Average-Case Quantum Advantage with Shallow Circuits

François Le Gall

Graduate School of Informatics, Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan
legall@i.kyoto-u.ac.jp

—— Abstract ——

Recently Bravyi, Gosset and König (Science 2018) proved an unconditional separation between the computational powers of small-depth quantum and classical circuits for a relation. In this paper we show a similar separation in the average-case setting that gives stronger evidence of the superiority of small-depth quantum computation: we construct a computational task that can be solved on all inputs by a quantum circuit of constant depth with bounded-fanin gates (a "shallow" quantum circuit) and show that any classical circuit with bounded-fanin gates solving this problem on a non-negligible fraction of the inputs must have logarithmic depth. Our results are obtained by introducing a technique to create quantum states exhibiting global quantum correlations from any graph, via a construction that we call the *extended graph*.

Similar results have been very recently (and independently) obtained by Coudron, Stark and Vidick (`arXiv:1810.04233`), and Bene Watts, Kothari, Schaeffer and Tal (STOC 2019).

1 Introduction

1.1 Background and our results

A fundamental problem in quantum complexity theory is to prove the superiority of quantum computation over classical computation. While this has been shown in constrained models of computation such as query complexity (see for instance [4] for a recent survey), in weak models of computation like finite-state automata [21], and when considering relativized complexity classes (see, e.g., [7] for the first results and [25] for the most recent breakthrough), no definite answer is known in standard computational models such as Turing machines or general circuits. Indeed, since the complexity class BQP corresponding to the problems that can be solved efficiently by a quantum computer satisfies the inclusions P ⊆ BQP ⊆ PSPACE, unconditionally separating P and BQP cannot be shown without separating P and PSPACE.

A recent active research area focuses on conditionally showing the superiority of quantum computation. Under several assumptions from computational complexity such as non-collapse of the polynomial hierarchy, the superiority of quantum computation with respect to classical computation has been shown in the standard circuit model in the worst-case setting [1, 2, 3, 11, 15, 17, 16, 23, 27] and even in the average-case setting [1, 2, 3, 8, 12, 13, 17].

Note that showing the superiority in the average-case setting is a much stronger evidence of the superiority of quantum computation than a proof for the worst-case setting.

A recent breakthrough by Bravyi, Gosset and König [9, 10] showed an *unconditional* separation between the computational powers of quantum and classical small-depth circuits: they constructed a computational problem that can be solved by quantum circuits of constant depth with bounded-fanin[1] gates ("shallow quantum circuits") and showed that any classical circuit with bounded fanin gates solving this problem on all inputs must have depth $\Omega(\log m)$, where m denotes the input size. Besides being the first such unconditional separation in the circuit model, this separation is also especially important since shallow quantum circuits are likely to be the easiest quantum circuits to experimentally implement, due to their robustness to noise and decoherence. (Note that separations were already known when allowing gates with unbounded fanin or fanout [18, 20, 26]. The strength of Bravyi, Gosset and König's result is that it holds for the weaker model of quantum circuits with bounded fanin and fanout.)

The original classical lower bound shown in [9] required the classical circuit to output the correct answer (with high probability) on each input, i.e., this was only a worst-case hardness result. Showing the advantages of shallow quantum circuits for a distribution (i.e., proving a corresponding average-case hardness result), which would give a significantly stronger evidence of the advantage of quantum shallow circuits, was discussed in [9, Section 5] and referred to as a "challenging open question". The recently published journal version [10] partially answers this open question: it presents an average-case lower bound showing that any classical circuit that outputs the correct answer on a constant fraction of some restricted subset of the inputs (which can be efficiently sampled) must have logarithmic depth. In other words, it shows that any sublogarithmic-depth classical algorithm will fail with some constant probability on a input chosen uniformly at random in this restricted subset.

In this work we give a stronger average-case hardness result. Our main result is the following theorem.

▶ **Theorem 1.** *There exists a relation $\mathcal{R} \subseteq \{0,1\}^M \times \{0,1\}^N$ for which the following two assertions hold.[2]*

   ▬ *There is a constant-depth quantum circuit with bounded-fanin gates (i.e., a shallow quantum circuit) that on any input $x \in \{0,1\}^M$ outputs an element in the set $\mathcal{R}(x)$ with probability 1.*

   ▬ *There is a constant $\gamma > 0$ such that any randomized circuit C with bounded-fanin gates satisfying*

$$\frac{1}{2^M} \sum_{x \in \{0,1\}^M} \Pr[C(x) \in \mathcal{R}(x)] \geq \frac{1}{\exp(\gamma\sqrt{M})}$$

   has depth $\Omega(\log M)$.

Theorem 1 thus shows the existence of a computational problem that can be solved by a shallow quantum circuit on all inputs but such that any classical circuit with bounded-fanin gates solving this problem on a non-negligible fraction of the inputs must have logarithmic

---

[1]   In this paper the term *bounded-fanin* means, as usual, that the fanin is bounded from above by a constant.

