COMMUTING QUANTUM CIRCUITS: EFFICIENT CLASSICAL SIMULATIONS VERSUS HARDNESS RESULTS

XIAOTONG NI,[a] MAARTEN VAN DEN NEST[b]
Max-Planck-Institut für Quantenoptik
Hans-Kopfermann-Straße 1, D-85748 Garching, Germany.

The study of quantum circuits composed of commuting gates is particularly useful to understand the delicate boundary between quantum and classical computation. Indeed, while being a restricted class, commuting circuits exhibit genuine quantum effects such as entanglement. In this paper we show that the computational power of commuting circuits exhibits a surprisingly rich structure. First we show that every 2-local commuting circuit acting on d-level systems and followed by single-qudit measurements can be efficiently simulated classically with high accuracy. In contrast, we prove that such strong simulations are hard for 3-local circuits. Using sampling methods we further show that all commuting circuits composed of exponentiated Pauli operators $e^{i\theta P}$ can be simulated efficiently classically when followed by single-qubit measurements. Finally, we show that commuting circuits can efficiently simulate certain non-commutative processes, related in particular to constant-depth quantum circuits. This gives evidence that the power of commuting circuits goes beyond classical computation.

Keywords: Commuting quantum circuits, Classical simulation

Communicated by: R Jozsa & B Terhal

1 Introduction

Since the discovery of Shor's factoring algorithm [1], the question whether quantum computers possess exponentially more power then classical computers has been one of the central problems in the field. Similar to other notorious problems in computational complexity theory, this question is very difficult. For example, a proof that P $\neq$ BQP would imply that P $\neq$ PSPACE, which is a longstanding open problem. A useful approach to gain insight into the relationship between quantum and classical computing power is to study restricted classes of quantum circuits and analyze their power. For several restricted but nontrivial classes of quantum circuits, it has been found that efficient classical simulations are possible. For instance, if in each step of a quantum circuit the entanglement (quantified by the p-blockedness [2] or by the Schmidt rank [3]) is bounded, the circuit can be simulated efficiently classically. Such results demonstrate that certain types of entanglement must be generated in sufficiently large amounts if a quantum algorithm is to yield an exponential speed-up. Certain other circuit classes can be simulated classically using entirely different arguments not based

---

[a]xiaotong.ni@mpq.mpg.de
[b]maarten.vandennest@mpq.mpg.de

on entanglement considerations [4–9], e.g. by using the Pauli stabilizer formalism [4] or the framework of matchgate tensors [5–7] .

Conversely, it has been shown that some restricted quantum circuits can perform tasks that appear to be hard classically [10–13]. For example, in [10,11] it was shown that constant depth quantum circuits are hard to simulate exactly or with a high accuracy. Besides theoretical importance, these results also lower the threshold to demonstrate nontrivial quantum computation in experiments. Restricted quantum computation schemes other then the circuit model have also been studied recently, including linear optical quantum computation [14] and non-adaptive measurement based quantum computation [15].

In this paper we focus on commuting quantum circuits. Several features make such circuits interesting. For example, commuting circuits exhibit genuine quantum effects, e.g. they can generate highly entangled states (such as cluster states [16]). Further, since commuting operations can be performed simultaneously, there is no time order in the computation, which is drastically different from other computational models. Moreover, all gates in the circuit can be diagonalized simultaneously. The latter property might at first sight suggest an intrinsic simplicity of this circuit class; however it is important to note that the diagonalizing unitary can by a complex entangling operator. In [12] evidence is given that commuting circuits indeed have nontrivial power beyond classical computation: it was shown that simulating the output probability distribution of commuting quantum circuits would imply a collapse of the polynomial hierarchy and is thus highly unlikely.

In the present paper we map out the computational power of commuting circuits in more detail. Compared to earlier work [12, 17, 18] which considered commuting gates that can be diagonalized in a local basis, we will consider general commuting gates acting on d-level systems. We will show that the computational power of commuting circuits exhibits a surprisingly rich structure. For example, the degree of hardness varies significantly depending on whether the gates are 2-local or 3-local. This indicates that commuting quantum circuits might serve as a interesting intermediate class between classical and universal quantum circuits. It is worth noting that for the same reason commuting operations have recently caught attention in different areas as well, such as the study of local Hamiltonian problem [19–22].

Our main results can be summarized as follows (here the terms "strongly" and "weakly" specify different notions of classical simulation, to be defined below):

- *2-local circuits are easy.* All uniform families of commuting circuits consisting of 2-local gates acting on d-level systems and followed by a single measurement can be strongly simulated by classical computation, for every d.

- *3-local circuits are hard.* Uniform families of commuting circuits consisting of 3-local gates acting on qubit systems and followed by a single measurement cannot be strongly simulated by classical computation, unless every problem in #P has a polynomial-time classical algorithm.

- *Commuting Pauli circuits.* All uniform families of commuting circuits consisting of exponentiated Pauli operators $e^{i\theta P}$ and followed by a single-qubit measurement can be efficiently simulated classically weakly. Furthermore, even when such circuits display a small degree of non-commutativity, an efficient classical simulation remains possible.

- *Mapping non-commuting circuits to commuting circuits.* Certain non-commuting quantum processes (related to bounded-depth circuits) can be efficiently simulated by purely commuting quantum circuits.

Finally, it is noteworthy that several distinct techniques were used to prove the above results, including tensor network methods, sampling methods as well as the Pauli stabilizer formalism. This is also an illustration of the rich structure displayed by commuting quantum circuits.

2   Preliminaries

2.1   Commuting quantum circuits

The quantum circuits considered in this work will always be unitary. The size of a circuit is the number of gates of which it consists. A d-level Hilbert space will sometimes generically be called a "qudit". For an operator A which acts on a system of n qudits labeled by $1 \cdots n$, the support of A is the subset of qudits on which it acts nontrivially. A quantum circuit acting on n qudits is said to be k-local if the support of each of its gates contains at most k qudits. A family of n-qubit quantum circuits $\mathcal{C}_n$ has *polynomial size m* if m scales polynomially with n, denoted by $m = \mathrm{poly}(n)$.

A *commuting circuit* is a quantum circuit consisting of pairwise commuting gates. A k-local commuting circuit is in *standard form* if for every subset $S \subseteq \{1, \ldots, n\}$ consisting of k qudits there is at most one gate G_i with $\mathrm{supp}(G_i) \subseteq S$. A k-local commuting circuit $\mathcal{C} = G_m \cdots G_1$ in standard form contains at most $\binom{n}{k}$ gates, so that the size of the circuit scales polynomially with n if k is constant. For example, a two-local commuting circuit is in standard form if for every $i, j = 1 \cdots n$ with $i < j$ there is at most one gate in the circuit with support contained in $\{i, j\}$; such circuit has size $O(n^2)$. Every k-local commuting circuit can be brought into standard form by replacing all gates in the circuit with support contained in S by a single gate given by the total product of these gates, for every subset S consisting of k qudits. Furthermore if k is constant and if the original circuit has size m then this procedure to bring a circuit in normal form can be carried out efficiently i.e. in $\mathrm{poly}(n, m)$ steps.

A simple example of a commuting circuit is $\mathcal{C} = G_m \cdots G_1$ with gates

$$G_i = U \otimes U D_i U^\dagger \otimes U^\dagger, \tag{1}$$

where U is a fixed single-qudit unitary operator (independent of i) and where each D_i is a diagonal unitary operator. In other words each gate is diagonal in the same local basis. This class of commuting circuits has been considered in [12, 17].

By their commutativity, all gates in any commuting circuit can be diagonalized simultaneously i.e. there exists a unitary operator V such that $V G_i V^\dagger$ is diagonal for every gate G_i. In the example (1), the diagonalizing operator is a simple tensor product $V = U \otimes \cdots \otimes U$. This example does however not represent the most general situation since V may be a global, entangling operation—even when the commuting circuit is k-local with k constant. Consider for example an n-qubit 3-local circuit with gates $G_j = e^{i\theta_j K_j}$ where the commuting operators K_j are defined as follows:

$$K_1 = X_1 Z_2, \quad K_i = Z_{i-1} X_i Z_i, \quad K_n = Z_{n-1} X_n, \quad \text{with} \quad i = 2, \ldots n-1. \tag{2}$$

Here Z_i and X_i denote the Pauli X and Z operators acting on qubit i. The operators K_j are the well known stabilizers of the 1D cluster state. Let H denote the Hadamard gate and

let $CZ = \text{diag}(1, 1, 1, -1)$ denote the controlled-Z gate. It is then easily verified that the entangling operation

$$V = H^{\otimes n} \prod_{i=1}^{n-1} CZ_{i,i+1} \tag{3}$$

sends $K_j \rightarrow V K_j V^{\dagger} = Z_j$ and thus simultaneously diagonalizes the K_j. Furthermore it can be shown that no tensor product of single-qubit operations can perform such a diagonalization .

