one axis of the 2D grid to time on a reduced quasi-1D hybrid simulator through a $O(d^2 n)$ time classical preprocessor. The quantum algorithm uses $d\sqrt{n}$ qubits, $O(\frac{d\sqrt{n}}{\delta^2})$ quantum depth, and $O(\frac{nd}{\delta^2})$ more gates than U. We show that folding our hybrid simulator in half in a 3D geometry can maintain the same number of gates as U while preserving the same qubit count and circuit depth as the unfolded simulator. Finally, we extend our same techniques to simulation of mean value approximation of 3D grids using a quasi-2D simulator in $(d+1)n^{2/3}$ qubits and $O(\frac{dn^{1/3}}{\delta^2})$ quantum depth.

It was shown in [3] a purely classical algorithm for additively approximating the mean value on a 2D grid that scales in time $O(\frac{n}{\delta^2} 2^{O(d^2)})$. This particular algorithm performs poorly when the depth becomes super-constant. In addition, the same authors discovered a classical algorithm

for mean value estimation of a 3D grid scaling in time $\delta^{-2}2^{O(n^{1/3})}$ for constant-depth circuits. [8] uses similar ideas from holographic simulation on 2D systems to classically estimate output probabilities of random circuits in time $O(nD^3)$ where D is the bond dimension. Tensor network simulators [9, 10, 11, 12] enable simulation of constant-depth 2D circuits with superpolynomial runtime $2^{O(\sqrt{n})}$. And finally, the general quantum algorithm straightforwardly estimates the mean value in depth $O(\frac{d}{\delta^2})$ on n qubits.

Compared to other work, our results demonstrate a bridge between efficient classical simulation and BQP computation of mean values in quantum circuits.

3 Notation

Let $[n] = \{1, ..., n\}$. A 2D grid of qubits is a $\sqrt{n} \times \sqrt{n}$ 2D grid, G, where at row i and column j of G is a qubit that can be denoted $G_{i,j}$. Only one and two qubit gates are allowed, and two qubit gates are restricted to interacting with nearest neighbors on the grid. A 3D grid is a similar $n^{1/3} \times n^{1/3} \times n^{1/3}$ 3D grid where qubits can be denoted $G_{i,j,k}$, $i, j, k \in [n^{\frac{1}{3}}]$ and two qubit gates are restricted to nearest neighbors.

Definition 3.1. *Given a measurement observable O_j with $j \in [n]$, denote the lightcone D_j of qubit j to be the largest subcircuit of U such that $D_j^\dagger(O_j \otimes I_{else})D_j = U^\dagger(O_j \otimes I_{else})U$. I_{else} applies the identity operator to every qubit except j. We extend this definition for sets of qubits in the natural way $D_{\{j_1,...,j_k\}}$.*

We note that on a 2D grid, the support of D_j in $[n]$ is centered at qubit j with a radius of size at most d. For each qubit $j \in [n]$ and time step $t \in [d]$, if there exists a gate in D_j at time t, then for all time steps $t' \leq t$, all gates supported on j in D_j also appear on j in U.

Definition 3.2. *The measurement outcome for observable O_j is denoted $m_j \in \{0, 1\}$.*

We use the convention that $m_j \in \{0, 1\}$ for ease of notation when multiplying several measurement outcomes later on, but $m_j = 0$ should be interpreted as a -1 eigenvalue outcome, and $m_j = 1$ interpreted as a $+1$ eigenvalue outcome.

4 Holographic quantum simulation

The holographic quantum simulator of [1, 2] is a conceptual quasi-1D system that simulates the interactions of a 2D system. Though one dimension of the 2D grid can be reduced significantly, the simulator still relies on 2D gates. The qubits of the simulation device can be divided into three categories: system, sink, and bath. In our holographic simulation, we only need two distinct subsystems: system and bath. The system at time t can be thought of as the t'th row of qubits in the 2D grid. The bath mediates the necessary correlation between the system qubits at different steps of the simulation. It was noted in [1] that the ratio of number of bath to system qubits should be proportional to the longest correlation length, and we consider this to be d in our simulator as an upper bound.