[2]   As usual in computational complexity, the subset $\mathcal{R} \subseteq \{0,1\}^M \times \{0,1\}^N$ is interpreted as the following computational problem: given an input $x \in \{0,1\}^M$, output any element of the set $\{z \in \{0,1\}^N \,|\, (x,z) \in \mathcal{R}\}$. Through this paper we will use the convenient notation $\mathcal{R}(x) = \{z \in \{0,1\}^N \,|\, (x,z) \in \mathcal{R}\}$, for any $x \in \{0,1\}^N$.

depth. This gives an average-case result that is a strengthening of the average-case result from [10] with respect to two aspects. First, our lower bound holds for any classical circuit that solves the problem on a non-negligible fraction of the inputs (even exponentially small), and not only on a constant fraction. Second, our statement does not make any restriction on the set of inputs for which the hardness is established, i.e., it shows that any sublogarithmic-depth classical algorithm will fail with high probability on an input chosen uniformly at random in the whole set $\{0,1\}^M$.

1.2 Overview of our techniques

Main technical result. Our central technical result is the following theorem.

▶ **Theorem 2.** *There exists a relation $R \subseteq \{0,1\}^m \times \{0,1\}^n$ for which the following two assertions hold.*

- *There is a constant-depth quantum circuit with bounded-fanin gates (i.e., a shallow quantum circuit) that on any input $x \in \{0,1\}^m$ outputs a string in the set $R(x)$ with probability 1.*

- *There is a constant $\alpha > 0$ such that any randomized circuit C with bounded-fanin gates satisfying*

$$\frac{1}{2^m} \sum_{x \in \{0,1\}^m} \Pr[C(x) \in R(x)] \geq 1 - \alpha$$

has depth $\Omega(\log m)$.

Theorem 1 is obtained from Theorem 2 by amplifying the soundness using standard techniques: the relation $\mathcal{R}$ is obtained by taking the direct product of t copies of the relation R for some sufficiently large integer t (the sizes of the inputs and outputs in $\mathcal{R}$ are thus $M = mt$ and $N = nt$). We show in Section 6 how the soundness is then amplified from $1 - \alpha$ to $(1 - \alpha)^{t'}$ with $t' \approx t$ by this process and observe that $(1 - \alpha)^{t'}$ is upper bounded by $1/\exp(\gamma\sqrt{M})$ for some constant γ. Note that this approach can also be applied to amplify the soundness of the average-case result in [10], which directly gives a result similar to Theorem 1 (but for a hard distribution that is not simply the uniform distribution).

Techniques from prior works. Before presenting our techniques we first describe how the result from [9, 10] was obtained. A central technical tool is a simple but fascinating result by Barrett et al. [5] that shows that correlations arising from local entanglement cannot be simulated classically without global interaction. This result was also used recently to show a separation between quantum and classical distributed computing [22]. More precisely, [5] considers the problem of simulating the measurement outcomes that occur when measuring each qubit of a well-chosen quantum state on n qubits (the graph state associated with a cycle of length n) in either the X-basis or the Y-basis (the choice of the basis depends on input bits), and shows that creating the resulting output distribution classically requires coordinating the outcomes of qubits located at distance $\Omega(n)$ on the cycle. This result can actually easily be adapted to show that any classical circuit with one-dimensional nearest-neighbor architecture and bounded-fanin gates requires logarithmic depth to create this distribution, since otherwise distant wires cannot interact. Since a graph state over a cycle (and more generally over any constant-degree graph) can be created using a shallow quantum circuit, this already gives an unconditional separation between the computational power of quantum shallow circuits and the computational power of this restricted class of small-depth classical circuits.

The main contribution of [9, 10] is to show how to get a similar separation without restricting the topology of the classical circuit (other than its depth, naturally). A first important observation is that while interactions can now naturally occur between distant wires, any sublogarithmic-depth bounded-fanin classical circuit C cannot create interactions between all pairs of wires. Ref. [9] showed that it is then always possible to find a large subset of wires S_C that are connected as a long cycle and in which distant wires do not interact. The key idea is then to consider a computational problem (called 2D Hidden Linear Function) where the input is divided in two parts: one part specifies the basis in which the qubits of the graph state are measured and the second part the topology of the graph state. By using the second part of the input to force the graph state to use only nodes corresponding to wires in S_C, the same argument as in [5] can be again applied on the cycle defined by S_C to conclude that the sublogarithmic-depth classical circuit C cannot output a valid output with high probability.

Our approach.   Let us now describe the main ideas of our approach to prove Theorem 2. Our main technical tool, described in Section 3 is a generalization of the construction from [5]: we show how to generate useful quantum correlations not only from a cycle but also from any undirected graph G. The key insight is to consider what we call the *extended graph of G*, denoted $\overline{G}$, which is obtained by adding a vertex on each edge of G. We show that when measuring the qubits of the graph state corresponding to $\overline{G}$ in either the X-basis or the Y-basis, we get probability distributions that satisfy global conditions related to properties of subgraphs (in particular paths and cycles) of $\overline{G}$. The conditions are described in Theorems 5 and 6.

In order to prove our separations we consider a $d^3 \times d^3$ square grid in which one vertex (called a control vertex) is placed at the center of each 1×1 square of the grid (and connected by 4 edges to the 4 corners of the square), and then adding one vertex on each edge. The final graph is denoted $\overline{\mathcal{G}}_d$. The construction is described in Section 4. Note that by construction $\overline{\mathcal{G}}_d$ is an extended graph. This means that the probability distributions arising when measuring the qubits of the graph state associated with this graph, which we denote $|\overline{\mathcal{G}}_d\rangle$, can be described by Theorems 5 and 6.