The example (2) shows that there exist k-local commuting circuits where the diagonalizing unitary V is a global, entangling operator. Nevertheless this example is still rather well-behaved as V can be computed efficiently and moreover has a relatively simple structure. In fact in section 5 we will investigate commuting circuits composed of exponentiated Pauli operators $e^{i\theta P}$ in more detail and show that such circuits have efficient classical simulations (relative to certain measurements). For general k-local commuting circuits, however, the unitary V may have a complex structure and be computationally difficult to determine. This feature is in part responsible for the complexity of commuting quantum circuits.

2.2 *Classical simulations of quantum circuits*

There are several valid notions of efficient classical simulations of quantum circuits. Two notions will be considered in this work viz. *strong* and *weak* simulations. Their main difference lies in the *accuracy* achieved in the classical simulation: roughly speaking, strong simulations achieve an exponential precision whereas weak simulations achieve polynomial precision. We mainly follow the definitions of [12].

Consider a uniform family of k-local n-qubit quantum circuits $\mathcal{C}_n$ for some constant k. The input states are standard basis states. The circuits are followed by measurement of the Pauli observable Z on the first qubit. The expectation value is denoted by $\langle Z_1 \rangle$.

Strong simulations. We say that $\mathcal{C}_n$ can be efficiently simulated classically in the strong sense if there exists a classical algorithm with runtime $\text{poly}(n, \log \frac{1}{\epsilon})$ which outputs a number E such that

$$|E - \langle Z_1 \rangle| \leq \epsilon. \tag{4}$$

Thus a strong simulation algorithm achieves an exponential accuracy $\epsilon = 2^{-\text{poly}(n)}$ in $\text{poly}(n)$ time.

Weak simulations. We say that $\mathcal{C}_n$ can be efficiently simulated classically in the weak sense if there exists a classical algorithm with runtime $\text{poly}(n, \frac{1}{\epsilon})$ which outputs a number E satisfying (4). Thus a weak simulation algorithm achieves *polynomial* accuracy $\epsilon = 1/\text{poly}(n)$ in polynomial time. We will often allow weak simulations to fail with an exponentially small probability. In this sense, we say that $\mathcal{C}_n$ can be efficiently simulated classically in the weak sense if there exists a probabilistic classical algorithm with runtime $\text{poly}(n, \frac{1}{\epsilon}, \log \frac{1}{1-p})$ which outputs a number E satisfying (4) with probability p. Thus for polynomial accuracies and for success probabilities which are exponentially (in n) close to 1, the classical simulation runs in $\text{poly}(n)$ time.

The motivation for the definition of a weak simulation originates from the fact the polynomial error scaling $\epsilon = 1/\mathrm{poly}(n)$ captures how accurately the expectation value $\langle Z_1 \rangle$ can be estimated by running the quantum circuit $\mathcal{C}_n$ polynomially many times. See section 2.3 below and [8] for a more extensive discussion.

The above definitions can readily be generalized to take into account more general inputs (e.g. arbitrary product states) and measurements (e.g. arbitrary single-qubit observables) as well as d-level systems. Finally, note that we will often use the term "simulation" as shorthand for "efficient classical simulation".

2.3   Chernoff-Hoeffding bound

The Chernoff-Hoeffding bound is a tool to bound how accurately the expectation value of a random variable may be approximated using of "sample averages". Let $X_1, \ldots X_K$ be i.i.d. real-valued random variables with $E := \mathbb{E}X_i$ and $X_i \in [-1, 1]$ for every $i = 1, \ldots, K$. Then the Chernoff-Hoeffding bound asserts that

$$\mathrm{Prob}\left\{\left|\frac{1}{K}\sum_{i=1}^{K} X_i - E\right| \le \epsilon\right\} \ge 1 - 2e^{-\frac{K\epsilon^2}{4}}. \tag{5}$$

For complex-valued X_i a similar bound can be obtained for $|X_i| \le 1$.

As an illustration, consider an n-qubit quantum circuit family $\mathcal{C}_n$ followed by measurement of Z_1 as in section 2.2. Suppose that the circuit is run K times, yielding an outcome $z_i \in \{1, -1\}$ in each run. Using (5) one shows that the number $E := [\sum z_i]/K$, where the sum is over all $i = 1 \cdots K$, satisfies $|E - \langle Z_1 \rangle| \le \epsilon$ with probability $p \ge 1 - 2e^{-K\epsilon^2/4}$. Consequently, for any $\epsilon = 1/\mathrm{poly}(n)$ there exists a suitable $K = \mathrm{poly}(n)$ such that $|E - \langle Z_1 \rangle| \le \epsilon$ holds with probability p exponentially close to 1. In other words, the above procedure allows to achieve a polynomial approximation of $\langle Z_1 \rangle$ in polynomial time with exponentially small probability of failure. This performance of the quantum computation corresponds precisely to the performance required of weak classical simulations, cf. section 2.2.

3   2-Local commuting circuits are easy

Here we consider 2-local commuting circuits acting on general d-level systems. The main conclusion of this section will be that such circuits, when followed by single-qudit measurements, cannot outperform classical computation. In fact we will show that their power is even *strictly* contained in P and give a concrete example of a simple function which cannot be computed with such commuting circuits.

3.1   Efficient strong simulation of one qudit

Theorem 1 (Strong simulations of 2-local commuting circuits) *Let $\mathcal{C}$ be a uniform family of 2-local n-qudit commuting circuits, acting on a product input state and followed by measurement of an observable O acting on qudit i for some i. Any such computation can be efficiently simulated classically in the strong sense.*

Proof.   We prove the result for $i = 1$; other i are treated fully analogously. Denote the input by $|\alpha\rangle = |\alpha_1\rangle \cdots |\alpha_n\rangle$ where each $|\alpha_i\rangle$ is a single-qudit state. We can assume without loss of generality that $\mathcal{C} = \prod U_{jk}$ is in standard form, where U_{jk} represents the unique gate

in the circuit with support $S \subseteq \{j, k\}$, for every $j, k = 1 \cdots n$ and $j < k$. If U_{jk} does not act on qudit 1, then this gate commutes with O. Hence in the product $\mathcal{C}^\dagger O \mathcal{C}$ we can commute U_{jk} through $\mathcal{C}$ and O to the left until it cancels out with $U_{jk}^\dagger$. By doing so, we can remove all gates that do not act on qudit 1. Therefore the expectation value of O is

$$\langle O \rangle = \langle \alpha | \mathcal{C}^\dagger O \mathcal{C} | \alpha \rangle = \langle \alpha | (\prod U_{1j})^\dagger O \prod U_{1j} | \alpha \rangle \tag{6}$$

where the products are over all $j \geq 2$. Now our strategy will be to trace out qudits one by one in the above equation. Denote $|\alpha^{(1)}\rangle = |\alpha\rangle$, $\mathcal{C}^{(1)} = \mathcal{C}$ and $O^{(1)} = O$. Furthermore for every $k = 2, \ldots, n-1$ define

$$
\begin{aligned}
|\alpha^{(k)}\rangle &= |\alpha_1\rangle |\alpha_{k+1}\rangle \cdots |\alpha_n\rangle \\
\mathcal{C}^{(k)} &= U_{1k+1} \cdots U_{1n} \\
O^{(k)} &= [I \otimes \langle \alpha_k |] \, U_{1k}^\dagger O^{(k-1)} U_{1k} \, [I \otimes |\alpha_k\rangle].
\end{aligned}
\tag{7}
$$

Remark that each $O^{(k)}$ acts on a single qudit (namely qudit 1). Furthermore each of these operators can be computed classically with exponential precision in polynomial time: $O^{(1)}$ is given as an input and each update from $O^{(j)}$ to $O^{(j+1)}$ involves simple multiplications of 2-qudit operations which can be done in constant time (taking $O(d^6)$ steps where d denotes the dimension of one qudit).