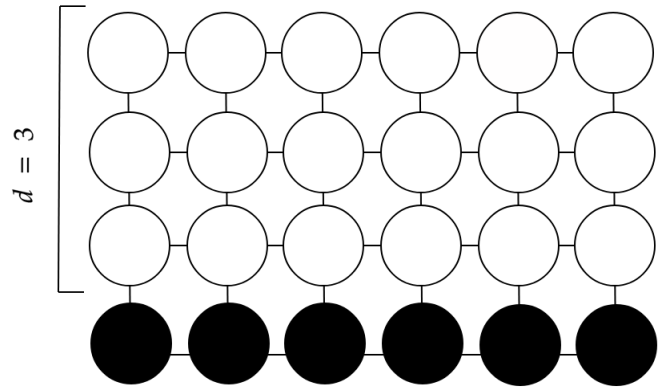

Fig. 1 *Structure of the physical system that implements the holographic quantum simulation with $d = 3$. The black vertices are the system qubits, and white vertices are the bath qubits. Two qubit gates are implemented over edges of the grid.*

The bath system can be thought of as an auxiliary system gliding over the rows of qubits in the grid. See Fig. 2. The $d \times \sqrt{n}$ auxiliary system works together with a single row of $\sqrt{n}$ qubits to simulate the circuit U and measure observables of the row.

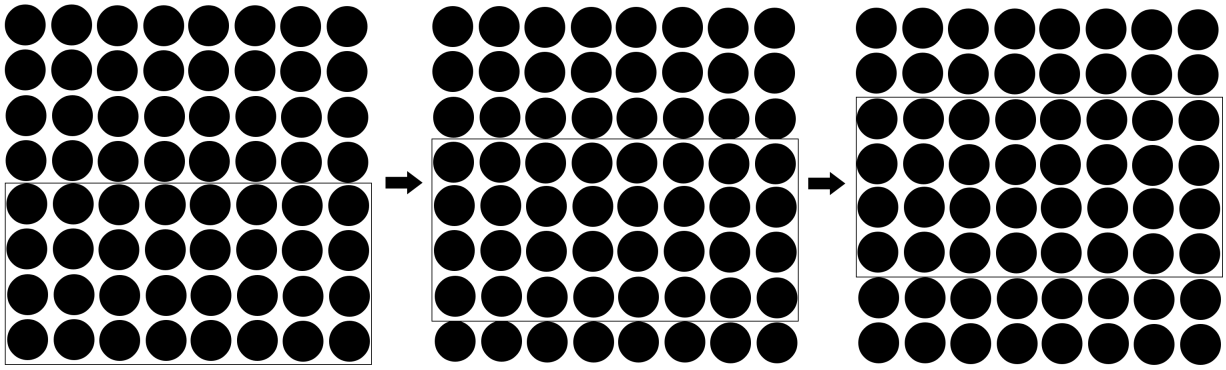

Fig. 2 *The smaller auxiliary system (box) glides over the qubits of the 2D grid with $d = 3$. The bottommost qubits in a box are the system, and the remaining qubits in the box are the bath.*

At the beginning of each time step, the system qubits are each initialized to the $|0\rangle$ basis state. Then a depth d circuit is applied to the bath-system composite to result in some desired state. To simulate the measurement of row t and column x of the 2D grid, we measure the x'th system qubit at the end of the time step. Note that although it is possible to defer all measurements until we are done processing all rows on a 2D grid computation, we can certainly measure once the auxiliary system and row of qubits is finished interacting. We formalize this idea later on, which becomes a key idea behind the holographic simulator. As with BQP mean value estimation, this simulation routine can be repeated a polynomial number of times to acquire an additive approximation to μ.

5 2D mean value approximation on a quasi-1D system

In this section, we show the algorithm for simulating the 2D grid mean value problem on a quasi-1D simulator.

Theorem 5.1. *Consider U a depth d circuit and measurement observable $O = O_1 \otimes ... \otimes O_n$ on a $\sqrt{n} \times \sqrt{n}$ 2D grid. There exists an algorithm with $O(d^2 n)$ time classical preprocessing, $(d+1)\sqrt{n}$ qubits, $O(\frac{d\sqrt{n}}{\delta^2})$ quantum depth, and $O(\frac{nd}{\delta^2})$ additional gates that δ-additively approximates μ with respect to O.*

We'll use the model described in section 4 as our quasi-1D system. We'll let the bath qubits in time step t simulate rows $t+1, ..., t+d$ to mediate long-range correlations. Recall the lightcone of a set of qubits S, D_S. We'll enact the operations in $D_{(\text{row t})}$ that haven't been simulated by previous time steps, then we'll measure the system qubits. We'll first show that for qubits of row t, it suffices to perform $D_{(\text{row t})}$ on the quasi-1D system and measure observables of the system qubits.