We can now describe the computational problem that we consider to show our separation. Let m denote the number of control vertices in $\overline{\mathcal{G}}_d$ and n denote the total number of vertices. Observe that $m = \Theta(d^6)$ and $n = \Theta(d^6)$. Given as input a string of bits $x \in \{0,1\}^m$, we consider the following process: measure each qubit of the quantum state $|\overline{\mathcal{G}}_d\rangle$ in the X-basis except the qubits corresponding to the control vertices, which are measured either in the X-basis or in the Y-basis depending on the value of x. The relation R considered to prove Theorem 2 simply asks, given $x \in \{0,1\}^m$ as input, to compute any sequence of measurement outcomes $z \in \{0,1\}^n$ that has non-zero probability of being obtained by this process. Note that this problem can be solved by shallow quantum circuits: the graph $\overline{\mathcal{G}}_d$ has constant degree and thus the graph state $|\overline{\mathcal{G}}_d\rangle$ can be constructed in constant depth.

In Section 5 we first show that for any sublogarithmic-depth bounded-fanin classical circuit C there exists a subset S_C of wires that are connected as a long cycle and in which distant wires do not interact. The proof of this claim is similar to what was done in [9, 10]. We then show that this claim, along with Theorems 5 and 6, are enough to prove that the sublogarithmic-depth classical circuit C cannot output a valid output with high probability. The key point of our argument – and the reason why our result holds for average-case hardness on the whole set $\{0,1\}^m$ of possible inputs and not only for worst-case hardness or average-case hardness on a restricted set of inputs – is that we do not need to construct the

graph state corresponding to the subgraph induced by S_C, i.e., we do not need to adapt the topology of the measured graph state to the circuit. Theorems 5 and 6 guarantee that we can instead work with the graph state $|\overline{\mathcal{G}}_d\rangle$ corresponding to the whole graph and simply look at the relevant part of the probability distribution (the part corresponding to the wires in S_C).

Related works. A similar result has been recently (and independently) obtained by Coudron, Stark and Vidick and expanded into a framework for robust randomness expansion [14]. The proof techniques are nevertheless different: [14] constructs a problem hard for small-depth classical circuits by starting with a non-local game and showing how to plant a polynomial number of copies of the game into a graph. Our approach, on the other hand, starts with a graph and shows how to create from it a quantum state exhibiting global quantum correlations that cannot be simulated by small-depth classical circuits with bounded-fanin gates.

An even stronger result has been very recently announced: Bene Watts, Kothari, Schaeffer and Tal [6] have shown that the 2D Hidden Linear Function introduced in [9, 10] cannot be solved on a non-negligible fraction of the inputs even by small-depth classical circuits with unbounded-fanin parity gates.

2 Preliminaries

2.1 General notations and a technical lemma

Given a Boolean function $f\colon A \to \{0,1\}$ on a finite set A, we write $|f|$ the number of elements $a \in A$ such that $f(a) = 1$, i.e., $|f| = \sum_{a\in A} f(a)$. Similarly, for any finite binary string $x \in \{0,1\}^*$, we denote $|x|$ the Hamming weight of x, i.e., the number of non-zero bits of x.

All the graphs considered in this paper will be undirected. Given a graph $G = (V, E)$ and any vertex $u \in V$, we denote

$$\mathcal{N}(u) = \{v \in V \mid \{u, v\} \in E\}$$

the set of neighbors of u. Given a path p in the graph G we will often be mainly interested only in the set of vertices on the path. For a vertex $v \in V$, we will thus use the convenient notation $v \in p$ to express the fact that v is on the path p.

The notation $\oplus$ will denote the addition modulo 2 (i.e., the bit parity). We will use the following lemma, which was first implicitly mentioned in [5], and stated formally (but in a form slightly different from the form we present below) in [9, 10]. For completeness we include a proof.

▶ **Lemma 3.** *([5, 9, 10]) Consider any affine function $q\colon \{0,1\}^3 \to \{0,1\}$ and any three affine functions $q_1\colon \{0,1\}^2 \to \{0,1\}$, $q_2\colon \{0,1\}^2 \to \{0,1\}$, $q_3\colon \{0,1\}^2 \to \{0,1\}$ such that*

$$q_1(b_2, b_3) \oplus q_2(b_1, b_3) \oplus q_3(b_1, b_2) = 0 \tag{1}$$

holds for any $(b_1, b_2, b_3) \in \{0,1\}^3$. Then at least one of the four following equalities does not hold:

$$q(0,0,0) \quad = 0, \tag{2}$$
$$q(0,1,1) \oplus q_1(1,1) \quad = 1, \tag{3}$$
$$q(1,0,1) \oplus q_2(1,1) \quad = 1, \tag{4}$$
$$q(1,1,0) \oplus q_3(1,1) \quad = 1. \tag{5}$$

Proof. Consider any affine function $q\colon \{0,1\}^3 \to \{0,1\}$ and any three affine functions $q_1, q_2, q_3\colon \{0,1\}^2 \to \{0,1\}$ satisfying Condition (1) for all $(b_1, b_2, b_3) \in \{0,1\}^3$. These four functions can be written as

$$
\begin{aligned}
q(b_1, b_2, b_3) &= \alpha_0 \oplus \alpha_1 b_1 \oplus \alpha_2 b_2 \oplus \alpha_3 b_3, & (6)\\
q_1(b_2, b_3) &= \beta_0 \oplus \beta_2 b_2 \oplus \beta_3 b_3, & (7)\\
q_2(b_1, b_3) &= \gamma_0 \oplus \gamma_1 b_1 \oplus \beta_3 b_3, & (8)\\
q_3(b_1, b_2) &= (\beta_0 \oplus \gamma_0) \oplus \gamma_1 b_1 \oplus \beta_2 b_2. & (9)
\end{aligned}
$$

for some coefficients $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_2, \beta_3, \gamma_0, \gamma_1 \in \{0,1\}$. Assume that these functions satisfy all the four equations (2)-(5). Equation (2) implies that $\alpha_0 = 0$. Consider the quantity

$$
\lambda = q(1,1,0) \oplus q_1(1,1) \oplus q(0,1,1) \oplus q_2(1,1) \oplus q(1,0,1) \oplus q_3(1,1).
$$