With the above definitions one finds, for every $k = 2, \ldots, n-1$:

$$\langle \alpha^{(k-1)} | [\mathcal{C}^{(k-1)}]^\dagger O^{(k-1)} \mathcal{C}^{(k-1)} | \alpha^{(k-1)} \rangle = \langle \alpha^{(k)} | [\mathcal{C}^{(k)}]^\dagger O^{(k)} \mathcal{C}^{(k)} | \alpha^{(k)} \rangle. \tag{8}$$

Using this equation iteratively, we get

$$\langle O \rangle = \langle \alpha^{(1)} | [\mathcal{C}^{(1)}]^\dagger O^{(1)} \mathcal{C}^{(1)} | \alpha^{(1)} \rangle = \cdots = \langle \alpha^{(n-1)} | U_{1n}^\dagger O^{(n-1)} U_{1n} | \alpha^{(n)} \rangle. \tag{9}$$

The last expression is easily computed since $|\alpha^{(n-1)}\rangle$ is a 2-qudit state and U_{1n} and $O^{(n-1)}$ act on at most 2 qudits. $\square$

The above result can readily be generalized in different ways. First, using a similar argument one shows that measurement of any observable acting on $O(\log n)$ qudits can be strongly simulated as well. Furthermore, interestingly, the result also generalizes to mutually *anticommuting* gates, and more generally to gates which commute "up to a phase" as follows. Let $\mathcal{C} = \prod G_i$ be a uniform family of 2-local n-qudit circuits such that $G_i G_j = \gamma_{ij} G_j G_i$ for all pairs of gates, where the γ_{ij} are complex phases. Input and measurement are as in theorem 1. Then such circuits can be efficiently simulated classically in the strong sense. Analogous to the first step in the proof of theorem 1, the proof starts by "removing" all gates which do not act on qudit i from the product $\mathcal{C}^\dagger O \mathcal{C}$ by commuting them through the circuit. This introduces an (easily computed) product of phases γ_{ij}. The remainder of the proof of theorem 1 carries over straightforwardly.

3.2 *2-local commuting circuits cannot compute all function families in* P

Here we show that two-local commuting circuits are not universal for classical computation by giving an explicit example of a function which is not computable with such circuits.

For every d we let $\mathbf{Z}_d$ denote the set of integers modulo d. Let $\mathcal{C}$ denote a two-local commuting circuit acting on m d-level systems. Consider a function $f : \mathbf{Z}_d^k \to \mathbf{Z}_d$. We say

that $\mathcal{C}$ computes f with probability at least p if the circuit $\mathcal{C}$ acing on $|x, 0\rangle$ (where 0 denotes a string of $m - k$ zeroes) and followed by a standard basis measurement of the first qudit yields the outcome $f(x)$ with probability at least p.

We will in particular consider the "inner product function" $f_{\text{ip}} : \mathbf{Z}_d^{2n} \to \mathbf{Z}_d$ defined by

$$f_{\text{ip}}(x^a, x^b) = (x^a)^T x^b \mod d, \tag{10}$$

for every $x^a, x^b \in \mathbf{Z}_d^n$.

Lemma 1 *Let $\sigma_1, \ldots \sigma_N$ be a collection of $d \times d$ density operators. For any $\epsilon > 0$, if $N > \left(\frac{5}{\epsilon}\right)^{2d^2}$, then there exists two operators σ_j and σ_k such that $\|\sigma_j - \sigma_k\|_{tr} \leq \epsilon$, where $\|A\|_{tr} \equiv \frac{1}{2} tr \sqrt{A^\dagger A}$ denotes the trace distance.*

Proof. We will show for any $\epsilon > 0$, there exists a finite set $\mathbf{E}$ of $d \times d$ density operators, such that for every density operator ρ, there exists $\sigma \in \mathbf{E}$ with $\|\rho - \sigma\|_{\text{tr}} < \epsilon$ (we call $\mathbf{E}$ a ϵ-net). To do this, first we recall that every density operator of dimension d has a purification by introducing an ancillary d-dimensional space R . And in [23], it was shown that for pure states of dimension d^2, there exists a ϵ-net $\mathbf{F}$ with cardinality $|\mathbf{F}| \leq \left(\frac{5}{2\epsilon}\right)^{2d^2} \equiv M$. We can then choose set $\mathbf{E}$ to be $tr_R \mathbf{F}$, which is the partial trace of each element of set $\mathbf{F}$. Since partial trace is a contractive operation [24], i.e. $\|tr_R(\mu - \tau)\|_{\text{tr}} \leq \|\mu - \tau\|_{\text{tr}}$, we know that set $\mathbf{E}$ obtained this way is indeed an ϵ-net.

Note that $|\mathbf{E}| \leq |\mathbf{F}| = M$. Now if there are more then M density operators, then there must be two density operators σ_j, σ_k that are ϵ-close to the same element of $\mathbf{E}$. Thus by triangle inequality $\|\sigma_j - \sigma_k\|_{\text{tr}} < 2\epsilon$. The proof can be finished by a rescaling of ϵ. $\square$

Theorem 2 *Consider an arbitrary d and an arbitrary constant $p > 1/2$. For sufficiently large n, the inner product function f_{ip} is not computable by any two-local commuting circuit.*

Proof. Suppose there exists an m-qudit quantum circuit $\mathcal{C}$, for some $m \geq 2n$, which computes f with probability $p > 1/2$. We show that this leads to a contradiction. Repeating the argument of theorem 1 we can remove all gates from the circuit which do not act on qudit 1. We denote this simplified circuit again by $\mathcal{C}$. Now write $\mathcal{C} = \mathcal{C}_b \mathcal{C}_a$, where $\mathcal{C}_a$ consists of all gates in the circuit acting on qudits $\{1, i\}$ with $i = 1 \ldots n$ and where $\mathcal{C}_b$ consists of all gates acting on qudits $\{1, j\}$ with $i = n + 1 \ldots m$. Furthermore, let $x = (x^a, x^b)$ be an arbitrary input of f. Finally, denote

$$\sigma(x^a) := \text{Tr}_{n \ldots 2} \, \mathcal{C}_a |x^a\rangle\langle x^a| \mathcal{C}_a^\dagger, \tag{11}$$

which is the reduced density operator for qudit 1 of the state $\mathcal{C}_a |x^a\rangle$.

The final state of the entire circuit is $\mathcal{C}|x, 0\rangle$ where 0 denotes a string of $m - n$ zeroes. With the notations above, the reduced density operator of the first qudit is

$$\begin{aligned} \rho(x^a, x^b) &:= \text{Tr}_{m \ldots 2} \, \mathcal{C}|x, 0\rangle\langle x, 0|\mathcal{C}^\dagger = \text{Tr}_{m \ldots n+1} \, \text{Tr}_{n \ldots 2} \, \mathcal{C}_b \mathcal{C}_a |x, 0\rangle\langle x, 0|\mathcal{C}_a^\dagger \mathcal{C}_b^\dagger \\ &= \text{Tr}_{m \ldots n+1} \, \mathcal{C}_b \left\{ \sigma(x^a) \otimes |x^b, 0\rangle\langle x^b, 0| \right\} \mathcal{C}_b^\dagger \end{aligned} \tag{12}$$

We now use lemma 1. This implies for every $\epsilon > 0$ there exists an n sufficiently large and two n-tuples $x^a \neq y^a$ such that $\|\sigma(x^a) - \sigma(y^a)\|_{\text{tr}} \leq \epsilon$. Using (12) and the fact that the trace norm is contractive, it follows that $\|\rho(x^a, x^b) - \sigma(y^a, x^b)\|_{\text{tr}} \leq \epsilon$ for *every* n-tuple x^b! This implies the following: if a standard basis measurement on $\rho(x^a, x^b)$ yield some outcome u with

probability $p(u)$, then standard basis measurement on $\rho(y^a, x^b)$ will yield the same outcome with probability $q(u)$ where $|p(u) - q(u)| \leq \epsilon$. Setting $\epsilon = p - \frac{1}{2}$ and using that $\mathcal{C}$ computes f with probability at least p, it then follows that $f(x^a, x^b) = f(y^a, x^b)$ for all x^b. Using the definition of f, this straightforwardly implies that $x^a = y^a$, thus leading to a contradiction. $\square$.

4 3-Local commuting circuits are hard

Next we show that strong simulations of 3-local commuting circuits are unlikely to exist.