Lemma 5.2. *For $j \in [n]$, let $S_j \subseteq [n]$ be the set of qubits supporting D_j. Then the measurement outcome probability of observable O_j after circuit U can be simulated by D_j operating on S_j, followed by measuring observable $O_j \otimes I_{else}$.*

Proof. Define the outcome probability for the observable O_j by

$$p = \langle 0^n | U^\dagger (O_j \otimes I_{else}) U | 0^n \rangle$$

By the definition of D_j, we have that

$$D_j^\dagger (O_j \otimes I_{else}) D_j = U^\dagger (O_j \otimes I_{else}) U$$

implying that the probability can be rewritten as

$$p = \langle 0^n | D_j^\dagger (O_j \otimes I_{else}) D_j | 0^n \rangle$$

Furthermore $D_j^\dagger (O_j \otimes I_{else}) D_j$ where I_{else} is restricted to S_j is supported on S_j, so the probability becomes

$$p = \left\langle 0^{|S_j|} \middle| D_j^\dagger (O_j \otimes I_{else}) D_j \middle| 0^{|S_j|} \right\rangle$$

□

Recall that the S_j is centered at qubit j with radius upper bounded by d. This means that it suffices to maintain a simulator where system qubits interact with qubits of distance at most d. At first glance, it may seem that this requires $2d\sqrt{n}$ qubits in the simulator for rows above and below, but as long as system qubits simulate *all* interactions with qubits of distance d above, then we can simualte all interactions with a $d\sqrt{n}$ qubit simulator.

Note that all individual measurements commute with one another, so the order in which we measure each qubit does not matter, as long as we perform D_j beforehand. We use this to perform measurements starting from row 1 to $\sqrt{n}$. Measurement is the final step in the original 2D mean value problem, but in our simulator, we are interspersing measurements with unitary

operations on our simulator. This can be done because the measurement outcome of row t is only dependent on $D_{(\text{row t})}$ acting on the basis state followed by measuring the corresponding qubits. Because by the end of the t'th time step we've performed all operations in $D_{(\text{row t})}$, then we achieve the correct measurement distribution after measuring the correct observables.

To piece together different time steps, we need a reliable way to reuse qubits while maintaining the correct correlations in the bath qubits. For example, after we measure the simulated row 1, those same physical qubits must become the system for row 2 in the next time step while preserving the necessary correlations of row 2. To do this reliably, we reset the system qubits to $|0\rangle_{row}$ after measuring m_j for all j in the current row, and then we swap the row all the way to the top of the simulator with depth d swapping. This swapping adds $\sqrt{n} \cdot d$ extra gates per time step. See Fig. 3 for an illustration of the process. Once $t = \sqrt{n} - d$, we simply simulate the remaining interactions among the top d rows and measure the remaining qubits.

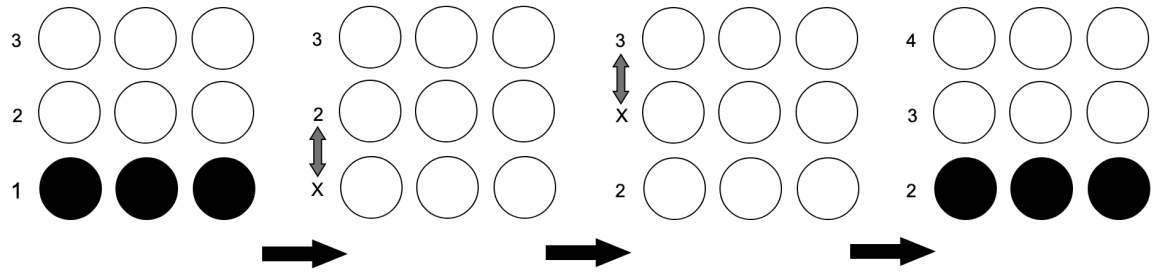

Fig. 3 *Subroutine for time step transition. Row 1 is measured, and reset to basis state "X". Row "X" is sequentially swapped to the top of the simulator until it becomes row 4. In the end, row 2 becomes the system.*

Now equipped with a process to measure the observables of each row for a given time step t and a process to piece together time steps in our simulator, we have an algorithm for one round of measurements, $m_1, ..., m_n$, with respect to O on the simulated 2D grid. Then the product $\prod_j m_j$ contributes to our estimation of the mean value. Each time step required a depth d circuit and depth d swapping, and with $\sqrt{n}$ time steps, we've simulated the 2D system in depth $O(d\sqrt{n})$.