Computing this quantity using the four equations (6)-(9) gives $\lambda = 3\alpha_0 = 0$. On the other hand, computing λ using the three equations (3)-(5) gives $\lambda = 1 \oplus 1 \oplus 1 = 1$, which leads to a contradiction and implies that the four equations (2)-(5) cannot hold simultaneously.   ◀

2.2   Quantum computation: graph states and their measurements

Quantum gates.   We assume that the reader is familiar with the basics of quantum computation and refer to [24] for a standard reference. We will use the Hadamard gate H and the Pauli X, Y and Z gates:

$$
H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
$$

where i denotes the imaginary unit of complex numbers. Note that $XZ = -ZX = -iY$. We will use two kinds of measurements: measurements in the X-basis and measurements in the Y-basis, which correspond to projective measurements with observables X and Y, respectively. Concretely, a measurement in the X-basis is realized by applying a Hadamard gate to this qubit and then measuring it in the computational basis $\{|0\rangle, |1\rangle\}$. A measurement in the Y-basis is realized by applying the gate

$$
\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}
$$

to this qubit and then measuring it in the computational basis.[3]

Graph states.   Graph states are quantum states that can be described using graphs [19]. Let $G = (V, E)$ be any undirected graph. The graph state associated with G is the quantum state on $|V|$ qubits obtained by first constructing the state

$$
\bigotimes_{u \in V} |0\rangle_{\mathsf{Q}_u},
$$

---

[3] The outcome of a measurement in the X-basis or the Y-basis is often defined as an element in $\{-1, 1\}$, i.e., the outcome corresponds to one of two eigenvalues of the observables X and Y. In our description the measurement outcome is a bit (the two bits 0 and 1 correspond to the two eigenvalues 1 and -1, respectively), which will be more convenient to describe our results.

where each $\mathbf{Q}_u$ represents a 1-qubit register, then applying a Hadamard gate on each register and, finally, applying a Controlled-Z gate on $(\mathbf{Q}_u, \mathbf{Q}_v)$ for any pair $\{u, v\} \in E$. We will write $|G\rangle$ the graph state associated with G.

Graph states can equivalently be defined using the stabilizer formalism. For each vertex $u \in V$ define the operator

$$\pi_u = X_u \otimes \bigotimes_{v \in \mathcal{N}(u)} Z_v,$$

where we use X_u to denote the Pauli operator X applied to Register $\mathbf{Q}_u$ and use Z_v to denote the Pauli operator Z applied to Register $\mathbf{Q}_v$. Observe that all these operators commute, and

$$\pi_u |G\rangle = |G\rangle$$

for each $u \in G$. The graph state $|G\rangle$ is thus the simultaneous eigenstate, associated with the eigenvalue 1, of all these operators .

Measurements of graph states. The description of graph states using the stabilizer formalism is especially convenient to derive the properties of measurements we describe below (we refer to [24] for details of the general discussion of measurements of stabilizer states and state below only the properties we will use in this paper).

Consider the graph state $|G\rangle$ of a graph $G = (V, E)$. Let $U_X, U_Y \subseteq V$ be any two disjoint subsets of vertices. Assume that we measure Register $\mathbf{Q}_u$, for each vertex $u \in U_X$, in the X-basis and measure Register $\mathbf{Q}_v$, for each vertex $v \in U_Y$, in the Y-basis. The observable corresponding to this measurement is

$$M = \prod_{u \in U_X} X_u \prod_{v \in U_Y} Y_v.$$

For each $u \in U_X \cup U_Y$, let $z_u \in \{0, 1\}$ denote the random variable corresponding to the measurement outcome of the measurement performed on Register $\mathbf{Q}_u$. Let us denote

$$z = \bigoplus_{u \in U_X \cup U_Y} z_u$$

the random variable corresponding to the parity of all the measurement outcomes. Using the stabilizer formalism it is easy to show that the value of this random variable is as follows:
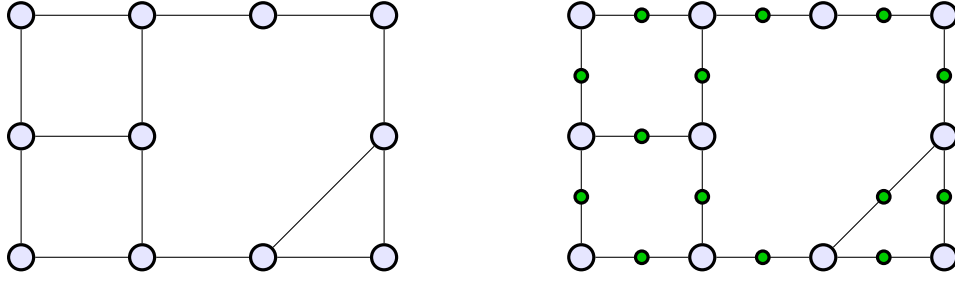
- if M can be written as $M = \prod_{u \in S} \pi_u$ for some set $S \subseteq V$ then $z = 0$ with probability 1;
- if M can be written as $M = -\prod_{u \in S} \pi_u$ for some set $S \subseteq V$ then $z = 1$ with probability 1;
- if M cannot be written as $M = \prod_{u \in S} \pi_u$ or $M = -\prod_{u \in S} \pi_u$ for some set $S \subseteq V$ then $z = 0$ with probability $1/2$ and $z = 1$ with probability $1/2$.