Theorem 3 (Hardness of simulating 3-local commuting circuits) *Let $\mathcal{C}$ be a uniform family of n-qubit 3-local commuting quantum circuits acting on the input $|0\rangle$ and followed by Z measurement of the first qubit. If all such circuits could be efficiently simulated classically in the strong sense then every problem in #P has a polynomial time algorithm.*

In other words, there is a drastic increase in complexity in the seemingly innocuous transition from 2-local to 3-local gates. Remark that hardness already holds for the simplest case i.e. qubit systems—even though d-level 2-local commuting circuits have efficient simulations for any d. Hardness of strong simulations does not necessarily imply that weak simulations are hard as well since strong and weak simulations are generally inequivalent concepts (cf. [8] for a discussion). In section 6 we will provide evidence that k-local commuting circuits with constant k can efficiently perform certain tasks that appear to be nontrivial for classical computers, thereby providing evidence that efficient weak simulations might not exist in general.

The proof of theorem 3 is given below. Our approach is to relate simulations of 3-local commuting circuits to the evaluation of matrix elements of *universal* unitary quantum circuits, which is known to be hard. The following three lemmata collect preliminary results. First we recall that the evaluation of matrix elements of universal quantum circuits is known to be hard. We denote $S := \text{diag}(1, e^{i\pi/4})$ and $CZ := \text{diag}(1, 1, 1, -1)$.

Lemma 2 *Let $\mathcal{U}$ be a uniform family of n-qubit quantum circuits composed of the gates H, S and CZ. If there existed an algorithm with runtime $poly(n, \log \frac{1}{\epsilon})$ which outputs an ϵ-approximation of $\langle 0|\mathcal{U}|0\rangle$ for any such circuit family, then every problem in #P has a polynomial-time algorithm.*

Proof. Consider an efficiently computable Boolean function $f : \{0, 1\}^n \to \{0, 1\}$. Let $s(f)$ denote the number of bit strings x satisfying $f(x) = 0$. The problem of computing $s(f)$ is well known to #P-complete. Now define the $(n + 1)$-qubit state $|f\rangle := 2^{-n/2} \sum_x |x, f(x)\rangle$ where the sum is over all n-bit strings x. Let $\mathcal{H}$ be the operator which acts as H on qubits 1 to n and as the identity on qubit $n + 1$. Then an easy calculation shows

$$\langle 0|\mathcal{H}|f\rangle = s(f)/2^n. \tag{13}$$

Since H, CZ and S form a universal gate set, the Solovay-Kitaev theorem implies that there exists a uniform circuit family $\mathcal{V}$ composed of these gates such that $\mathcal{V}|0\rangle$ is δ-close to $|f\rangle$ with $\delta := 2^{-n^2}$. Denote the circuit $\mathcal{U} := \mathcal{H}\mathcal{V}$. Using (13) it follows that

$$|\langle 0|\mathcal{U}|0\rangle - \frac{s(f)}{2^n}| \leq \delta. \tag{14}$$
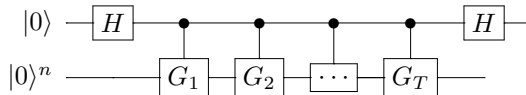
Fig. 1. The Hadamard test

Now suppose that there exists a $\text{poly}(n, \log \frac{1}{\epsilon})$ classical algorithm to compute $\langle 0|\mathcal{U}|0\rangle$ with accuracy ϵ. Setting $\epsilon = \delta$, this would imply the existence of a polynomial time classical algorithm that outputs an δ-approximation γ of $\langle 0|\mathcal{U}|0\rangle$. Using (14) and the triangle inequality this implies that γ approximates $s(f)/2^n$ with accuracy 2δ. Since $s(f)/2^n = k/2^n$ for some integer between 0 and 2^n, this accuracy would allow to compute $s(f)$ exactly in polynomial time, hence implying that every problem in #P has a poly-time algorithm. □.

Second, we recall a result from [12] which relates universal quantum circuits to postselected 2-local commuting circuits.

Lemma 3 *Let $\mathcal{U}$ be an n-qubit quantum circuit composed of the gates H, S and CZ and denote $|\psi\rangle = \mathcal{U}|0\rangle^n$. Then there exists a 2-local commuting circuit $\mathcal{C}$ on $k+n$ qubits such that $|\psi\rangle$ is obtained by postselecting $\mathcal{C}|0\rangle^{k+n}$ on the first k qubits; more precisely*

$$|0\rangle^k|\psi\rangle = \sqrt{2}^k \mathcal{P}\mathcal{C}|0\rangle^{k+n}. \tag{15}$$

Here $\mathcal{P}$ denotes the projector $|0\rangle\langle 0|$ acting on the first k qubits. Furthermore $k = poly(n)$ and the description of $\mathcal{C}$ can be computed efficiently on input of the description of $\mathcal{U}$.

Combining the above two lemmata shows that approximating matrix elements of commuting 2-local circuits is hard.

Lemma 4 *Let $\mathcal{C}$ be a uniform family of n-qubit 2-local commuting quantum circuits. If there existed a classical algorithm with runtime $poly(n, \log \frac{1}{\epsilon})$ which outputs an ϵ-approximation of $\langle 0|\mathcal{C}|0\rangle$ for any such $\mathcal{C}$, then every problem in #P has a poly-time algorithm.*

Proof. Let $\mathcal{U}$ be a uniform family of n-qubit quantum circuits composed of the gates H, S and CZ and let $\mathcal{C}$ be the associated commuting circuit family as in lemma 3. Using (15) one finds

$$\langle 0|^n \mathcal{U}|0\rangle^n = \sqrt{2}^k \langle 0|^{n+k} \mathcal{C}|0\rangle^{n+k}. \tag{16}$$

If an efficient classical algorithm existed to estimate $\langle 0|\mathcal{C}|0\rangle$ with exponential precision, then there also exists an algorithm to estimate $\langle 0|\mathcal{U}|0\rangle$ with exponential precision. This implies that every problem in #P has a poly-time algorithm owing to lemma 2. □.

The proof of theorem 3 now proceeds by relating the simulation of 3-local commuting circuits to the evaluation of matrix elements of 2-local commuting circuits, via the Hadamard test.

Proof of theorem 3: Suppose that an efficient algorithm existed to strongly simulate the circuits described in the theorem. Consider an arbitrary n-qubit commuting circuit $\mathcal{C} = G_T \cdots G_1$ with two-qubit gates G_i. Consider the following $(n+1)$-qubit quantum circuit (the "Hadamard test") with input $|0\rangle$ as depicted in Fig. 1. First H is applied to the first qubit. Then each gate G_i is applied controlled on the first qubit being in the state $|1\rangle$; we denote these 3-qubit gates by CG_i. Finally, H is again applied to the first qubit. Measuring the first

qubit yields the outcome 0 with probability

$$p(0) = \frac{1}{2}(1 + \text{Re}(\langle 0|\mathcal{C}|0\rangle)). \tag{17}$$

Now for each i define the 3-qubit gate $U_i := [H \otimes I]CG_i[H \otimes I]$, where H acts on the first qubit, and let $\mathcal{C}'$ denote the circuit composed of the gates U_i. Since the gates G_i commute, also the gates U_i commute. Furthermore, it is straightforward to show that the circuit $\mathcal{C}'$ acting on $|0\rangle$ and followed by measurement of the first qubit is equivalent to the circuit in Fig. 1, since the hadamard operations "in the middle" cancel out. Thus $\mathcal{C}'$ also yields the outcome 0 with probability $p(0)$. It follows that the existence of an efficient classical algorithm to strongly simulate the circuit $\mathcal{C}'$ yields an efficient classical algorithm to compute the real part of $\langle 0|\mathcal{C}|0\rangle$ with exponential precision. Adding $P = \text{diag}(1, i)$ before the second Hadamard gate in Fig. 1 and arguing analogously yields an efficient algorithm to estimate the imaginary part of $\langle 0|\mathcal{C}|0\rangle$ with exponential precision. Using lemma 3 we conclude that this would imply that every problem in $\#P$ has a poly-time algorithm. $\square$

5   Efficient simulation of commuting Pauli Circuits

A circuit composed of unitary operators of the form $e^{i\theta P}$, where the Ps are (Hermitian) Pauli operators, is called a Pauli circuit. By criterion in [25], Pauli circuits can be shown to be universal for quantum computation. Here we investigate *commuting* Pauli circuits. We allow P to act on arbitrarily many qubits i.e. we do not restrict to local gates[c].