We repeat this process S times, and we let X_i be the indicator variable for a positive measurement outcome after the i'th simulation. Clearly X_i are independent variables. Then define the empirical mean $\widetilde{\mu} = S^{-1} \sum_i X_i$. $\widetilde{\mu}$ has variance

$$\frac{\mu(1-\mu)}{S}$$

and expected value μ. Then choose the number of samples S to be $\frac{0.75}{\delta^2}$. By Chebychev inequality, $|\widetilde{\mu} - \mu| < \delta$ with probability greater than $\frac{2}{3}$. So we can estimate the mean value in quantum depth $O(\frac{d\sqrt{n}}{\delta^2})$.

To construct the necessary circuit on the simulator, we have to classically determine the subset of $D_{(\text{row t})}$ to implement for a given time step. For each row t of the 2D grid, we can search for the necessary subset of $D_{(\text{row t})}$ by recursively doing the following for the qubit in

column $c \in \{1,...,\sqrt{n}\}$ of the row: We begin by including every gate acting on qubit (t,c) that hasn't been considered yet. We then search from the beginning of the d'th layer of U. For the greatest $d' \leq d$ such that there is a two qubit gate at layer d' between qubit (t,c) and $(t+1,c)$, we include every gate on qubit $(t+1,c)$ prior to and including layer d' that hasn't been included yet. Then we repeat the same process substituting d' for d and qubit $(t+1,c)$ for (t,c). This process can be done in $O(d^2\sqrt{n})$ time for each row, and $O(d^2 n)$ time for the entire circuit. Thus we've constructed the simulating circuit on the quasi-1D system for all rows.

6 2D mean value approximation on 3D geometry without extra gates

The only additional gates we introduced when compared with the original U were the swaps necessary for the qubit-saving operations of the time step transition. We'll show how to sidestep this swapping by layering two quasi-1D systems together.

Theorem 6.1. *The mean value of a 2D grid can be additively approximated by $O(d^2 n)$ time classical preprocessing, $d\sqrt{n}$ qubits, $O(\frac{d\sqrt{n}}{\delta^2})$ depth, and no additional gates by a simulator on a $\sqrt{n} \times (\frac{d+1}{2}) \times 2$ grid.*

Proof. Assume that d is odd. If d is even then we pad the simulator with an additional row. Then the quasi-1D simulator has an even $d+1$ rows. Take the rows $\frac{d+1}{2}+1,...,d+1$, and "fold" them around the quasi-1D system such that rows $d+1$ and 1 are nearest neighbors. Note that row $d+1$ is a bath row and row 1 is the system. To transition time steps from t to $t+1$ after the system qubits have been measured and reset, row $t+1$ simply takes on the logical role of the system qubits and the reset row becomes the $t+1+d$ row. See Fig. 4 for an illustration of the folded simulator.

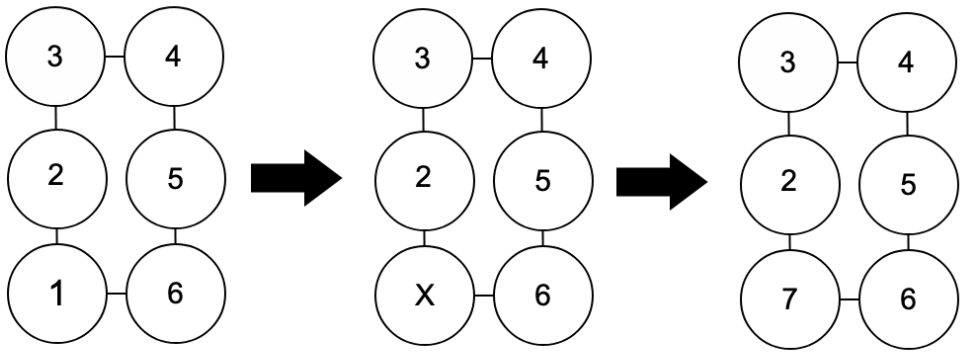

Fig. 4 *Side view of simulator. Row 6 is folded over as a nearest neighbor to row 1. To transition time steps, the system qubits are reset to "X" and they are replaced by the new row of the next time step.*

This process cycles around the folded system until the mean value computation concludes. We follow the same logical simulation techniques during each time step as the quasi-1D system. We also repeat the process $\frac{0.75}{\delta^2}$ times to additively approximate the mean value with high probability.