3 Extended Graphs and their Graph States

In this section we describe the general construction on which our results are based.

For any undirected graph $G = (V, E)$, let $\overline{G}$ denote the graph with $|V| + |E|$ vertices and $2|E|$ edges obtained from G by inserting a vertex at the middle of each edge of G. We call $\overline{G}$ the *extended graph of G*. We will write V^* the set of inserted vertices and consider $\overline{G}$ as a graph over the vertex set $V \cup V^*$. We refer to Figure 1 for an illustration.
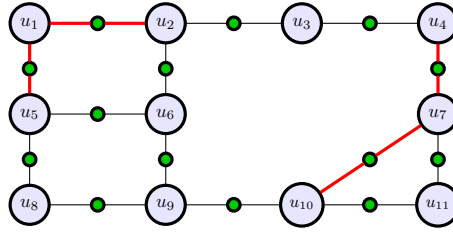
We now define the concept of *f-covering* of a graph.

■ **Figure 1** Example for our construction. The graph $G = (V, E)$ is represented on the left. The extended graph $\overline{G}$ is represented on the right. In this figure the large circles represent the vertices in V, while the small circles represent the vertices in V^*.

▶ **Definition 4.** *Let $G = (V, E)$ be an undirected graph and $f : V \to \{0, 1\}$ be any function such that $|f|$ is even. An f-covering of $\overline{G}$ is a set of $|f|/2$ paths of $\overline{G}$ such that each vertex in $\{v \in V \mid f(v) = 1\}$ appears once as an endpoint of one of these paths.*

We refer to Figure 2 for an illustration. Note that the $|f|/2$ paths of an f-covering do not need to be edge-disjoint.

■ **Figure 2** Illustration of the concept of f-covering. Here $V = \{u_1, \dots, u_{11}\}$ and $f : V \to \{0, 1\}$ is defined as follows: $f(u_2) = f(u_4) = f(u_5) = f(u_{10}) = 1$ and $f(u_1) = f(u_3) = f(u_6) = f(u_7) = f(u_8) = f(u_9) = f(u_{11}) = 0$. The two paths depicted in red form an f-covering.

---

1. Construct the graph state over $\overline{G}$.
2. For each $v \in V$ such that $f(v) = 1$, measure the qubit of the node v in the Y-basis. Measure the qubits of all the other nodes of $\overline{G}$ in the X-basis.

■ **Figure 3** The process $\mathcal{P}(G, f)$.

Given a graph $G = (V, E)$ and a function $f : V \to \{0, 1\}$, consider the process $\mathcal{P}(G, f)$ described in Figure 3. For any vertex $v \in V \cup V^*$, let z_v denote the random variable corresponding to the outcome of the measurement performed on the qubit of node v. The following two theorems describe the correlations among these random variables.

▶ **Theorem 5.** *For any cycle $\mathcal{C}$ of $\overline{G}$ the following equality holds with probability 1:*

$$\bigoplus_{v \in \mathcal{C} \cap V^*} z_v = 0. \tag{10}$$

Proof. In Process $\mathcal{P}(G, f)$ all the vertices in $\mathcal{C} \cap V^*$ are measured in the X-basis. Since $\mathcal{C}$ is a cycle we have

$$\prod_{v \in \mathcal{C} \cap V^*} \pi_v = \prod_{v \in \mathcal{C} \cap V^*} X_v,$$

which is the measurement operator corresponding to this measurement. The discussion of Section 2.2 implies that the parity of all the measurement outcomes is always zero, as claimed. ◄

▶ **Theorem 6.** *Assume that $|f|$ is even and let $\{p_1, \ldots, p_{|f|/2}\}$ be any f-covering of $\overline{G}$. Let us write*

$$z_V = \bigoplus_{v \in V} z_v.$$

Then the following equality holds with probability 1:

$$z_V \oplus \bigoplus_{i=1}^{|f|/2} \bigoplus_{v \in p_i \cap V^*} z_v = \begin{cases} 0 & \text{if } |f| \bmod 4 = 0, \\ 1 & \text{if } |f| \bmod 4 = 2. \end{cases} \tag{11}$$

Proof. Let $V_1 = \{u_1, \ldots, u_{|f|/2}, v_1, \ldots, v_{|f|/2}\} \subseteq V$ denote the set of vertices that appear as an endpoint of one of the paths. Let $V_2 \subseteq V^*$ denote the set of vertices in V^* that appear on an odd number of paths (remember that the paths in an f-covering do not need to be disjoint). Note that the equation we want to show (Equation (11)) can be rewritten as

$$z_V \oplus \bigoplus_{v \in V_2} z_v = \begin{cases} 0 & \text{if } |f| \bmod 4 = 0, \\ 1 & \text{if } |f| \bmod 4 = 2. \end{cases} \tag{12}$$