Given the distinguished status of Pauli operators, commuting Pauli circuits constitute a simple and natural class of commuting quantum circuits. This class in fact encompasses the model of "instantaneous quantum computation" (IQP) introduced in [17]. IQP corresponds to the subclass of commuting Pauli circuits where each P is restricted to be a tensor product of identities and Pauli X matrices, so that every gate $e^{i\theta P}$ is diagonalized by the tensor product operator $H \otimes \cdots \otimes H$. Generalizing IQP to arbitrary commuting Pauli circuits adds the interesting feature that the unitary operator which simultaneously diagonalizes the gates in the circuit is generally no longer a tensor product of single-qubit operators, but rather a global entangling operation; see example (2).

Whereas arbitrary Pauli circuits are universal, we will show that commuting Pauli circuits can be efficiently simulated classically in the following sense.

Theorem 4 (Weak simulation of Commuting Pauli circuits) *Every uniform family of commuting Pauli circuits acting on a standard basis input and followed by measurement of Z acting on one of the qubits can be weakly simulated classically.*

It was shown in [12] that IQP circuits followed by single-qubit standard basis measurements can be simulated efficiently weakly[d]. Theorem 4 hence generalizes this result to arbitrary commuting Pauli circuits. Furthermore, in [12] it was shown that efficient weak classical simulation (relative to a certain special type of approximations viz. multiplicative approximations) of 2-local IQP circuits followed by $O(n)$ computational basis measurements are highly unlikely to exist: the existence of such simulations would imply a collapse of the polynomial hierarchy

---

[c]Remark that, even for such non-local gates, every gate $e^{i\theta P}$ can be efficiently implemented on a quantum computer i.e. it can be realized by a polynomial size quantum circuit of elementary gates

[d]In fact classical simulations were also achieved in [12] for $O(\log n)$ measurements; our results can also be generalized to simulate such measurements for arbitrary commuting Pauli circuits.

to its third level. Thus a fortiori simulations of $O(n)$ computational basis measurements are unlikely to exist for general commuting Pauli circuits as well.

One can in fact show a stronger version of theorem 4: a general Pauli circuit containing a limited degree of non-commutativity can still be simulated classically efficiently.

Theorem 5 (Weak simulation of slightly non-commuting Pauli circuits) *Consider a uniform family of n-qubit commuting Pauli circuits interspersed with $O(\log n)$ gates of the form $e^{i\theta Q}$ with Q an arbitrary (Hermitian) Pauli operator. Any such circuit family acting on standard basis input and followed by measurement of Z acting on one of the qubits can be weakly simulated classically.*

The proofs of theorem 4 and 5 are given in section 5.3. In the preceding sections we develop the necessary tools. It is interesting that the simulation techniques used here are completely different from those used our simulations of 2-local commuting circuits (theorem 1). In particular the latter involved strong simulations whereas commuting Pauli circuits will be simulated using weak simulations combined with stabilizer methods.

5.1   Pauli and Clifford operators

A Pauli operator on n qubits has the form $P = \alpha P_1 \otimes \ldots \otimes P_n$, where $\alpha \in \{\pm 1, \pm i\}$ and where each P_j is one of the Pauli matrices X, Y, Z or the identity. A Pauli operator is said to be of Z-type if each P_j is either Z or the identity; X-type Pauli operators are defined analogously. Since X, Y and Z are Hermitian, a Pauli operator is Hermitian if and only if $\alpha \in \{1, -1\}$. Letting Z_k and X_k denote the operators Z and X acting on qubit k, respectively, it can be verified that every Pauli operator P can be written as

$$P = i^t \prod_k X_k^{a_k} Z_k^{b_k}, \quad \text{where } t \in \{0, 1, 2, 3\}, \ a_k, b_k \in \{0, 1\}. \tag{18}$$

Defining the $2n$-dimensional bit string

$$r(P) = (a_1, \cdots, a_n, b_1, \cdots, b_n), \tag{19}$$

it is easily verified that $r(PQ) = r(P) + r(Q)$ for all Pauli operators P and Q, where addition is modulo 2.

An n-qubit operator U is a Clifford operation if $UPU^\dagger$ is a Pauli operator for every Pauli operator P. The set of all n-qubit Clifford operations is a group, called the Clifford group. A Clifford circuit is a quantum circuit composed of H, $CNOT$ and $P = \text{diag}(1, i)$. It is well known that every Clifford circuit realizes a Clifford operator, and that every Clifford operator can be realized as a (polynomial-size) Clifford circuit.

Lemma 5 *Let $P_1, \ldots, P_m$ be a collection of commuting n-qubit Pauli operators. Then there exists a Clifford operation $\mathcal{C}$ such that $\mathcal{C}^\dagger P_i \mathcal{C} = Q_i$ for every i, where each Q_i is a Z-type Pauli operator. Moreover each Q_j as well as the description of a poly-size Clifford circuit realizing $\mathcal{C}$ can be determined efficiently.*

Proof. It suffices to prove the result for Hermitian Pauli operators since every Pauli operator can be made Hermitian by providing it with a suitable overall phase. Thus henceforth we assume that the P_i are Hermitian. We can write all m vectors $r(P_i)$ in a $m \times 2n$ matrix and pick out a maximal set of independent row vectors over $\mathbb{Z}_2$ efficiently by Gaussian elimination. W.l.o.g. we assume these are the first l vectors. The corresponding Pauli operators

$\{P_1, \ldots, P_l\} =: \mathcal{S}$ form an independent set i.e. no operator in $\mathcal{S}$ can be written as a product of the other elements of $\mathcal{S}$. In addition, no product of operators in $\mathcal{S}$ yields $-I$. Indeed suppose there exist bits x_j, not all zero, such that $P_1^{x_1} \ldots P_l^{x_l} = -I$. This would imply that $\sum x_j r(P_j) = 0$, contradicting with the linear independence of the $r(P_j)$. Since the operators in $\mathcal{S}$ are Hermitian, independent and commuting and since no product of some of these operators yields $-I$, there exists a stabilizer code $\mathcal{V}$ of dimension 2^{n-l} stabilized by $\mathcal{S}$ [24]. This implies in particular that $l \leq n$. Using standard stabilizer techniques one can efficiently compute additional Hermitian Pauli operators $\mathcal{S}' = \{R_{l+1}, \ldots, R_n\}$ such that all operators in the set $\mathcal{T} = \mathcal{S} \cup \mathcal{S}'$ mutually commute, are independent and no product of these operators yields $-I$ [24]. These n operators are the stabilizers of a 1-dimensional stabilizer code i.e. a stabilizer state $|\psi\rangle$. In other words $|\psi\rangle$ satisfies $P_i|\psi\rangle = |\psi\rangle = R_j|\psi\rangle$ for every $i = 1, \ldots, l$ and $j = l+1, \ldots, n$, and moreover it is the unique state doing so. It is well known that there exists a poly-size n-qubit Clifford circuit $\mathcal{C}$ such that $|\psi\rangle = \gamma \mathcal{C}|0\rangle^n$ for some global phase γ; moreover a description of $\mathcal{C}$ can be computed efficiently [26]. Now define $Q_i = \mathcal{C}^\dagger P_i \mathcal{C}$ for every $i = 1, \ldots, m$. Each Q_i is an efficiently computable Pauli operator since $\mathcal{C}$ is a poly-size Clifford circuit. Since $\mathcal{C}|0\rangle = |\psi\rangle$ and $P_j|\psi\rangle = |\psi\rangle$ for every $P_j \in \mathcal{S}$ one has $Q_j|0\rangle = |0\rangle$. This last property together with the fact that each Q_j is a Pauli operator implies that Q_j must be of Z-type. Finally, since each P_k with $k \geq l+1$ can be written, up to a global phase, as a product of operators within $\mathcal{S}$ and since products of Z-type Pauli operators are again of Z-type, it follows that also Q_k is of Z-type. $\square$

5.2 CT states

Here we recall a result (theorem 6 below) stating that a general class of quantum processes can be simulated weakly. First we need some definitions. Consider a family of n-qubit states $|\psi_n\rangle \equiv |\psi\rangle$ specified in terms of some classical description, say a quantum circuit preparing $|\psi\rangle$ from the state $|0\rangle$. Following [8], $|\psi\rangle$ is said to be *computationally tractable (CT)* (relative to this description) if

(a) it is possible to sample in poly(n) time with classical means from the probability distribution $\mathrm{Prob}(x) = |\langle x|\psi\rangle|^2$ on the set of n-bit strings x, and

(b) for any bit string x, the coefficient $\langle x|\psi\rangle$ can be computed in poly(n) time on a classical computer with exponential precision.