□

7 3D mean value approximation on a quasi-2D system

We extend the techniques of section 5 in a very natural way to obtain an efficient quasi-2D simulator of a 3D grid.

Theorem 7.1. *Consider U a depth d circuit and measurement observable $O = O_1 \otimes ... \otimes O_n$ on a $n^{1/3} \times n^{1/3} \times n^{1/3}$ 3D grid. There exists an algorithm with $O(d^2 n)$ time classical preprocessing, $(d+1)n^{2/3}$ qubits and $O(\frac{dn^{1/3}}{\delta^2})$ quantum depth that δ-additively approximates μ with respect to O.*

Proof. Let the simulator be a $n^{1/3} \times n^{1/3} \times (d+1)$ grid. For each time step $t \in [n^{1/3}]$, we let the system qubits simulate a 2D subgrid and the bath qubits simulate the forward correlations in the 3D grid G. This results in the system qubits simulating $G_{i,j,t}$ for $i, j \in [n^{1/3}]$ and the bath qubits simulating the qubits $G_{i,j,k}$ for $i, j \in [n^{1/3}]$ and $k \in \{t+1, ..., t+d\}$ at time step t. We do the same classical preprocessing to determine the lightcone of 2D subgrid t for every time step t, $D_{(\text{2D grid t})}$.

Then during the quantum computation, we operate the subcircuit of $D_{(\text{2D grid t})}$ to enact, determined by the classical preprocessor. At the end of the time step, we measure the 2D subgrid t and reset it just like on the quasi-1D simulator. Finally, we swap up the newly reset qubits and begin the next time step. We can repeat this process $\frac{0.75}{\delta^2}$ times and δ-additively approximate μ.

The classical preprocessing can be done in the same way as on the quasi-1D simulator by extending the lightcones to 2D grids instead of rows, and this is similarly done in time $O(d^2 n)$. Each time step is bounded by a depth d computation while swapping is also a depth d computation, and we repeat this for $n^{1/3}$ sets of system qubits. This results in $O(dn^{1/3})$ quantum depth per simulation run, and we approximate the mean value by repeating $O(\frac{1}{\delta^2})$ times for an overall $O(\frac{dn^{1/3}}{\delta^2})$ quantum depth computation. □

References

[1] Isaac H. Kim. Holographic quantum simulation, 2017; arXiv:1702.02093.

[2] Isaac H. Kim. Noise-resilient preparation of quantum many-body ground states, 2017; arXiv:1703.00032.

[3] Sergey Bravyi, David Gosset and Ramis Movassagh. Classical algorithms for quantum mean values, 2019; arXiv:1909.11485.

[4] Alberto Peruzzo and Jarrod McClean and Peter Shadbolt and Man-Hong Yung and Xiao-Qi Zhou and Peter J. Love and Alán Aspuru-Guzik and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor; Nat Commun 5, 4213 (2014).

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

[6] Maria Schuld and Nathan Killoran. Quantum machine learning in feature Hilbert spaces. Phys. Rev. Lett., 122(4):040504, 2019.

[7] Vojtech Havlicek, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, Jay M. Gambetta Supervised learning with quantum-enhanced feature spaces. Nature, 567(7747):209, 2019.

[8] John Napp, Rolando L. La Placa, Alexander M. Dalzell, Fernando G. S. L. Brandao and Aram W. Harrow. Efficient classical simulation of random shallow 2D quantum circuits, 2019; arXiv:2001.00021.

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

[10] Edwin Pednault, John A. Gunnels, Giacomo Nannicini, Lior Horesh, Thomas Magerlein, Edgar Solomonik, Erik W. Draeger, Eric T. Holland and Robert Wisnieff. Breaking the 49-Qubit Barrier in the Simulation of Quantum Circuits, 2017; arXiv:1710.05867.

[11] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy and Hartmut Neven. Simulation of low-depth quantum circuits as complex undirected graphical models, 2017; arXiv:1712.05384.

[12] Benjamin Villalonga, Dmitry Lyakh, Sergio Boixo, Hartmut Neven, Travis S. Humble, Rupak Biswas, Eleanor G. Rieffel, Alan Ho and Salvatore Mandrà. Establishing the Quantum Supremacy Frontier with a 281 Pflop/s Simulation, 2019; arXiv:1905.00444.