From the definition of an f-covering, we have $V_1 = \{v \in V \mid f(v) = 1\}$, and thus in Process $\mathcal{P}(G, f)$ all the vertices in V_1 are measured in the Y-basis, while the vertices in $V \setminus V_1$ and the vertices in V_2 are measured in the X-basis. The observable corresponding to this measurement is thus

$$\prod_{u \in V_1} Y_u \prod_{v \in (V \setminus V_1) \cup V_2} X_v. \tag{13}$$

Observe that

$$\prod_{v \in V} \pi_v = \prod_{v \in V} X_v.$$

This simple but crucial property follows from our construction: $\overline{G}$ is obtained from G by inserting a vertex on each edge of G. For each $i \in \{1, \ldots, |f|/2\}$ we also have

$$\prod_{v \in p_i \cap V^*} \pi_v = Z_{u_i} \left(\prod_{v \in p_i \cap V^*} X_v \right) Z_{v_i}.$$

Thus

$$\left(\prod_{v \in V} \pi_v \right) \times \left(\prod_{i=1}^{|f|/2} \prod_{v \in p_i \cap V^*} \pi_v \right) = \left(\prod_{u \in V_1} X_u Z_u \right) \times \left(\prod_{v \in V \setminus V_1} X_v \right) \times \left(\prod_{v \in V_2} X_v \right)$$
$$= (-1)^{|f|/2} \prod_{u \in V_1} Y_u \prod_{v \in (V \setminus V_1) \cup V_2} X_v.$$

When $|f| \bmod 4 = 0$ the observable of Equation (13) can then be written as a product of generators of the graph state, and thus the parity of all the measurement outcomes is 0. When $|f| \bmod 4 = 2$ the additive inverse of this observable can be written as a product of generators of the graph states, and thus the parity of all the measurement outcomes is 1. This proves Equation (12), and thus Equation (11). ◄

Remark.   The conditions of Equations (10) for all the cycles $\mathcal{C}$ of $\overline{G}$ and the condition of Equation (11) together actually completely characterize the distribution of the outcomes of $\mathcal{P}(G, f)$: the variables $\{z_v\}_{v \in V \cup V^*}$ are uniformly distributed over the set of all values satisfying all these equations. Note that when G is a connected graph then this corresponds to satisfying exactly $|E| - |V| + 2$ independent linear equations. Indeed, $|E| - (|V| - 1)$ equations suffice to guarantee that Equation (10) holds for all the cycles $\mathcal{C}$ of $\overline{G}$, as can be seen by considering a spanning tree of G: the spanning tree contains $|V| - 1$ edges and each of the remaining $|E| - (|V| - 1)$ edges gives rise to a cycle in G (and thus to a new linear equation) when added to the spanning tree. A similar characterization can be easily obtained when G is not connected as well, by considering separately each connected component.

4   Description of the Relation R

In this section we describe the computational problem we use to prove Theorem 2.

4.1   Our graph construction

For any even positive integer d, we explain how to construct two graphs $\mathcal{G}_d$ and $\overline{\mathcal{G}}_d$ that we will use to define the computational problems. The construction is illustrated in Figure 4.

The graph $\mathcal{G}_d$ is the graph with vertex set $V_d = V_d^1 \cup V_d^2$ defined as follows. We start with a $d^3 \times d^3$ square grid and denote V_d^1 the set of vertices of this grid (observe that $|V_d^1| = d^6$). This grid can be divided into d^4 contiguous square regions each of size $d \times d$. We call each region a box. In each box we place a vertex at the center of each 1×1 square and connect it to the four corners of the square. Let V_d^2 denote the set of all these new vertices. We have $|V_d^2| = d^4(d-1)^2$. It will be convenient to denote those vertices u_{ij} for $i, j \in \{1, \ldots, k\}$, where $k = d^2(d-1)$, with the index i representing the horizontal position and the index j representing the vertical position. This completes the description of $\mathcal{G}_d$.

The graph $\overline{\mathcal{G}}_d$ is obtained from $\mathcal{G}_d$ by the construction described in Section 3: one vertex is inserted on each edge of $\mathcal{G}_d$. Let V_d^* denote the introduced vertices. Note that $|V_d^*| = 2d^3(d^3 - 1) + 4d^4(d-1)^2$. Let us denote $\overline{V}_d = V_d^1 \cup V_d^2 \cup V_d^*$ the set of all vertices in $\overline{\mathcal{G}}_d$ and write