Second, an n-qubit unitary operator U is said to be monomial if there exists a permutation π on the set of n-bit strings and a family of complex phases λ_x, such that

$$U|x\rangle = \lambda_x |\pi(x)\rangle \quad \text{for every } |x\rangle. \tag{20}$$

In other words U maps each standard basis state to another one, up to a global phase. Equivalently, one has $U = PD$ where $D = \sum \lambda_x |x\rangle\langle x|$ is a diagonal matrix and $P = \sum |\pi(x)\rangle\langle x|$ is a permutation matrix. The operation U is said to be efficiently computable if the functions

$$x \to \lambda_x, \quad x \to \pi(x) \quad \text{and} \quad x \to \pi^{-1}(x) \tag{21}$$

can be computed efficiently.

In our simulation of commuting Pauli circuits we will use the following classical simulation result proved in [8].

Theorem 6 (CT states) *Let $|\psi\rangle$ and $|\varphi\rangle$ be n-qubit CT states and let U be an n-qubit efficiently computable monomial operation. Then there exists a polynomial time classical algorithm to approximate $\langle\psi|U|\varphi\rangle$ with polynomial accuracy (and exponentially small probability of failure).*

For our purposes, it will be relevant that every stabilizer state is CT. More precisely, for every polynomial-size Clifford circuit $\mathcal{C}$ and standard basis state $|x\rangle$ (where x is an n-bit string), the state $|\psi\rangle = \mathcal{C}|x\rangle$ is CT relative to the description of $\mathcal{C}$ and the input x. Property (a) is the content of the Gottesman-Knill theorem [4]. Property (b) was shown in [26]; in fact for every stabilizer state $|\psi\rangle$ the standard basis coefficients $\langle y|\psi\rangle$ can be computed exactly. We refer to [8] for a more extensive discussion of CT states.

As for monomial operators, it is easily shown using (18) that every Pauli operator is unitary, monomial and efficiently computable. Second, every unitary operator of the form $\exp[i\theta Q]$, where Q is any (Hermitian) Z-type Pauli operator, is diagonal and hence monomial. Furthermore it is straightforward to show that any such operator is efficiently computable. More generally, it is useful to note (and easy to show):

Lemma 6 *If $U_1, \ldots, U_k$ are efficiently computable monomial unitary n-qubit operators and $k = poly(n)$, then also $\prod_{i=1}^{k} U_i$ is efficiently computable monomial.*

5.3   Proof of theorem 4

For clarity we prove theorem 4 separately even though it is superseded by theorem 5. Denote the input by $|x\rangle$ where x is an n-bit string. Denote the Pauli circuit by $\mathcal{U}$ and let $e^{i\theta_j P_j}$ denote its gates ($1 \le j \le m$). Let $\langle Z_i \rangle$ denote the expectation value of Z_i. First we invoke lemma 5, yielding a Clifford circuit $\mathcal{C}$ satisfying $\mathcal{C}^\dagger P_j \mathcal{C} = Q_j$ for some efficiently computable Hermitian Z-type operators Q_j. It follows that

$$e^{i\theta_j P_j} = \mathcal{C} e^{i\theta_j Q_j} \mathcal{C}^\dagger \tag{22}$$

and therefore $\mathcal{U} = \mathcal{C}\mathcal{D}\mathcal{C}^\dagger$ where $\mathcal{D}$ is given by the product of the m diagonal operators $e^{i\theta_j Q_j}$. Denote $P = \mathcal{C}^\dagger Z_i \mathcal{C}$ which is an efficiently computable Pauli operator. Furthermore denote $|\psi\rangle := \mathcal{C}^\dagger|x\rangle$. Then

$$\langle Z_i \rangle = \langle x|\mathcal{U}^\dagger Z_i \mathcal{U}|x\rangle = \langle\psi|\mathcal{D}^\dagger P \mathcal{D}|\psi\rangle. \tag{23}$$

Since $\mathcal{C}$ is a Clifford circuit, $|\psi\rangle$ is a CT state. Finally $M := \mathcal{D}^\dagger P \mathcal{D}$ is monomial and efficiently computable: indeed the Pauli operator P as well as each $e^{\theta_j Q_j}$ are efficiently computable monomial, as discussed in section 5.2. Applying lemma 6 then shows that M is efficiently computable monomial as well. Theorem 6 can now be applied.

5.4   Proof of theorem 5

We assume w.l.o.g. that Z is measured on the first qubit. Let $\mathcal{C}'$ be obtained by interspersing the commuting Pauli circuit $\mathcal{C} = \prod e^{i\theta P_j}$ with k additional gates $e^{i\theta Q_j}$ at arbitrary places in the circuit. Write

$$e^{i\theta Q_j} = [\cos\theta]I + [i\sin\theta]Q_j \tag{24}$$

for ever such additional gate. Doing so, the circuit $\mathcal{C}'$ is written as a linear combination of 2^k circuits (with coefficients of the form $(\cos\theta)^l(i\sin\theta)^{k-l}$), each of which being obtained by

replacing $e^{i\theta Q_j}$ by either I or Q_j. Thus every circuit in the linear combination is obtained by interspersing $\mathcal{C}$ with k Pauli operators. Using that $e^{i\theta P}Q = Qe^{\pm i\theta P}$ for every two Pauli operators P and Q, the Q_j can all be commuted to the right. As a result, we find that $\mathcal{C}'$ is written in the form

$$\mathcal{C}' = \sum_{\alpha=1}^{2^k} a_\alpha \mathcal{C}_\alpha \Sigma_\alpha, \tag{25}$$

where each coefficient a_α is efficiently computable, where each Σ_α is a Pauli operator and where each $\mathcal{C}_\alpha$ is a commuting Pauli circuit obtained by flipping a subset of the signs $P_j \to -P_j$ in the commuting circuit $\mathcal{C}$. Furthermore there are only poly(n) terms in the sum since $k = O(\log n)$ by assumption. To arrive at an efficient weak simulation of $\mathcal{C}'$ followed by measurement of Z_1, it suffices to show that each of the matrix elements

$$\langle x | \Sigma_\alpha \mathcal{C}_\alpha^\dagger Z_1 \mathcal{C}_\beta \Sigma_\beta | x \rangle \tag{26}$$

can be estimated efficiently with polynomial accuracy. First we can commute Z_1 to the right, transforming $\mathcal{C}_\beta$ into a commuting Pauli circuit $\overline{\mathcal{C}}_\beta$ obtained by changing some of the signs $P_j \to \pm P_j$ as before. Note that the combined circuit $\mathcal{C}_\alpha^\dagger \overline{\mathcal{C}}_\beta$ is a commuting Pauli circuit since all gates have the form $e^{\pm i\theta_j P_j}$. Furthermore $\Sigma_\alpha|x\rangle$ and $Z_1\Sigma_\beta|x\rangle$ are, up to global phases, simple standard basis states, say $|y\rangle$ and $|z\rangle$ resp., which can be computed efficiently. Analogous to the proof of theorem 4 we write $\overline{\mathcal{C}}_\alpha^\dagger \mathcal{C}_\beta = \mathcal{U}\mathcal{D}\mathcal{U}^\dagger$ where $\mathcal{U}$ is a polynomial size Clifford circuit and $\mathcal{D}$ is a product of diagonal gates. Putting everything together we find that (26) is, up to an efficiently computable overall phase, of the form $\langle y|\mathcal{U}\mathcal{D}\mathcal{U}^\dagger|z\rangle$ for some standard basis states $|y\rangle$ and $|z\rangle$. Since $\mathcal{U}^\dagger|y\rangle$ and $\mathcal{U}^\dagger|z\rangle$ are CT states (see section 5.2) and since $\mathcal{D}$ is efficiently computable monomial, we can apply theorem 6 yielding an efficient classical algorithm to estimate (26). This proves the result.

6 Mapping non-commuting circuits to commuting circuits

Here we show that commuting circuits can be used to efficiently reproduce the output of certain non-commutative processes. These results will provide evidence that commuting circuits can be used to solve tasks that appear nontrivial for classical computers.

6.1 Two-layer circuits

For every constant k we let Γ^k denote a computational model involving a universal classical computer supplemented with a restricted quantum computer operating with uniformly generated families of k-local commuting circuits acting on an arbitrary product input state and followed by Z measurement of the first qubit. By construction, Γ^k has the power to efficiently solve every problem in the complexity class P, for every k. Our goal is to investigate whether Γ^k-computations have the potential to outperform classical computers.

Theorem 7 (Mapping k-local non-commuting to $(k + 1)$-local commuting circuits) *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be uniform families of k-local n-qubit commuting circuits, where the gates in $\mathcal{C}_1$ need not commute with those in $\mathcal{C}_2$. Then there exists a polynomial time Γ^{k+1}-algorithm which approximates $\langle 0|\mathcal{C}_1\mathcal{C}_2|0\rangle$ with polynomial accuracy (with success probability exponentially close to 1).*

The above result shows that the non-commutativity in the two-layer circuit $\mathcal{C}_1\mathcal{C}_2$ can be "removed" by allowing gates to act on $k + 1$ qubits. The proof is an immediate consequence of the following alternate version of the Hadamard test (which regards arbitrary, i.e. not necessarily commuting, circuits).

Lemma 7 (Alternate Hadamard test) *Let $\mathcal{U} = U_{2m} \cdots U_1$ be an n-qubit quantum circuit of even size $2m$. Add one extra qubit line (henceforth called qubit 1) and for every $i = 1 \cdots m$ define the gate*

$$W_i = |0\rangle\langle 0| \otimes U_{2m+1-i}^{\dagger} + |1\rangle\langle 1| \otimes U_i, \tag{27}$$

which acts on qubit 1 and the qubits on which U_i and U_{2m+1-i} acted in the initial circuit $\mathcal{U}$. Consider the following circuit $\mathcal{U}'$ acting on the $(n+1)$-qubit input $|0\rangle$: first, apply H to qubit 1; second, apply the gates $W_1, \ldots, W_m$; third, apply H to qubit 1; finally measure Z on qubit 1. Then the probability of outputting 0 is

$$p(0) = \frac{1}{2}(1 + Re\langle 0|\mathcal{U}|0\rangle). \tag{28}$$

Analogously, replacing H in the third step by HP with $P = diag(1, i)$ yields the imaginary part of $\langle 0|\mathcal{U}|0\rangle$.

Remark that lemma 7 requires $\mathcal{U}$ to have even size. This is however not an essential requirement since a circuit of odd size $2m + 1$ can be "padded" with an additional identity. This yields a circuit $\mathcal{U}'$ of size $m + 1$.

The proof of the lemma is obtained by directly computing $p(0)$. Similar to the Hadamard test, the above result provides a simple quantum algorithm to estimate matrix elements of unitary quantum circuits with polynomial accuracy (and with success probability exponentially close to 1). Different from the standard Hadamard test, however, is that the *size* of the circuit $\mathcal{U}'$ used in lemma 7 is *half* the size of the original circuit $\mathcal{U}$ i.e. the alternate Hadamard test is "twice as fast". The price to pay for this is that the gates in $\mathcal{U}'$ act on a larger number of qubits: if $\mathcal{U}$ is a k-local circuit then $\mathcal{U}'$ can be as much as $(2k + 1)$-local.

Proof of theorem 7: Without loss of generality we can assume that $\mathcal{C}_1$ and $\mathcal{C}_2$ are in standard form, say $\mathcal{C}_1 = G_m \cdots G_1$ and $\mathcal{C}_2 = G'_m \cdots G'_1$ where $m = \binom{n}{k}$. By definition of the standard form, for every subset S of k qubits there is precisely one gate G_i and one gate G'_j such that $supp(G_i) \subseteq S$ and $supp(G'_j) \subseteq S$. By suitably labeling the gates in both circuits we can ensure that always $j = m + 1 - i$. Now apply lemma 7 to the circuit $\mathcal{U} := \mathcal{C}_1\mathcal{C}_2$ with the identification $U_i := G_i$ and $U_{m+i} := G'_i$ for every $i = 1 \cdots m$. Then each gate (27) acts on the qubits in S together with qubit 1 so that this gate is $(k+1)$-local (at most). Note furthermore that all gates W_i mutually commute. Finally, define $W'_i := [H \otimes I]W_i[H \otimes I]$ where H acts on qubit 1. Since all H gates in the middle cancel out, the $(k + 1)$-local commuting circuit $\mathcal{C} = \prod_i W'_i$ acting on $|0\rangle$ followed by measurement of Z_1 yields the same output as the circuit $\mathcal{U}'$ of lemma 7. This allows to estimate the real part of $\langle 0|\mathcal{C}_1\mathcal{C}_2|0\rangle$ with polynomial accuracy within the class Γ^{k+1}. The imaginary part is treated analogously. $\qquad\square$

6.2   *Constant-depth circuits*

Here we will relate commuting circuits with constant-depth circuits comprising *arbitrary* gates.

Theorem 8 (Estimating constant-depth matrix elements) *Let $\mathcal{U}$ be a n-qubit quantum circuit from a uniform family of circuits that has constant depth m. Then there exists a polynomial time Γ^k-algorithm to approximate $|\langle 0|\mathcal{U}|0\rangle|^2$ with polynomial accuracy (and with success probability exponentially close to 1) where $k = 2^m + 1$.*

Recall that the problem of estimating matrix elements $|\langle 0|\mathcal{U}|0\rangle|^2$ of polynomial size quantum circuits of arbitrary depth is known to be BQP-hard (and the naturally corresponding decision problem is BQP-complete). Theorem 8 shows that such matrix elements can be estimated efficiently with k-local commuting circuits with constant k as long as $\mathcal{U}$ has constant depth (with an exponential scaling of k with m). Although one would not expect the constant-depth matrix problem (with polynomially small error) to be BQP-hard, this task appears to be nontrivial for classical computers and, to our knowledge, no efficient classical algorithm is known. On the other hand, it is known to be hard to estimate constant-depth matrix elements with high accuracy [10, 11]. For example, using the methods of [10] it can be shown that if we can classically estimate constant-depth matrix elements with exponentially small error in polynomial time, we could solve all #P problems efficiently.

Proof of theorem 8: Letting Z_j denote the operator Z acting on qubit j, we define $Z(S) = \prod_{j \in S} Z_j$ for every subset $S \subseteq \{1, \ldots, n\}$. Using

$$|0\rangle\langle 0| = \frac{1}{2^n} \sum Z(S), \tag{29}$$

where the sum is over all subsets S, one finds

$$|\langle 0|\mathcal{U}|0\rangle|^2 = \langle 0|\mathcal{U}^\dagger|0\rangle\langle 0|\mathcal{U}|0\rangle = \frac{1}{2^n} \sum \langle 0|\mathcal{U}^\dagger Z(S)\mathcal{U}|0\rangle. \tag{30}$$

Setting $G_j := \mathcal{U}^\dagger Z_j \mathcal{U}$ yields

$$\langle 0|\mathcal{U}^\dagger Z(S)\mathcal{U}|0\rangle = \langle 0| \prod_{j \in S} G_j |0\rangle =: F(S) \tag{31}$$

for every subset S. Since the Z_j mutually commute, the G_j mutually commute as well as these operators are obtained by simultaneously conjugating the Z_j. Furthermore since $\mathcal{U}$ has depth m, each G_j acts on at most 2^m qubits. Thus $F(S)$ is a matrix element of a 2^m-local commuting circuit. Via the standard Hadamard test (recall Fig. 1 and the proof of theorem 3) one constructs a k-local commuting circuit with $k = 2^m + 1$ which allows to estimate any such matrix element with polynomial accuracy in polynomial time, with success probability exponentially close to 1.