$$n = |\overline{V}_d| = \Theta(d^6).$$

For any vertex $u \in V_d$, let $\mathrm{Box}(u)$ denote the unique $d \times d$ box in which u is included. Finally, we denote $\partial(\overline{\mathcal{G}}_d)$ the external border of the graph, i.e., the perimeter of the whole grid.
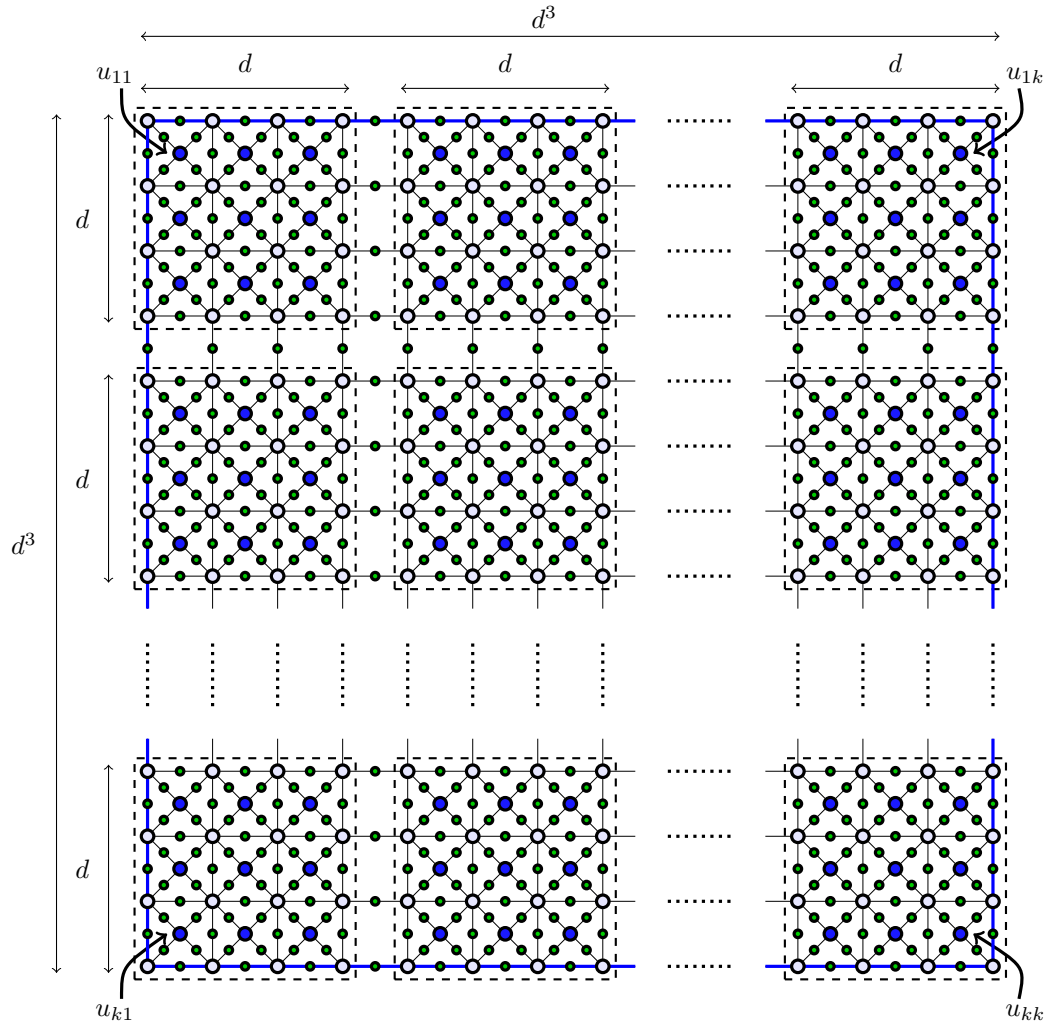
4.2   Definition of the relation

Let d, k and n be as in Section 4.1. Given a matrix $A \in \{0, 1\}^{k \times k}$, consider the process $\mathcal{P}_d(A)$ described in Figure 5.

In this process each node of $\overline{\mathcal{G}}_d$ performs a measurement and outputs one bit. We represent the whole output by a binary string of length n by fixing an arbitrary ordering of the n nodes of $\overline{\mathcal{G}}_d$. With this representation of measurement outcomes as strings, let

$$\Lambda_d(A) \subseteq \{0, 1\}^n$$

denote the set of all the strings that occur with non-zero probability in Process $\mathcal{P}_d(A)$.

Figure 4 The graph $\overline{\mathcal{G}}_d$, here represented for $d = 4$. The vertices in V_d^1 are represented in white, the vertices in V_d^2 are represented in blue and the vertices in V_d^* are represented in green. The blue line represents the external border of the graph. The dashed squares represent the boxes.

---

1. Construct the graph state over $\overline{\mathcal{G}}_d$.
2. For each vertex $u_{ij} \in V_d^2$, measure the qubit of the vertex u_{ij} in the X-basis if $A_{ij} = 0$, and measure it in the Y-basis if $A_{ij} = 1$.
   For each vertex $u \in V_d^1 \cup V_d^*$, measure the qubit of the vertex u in the X-basis.

Figure 5 The process $\mathcal{P}_d(A)$.

Definition of the relation R. For any even positive integer d, the computational problem that we consider is as follows: given a matrix $A \in \{0,1\}^{k \times k}$ as input, where $k = d^2(d-1)$, compute a string from $\Lambda_d(A)$. Note that since $|\Lambda_d(A)| > 1$ there are more than one valid output. This computational problem corresponds to the relation

$$R = \left\{ (A, z) \mid A \in \{0,1\}^{k \times k} \text{ and } z \in \Lambda_d(A) \right\} \subseteq \{0,1\}^{k \times k} \times \{0,1\}^n.$$

By setting $m = k^2$ and identifying $\{0,1\}^{k \times k}$ with $\{0,1\}^m$, we interpret R as a subset of $\{0,1\}^m \times \{0,1\}^n$. This relation R is the relation that appears in the statement of Theorem 2. To avoid confusion it will be preferable to make explicit the dependence on the parameter d. We will thus denote this relation by R_d instead of R in the next sections.

5 Proof of Theorem 2

In this section we prove Theorem 2. Let R_d be the relation defined in Section 4.2.

In the quantum setting, the computational problem corresponding to R_d can obviously be solved by directly implementing the process $\mathcal{P}_d(A)$. This can be done by a constant-depth quantum circuit since the graph $\overline{\mathcal{G}}_d$, which has constant degree, can be constructed in constant depth. Note that the description of the quantum circuit can be computed easily, e.g., by a logarithmic-space classical Turing machine.