We now use these findings to give an efficient Γ^k-algorithm to estimate $\gamma := |\langle 0|\mathcal{U}|0\rangle|^2$ with polynomial accuracy. Owing to (30)-(31), one has $\gamma := 2^{-n} \sum F(S)$. Thus γ equals the expectation value of a random variable over the collection of all 2^n subsets S which takes the value $F(S)$ with uniform probability. Fix $\epsilon > 0$. First we generate K subsets $S_\alpha \subseteq \{1, \ldots, n\}$ uniformly at random. Applying the Chernoff-Hoeffding bound we find that, for some sufficiently large $K = \text{poly}(n, 1/\epsilon)$, one has

$$\left| \frac{1}{K} \sum_{\alpha=1}^{K} F(S_\alpha) - \gamma \right| \leq \epsilon/2 \tag{32}$$

with probability exponentially close to 1. Next, as described above we can efficiently compute an estimate f_α of each $F(S_\alpha)$ using Γ^k-circuits with $k = 2^m + 1$; more precisely, we compute K numbers f_α satisfying $|f_\alpha - F(S_\alpha)| \leq \epsilon/2$. The runtime of the computation will be $\text{poly}(n, 1/\epsilon)$ and the success probability exponentially close to 1. Finally, we compute $c := [\sum f_\alpha]/K$ which takes $\text{poly}(n, 1/\epsilon)$ time as well. Using (32) and the triangle inequality it follows that $|c - \gamma| \leq \epsilon$. Thus c is our desired polynomial approximation of γ.   □

Finally we note that theorem 8 can be generalized in the following rather intriguing sense: using Γ^k-circuits one can also efficiently estimate matrix elements of the form $|\langle 0|\mathcal{U}\mathcal{C}|0\rangle|^2$ where $\mathcal{U}$ is again a constant-depth circuit and where $\mathcal{C}$ represents an arbitrary uniform family of *Clifford circuits*. Interestingly, these Clifford circuits need not have constant depth. The proof, which is given in Appendix A, uses an argument analogous to the proof of theorem 8 combined with the alternate Hadamard test given in lemma 7.

Acknowledgements

References

1. P. W. Shor (1999), *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM review, vol. 41, no. 2, pp. 303–332.
2. R. Jozsa and N. Linden (2003), *On the role of entanglement in quantum computational speed-up*, Proc. R. Soc. A, vol. 459, pp. 2011–2032, quant-ph/0201143.
3. G. Vidal (2003), *Efficient classical simulation of slightly entangled quantum computations*, Phys. Rev. Lett., vol. 91, p. 147902, quant-ph/0301063.
4. D. Gottesman (1997), *Stabilizer Codes and Quantum Error Correction*, quant-ph/9705052.
5. L. G. Valiant (2002), *Quantum Circuits That Can Be Simulated Classically in Polynomial Time*, SIAM J. Comput., vol. 31, pp. 1229–1254.
6. E. Knill (2001), *Fermionic Linear Optics and Matchgates*, quant-ph/0108033.
7. R. Jozsa and A. Miyake (2008), *Matchgates and classical simulation of quantum circuits*, Proc. R. Soc. A, vol. 464, pp. 3089–3106, arXiv:0804.4050.
8. M. Van den Nest (2010), *Simulating quantum computers with probabilistic methods*, Quantum Inf. and Comp., vol. 11, pp. 784–812, arXiv:0911.1624.
9. M. Van den Nest (2012), *Efficient classical simulations of quantum Fourier transforms and normalizer circuits over Abelian groups*, arXiv:1201.4867.
10. B. Terhal and D. DiVincenzo (2004), *Adptive quantum computation, constant depth quantum circuits and arthur-merlin games*, Quantum Inf. and Comp., vol. 4, no. 2, pp. 134–145, quant-ph/0205133.
11. S. Fenner, F. Green, S. Homer, and Y. Zhang (2005), *Bounds on the power of constant-depth quantum circuits* in *Fundamentals of Computation Theory*, pp. 44–55, Springer, quant-ph/0312209.
12. M. J. Bremner, R. Jozsa, and D. J. Shepherd (2011), *Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy*, Proc. R. Soc. A, vol. 467, pp. 459–472, arXiv:1005.1407.
13. S. Jordan (2010), *Permutational quantum computing*, Quantum Inf. and Comp., vol. 10, no. 5, pp. 470–497, arXiv:0906.2508.
14. S. Aaronson and A. Arkhipov (2011), *The computational complexity of linear optics* in *Proceedings*

of the 43rd annual ACM symposium on Theory of computing, pp. 333–342, ACM, arXiv:1011.3245.

15. M. J. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne (2011), *Non-adaptive Measurement-based Quantum Computation and Multi-party Bell Inequalities*, New J. Phys, vol. 13, p. 023014, arXiv:1009.5213.

16. H. Briegel and R. Raussendorf (2001), *Persistent entanglement in arrays of interacting particles*, Physical Review Letters, vol. 86, no. 5, pp. 910–913, quant-ph/0004051.

17. D. Shepherd and M. J. Bremner (2009), *Temporally unstructured quantum computation*, Proc. R. Soc. A, vol. 465, pp. 1413–1439, arXiv:0809.0847.

18. D. Shepherd (2010), *Binary Matroids and Quantum Probability Distributions*, arXiv:1005.1744.

19. S. Bravyi and M. Vyalyi (2005), *Commutative version of the k-local Hamiltonian problem and common eigenspace problem*, Quantum Inf. and Comp., vol. 5, pp. 187–215, quant-ph/0308021.

20. D. Aharonov and L. Eldar (2011), *On the complexity of Commuting Local Hamiltonians, and tight conditions for Topological Order in such systems* in *IEEE 52nd Annual Symposium on Foundations of Computer Science 2011*, pp. 334–343, IEEE, arXiv:1102.0770.

21. N. Schuch (2011), *Complexity of commuting Hamiltonians on a square lattice of qubits*, Quantum Inf. and Comp., vol. 11, no. 11-12, pp. 901–912, arXiv:1105.2843.

22. M. Hastings (2012), *Trivial Low Energy States for Commuting Hamiltonians, and the Quantum PCP Conjecture*, arXiv:1201.3387.

23. P. Hayden, D. Leung, P. Shor, and A. Winter (2004), *Randomizing quantum states: Constructions and applications*, Commun. Math. Phys., vol. 250, no. 2, pp. 371–391, quant-ph/0307104.

24. M. A. Nielsen and I. L. Chuang (2000), *Quantum computation and quantum information.* Cambridge University Press.

25. J.-L. Brylinski and R. Brylinski (2002), *Universal quantum gates*, Mathematics of Quantum Computation, quant-ph/0108062.

26. J. Dehaene and B. De Moor (2003), *The Clifford group, stabilizer states, and linear and quadratic operations over GF(2)*, Phys. Rev. A, vol. 68, p. 042318, quant-ph/0304125.

Appendix A Generalization of theorem 8

Theorem A.1 *Let $\mathcal{U}$ be a (uniform family of) n-qubit quantum circuit(s) of depth m. Let $\mathcal{C}$ be a (uniform family of) n-qubit Clifford circuit(s). Then the problem of estimating the matrix element $|\langle 0|\mathcal{C}\mathcal{U}|0\rangle|^2$ with polynomial accuracy and with success probability exponentially (in n) close to 1 is in Γ^k with $k = 2^m + 1$.*

Proof. Similar to (30) one has

$$|\langle 0|\mathcal{C}\mathcal{U}|0\rangle|^2 = \frac{1}{2^n} \sum \langle 0|\mathcal{U}^\dagger \mathcal{C}^\dagger Z(S)\mathcal{C}\mathcal{U}|0\rangle. \tag{A.1}$$

Since $\mathcal{C}$ is Clifford, $\mathcal{C}^\dagger Z(S)\mathcal{C} =: P$ is a Pauli operator which can moreover be determined efficiently; that we suppress dependence of P on S to simplify notation. Following (18), we can write

$$P = i^t \prod X_k^{a_k} Z_k^{b_k}, \quad \text{where } t \in \{0, 1, 2, 3\}, \ a_k, b_k \in \{0, 1\}. \tag{A.2}$$

Now define $G_k := \mathcal{U}^\dagger X_k^{a_k}\mathcal{U}$ and $H_k := \mathcal{U}^\dagger Z_k^{a_k}\mathcal{U}$ as well as $\mathcal{C}_1 := \prod G_k$ and $\mathcal{C}_2 := \prod H_k$. Then

$$\langle 0|\mathcal{U}^\dagger \mathcal{C}^\dagger Z(S)\mathcal{C}\mathcal{U}|0\rangle = i^t \langle 0|\mathcal{C}_1\mathcal{C}_2|0\rangle. \tag{A.3}$$

Since the Z_k mutually commute, the G_k mutually commute as well. Furthermore each G_j acts on at most 2^m qubits. Therefore $\mathcal{C}_1$ is a 2^m-local commuting circuit. Similarly, $\mathcal{C}_2$ is a

2^m-local commuting circuit as well. We can now apply theorem 7, showing that $\langle 0|\mathcal{C}_1\mathcal{C}_2|0\rangle$ can be estimated with polynomial accuracy using Γ^k-circuits with $k = 2^m + 1$. Continuing the argument as in the proof of theorem 8 completes the proof. $\square$.