We now show the classical lower bound, i.e., show that any classical circuit of sublogarithmic depth with bounded-fanin gates cannot output a string in $\Lambda_d(A)$ with high probability on a non-negligible fraction of the inputs A. For concreteness (and without loss of generality) we will assume in this section that all the gates in the classical circuit have fanin at most 2.

Consider any randomized classical circuit C_d, with gates of fanin at most 2, of depth at most $\frac{1}{8} \log_2 m$ for the relation R_d. The circuit has $m = k^2 = \Theta(d^6)$ input wires to receive the matrix A and n output wires. Remember that $n = \Theta(d^6)$. To simplify the presentation we assume that d is large enough so that the inequality

$$3n^{1/7} < d - 2 \tag{14}$$

holds. In Section 5.1 below we show how to associate the wires of C_d to the nodes of $\overline{\mathcal{G}}_d$. In Section 5.2 we present technical results that exploit this correspondence. Finally, in Section 5.3 we give an upper bound on the success probability of C_d and conclude the proof of Theorem 2.

5.1 Correspondence between C_d and $\overline{\mathcal{G}}_d$

We associate the wires of C_d to the nodes of $\overline{\mathcal{G}}_d$ in the following way. For any vertex $u_{ij} \in V_d^2$, we denote $x_{u_{ij}}$ the input wire of C_d that receives the entry A_{ij} of A. For any vertex $u \in \overline{V}_d$, we denote z_u the output wire of C_d that should output the outcome of the measurement performed at vertex u.

For any vertex $u \in \overline{V}_d$, we denote $L(z_u)$ the set of all vertices $v \in V_d^2$ such that the input wire x_v is in the lightcone of z_u (i.e., the value of z_u depends on the value of x_v). For any $u \in V_d^2$, we denote $L(x_u)$ the set of all vertices $v \in \overline{V}_d$ such that the output wire z_v is in the lightcone of x_u (i.e., the value of z_v depends on the value of x_u). Since the depth of C_d is at most $\frac{1}{8} \log_2 m$ and since each gate of C_d has fanin at most 2, we have $|L(z_u)| \leq m^{1/8} \leq n^{1/8}$ for each $u \in \overline{V}_d$. Let us define the set

$$\Gamma = \{u \in V_d^2 \mid L(x_u) > n^{1/7}\}.$$

Since the number of input wires is $|V_d^2| = \Theta(n)$, a simple counting argument shows that $|\Gamma| = O(n^{55/56})$, i.e., most input wires have small lightcones as well.

Define the sets $\mathcal{U}, \mathcal{V}, \mathcal{W} \subseteq V_d^2$ as follows:

$$\mathcal{U} = \{u_{ij} \mid i \in \{1, \ldots, \lfloor k/3 \rfloor\} \text{ and } j \in \{1, \ldots, \lfloor k/3 \rfloor\}\} \setminus \Gamma,$$
$$\mathcal{V} = \{u_{ij} \mid i \in \{\lceil 2k/3 \rceil, \ldots, k\} \text{ and } j \in \{1, \ldots, \lfloor k/3 \rfloor\}\} \setminus \Gamma,$$
$$\mathcal{W} = \{u_{ij} \mid i \in \{\lceil 2k/3 \rceil, \ldots, k\} \text{ and } j \in \{\lceil 2k/3 \rceil, \ldots, k\}\} \setminus \Gamma.$$

These three sets represent the vertices in $V_d^2 \cap \Gamma$ that are in the upper left part, the upper right part, and the lower right part of the graph $\overline{\mathcal{G}}_d$, respectively. From the above discussion we have $|\mathcal{U}| = \Theta(n)$, $|\mathcal{V}| = \Theta(n)$ and $|\mathcal{W}| = \Theta(n)$.

5.2 Graph-theoretic arguments

We start with a first lemma, which is similar to [9, Claim 6].

▶ **Lemma 7.** *The number of triples* $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times \mathcal{W}$ *such that the three conditions*
- $L(x_u) \cap \mathrm{Box}(v) = \emptyset$ *and* $L(x_u) \cap \mathrm{Box}(w) = \emptyset$;
- $L(x_v) \cap \mathrm{Box}(u) = \emptyset$ *and* $L(x_v) \cap \mathrm{Box}(w) = \emptyset$;
- $L(x_w) \cap \mathrm{Box}(u) = \emptyset$ *and* $L(x_w) \cap \mathrm{Box}(v) = \emptyset$.
do not simultaneously hold is $O(n^{2+10/21})$.

Proof. Observe that for each $u \in \mathcal{U}$, there are at most $n^{1/7}$ boxes that intersect $L(x_u)$. Since each box contains $(d-1)^2 = O(n^{1/3})$ vertices in V_d^2, there are at most $O(n^{10/21})$ vertices $v \in \mathcal{V}$ such that $\mathrm{Box}(v)$ intersects $L(x_u)$. Assume that we choose a vertex v uniformly at random in $\mathcal{V}$. Then we have

$$\Pr_{v \in \mathcal{V}} [L(x_u) \cap \mathrm{Box}(v) \neq \emptyset] = O\left(n^{-11/21}\right).$$

Applying the union bound shows that if we choose a triple (u, v, w) uniformly at random in $\mathcal{U} \times \mathcal{V} \times \mathcal{W}$, then the probability that this triple does not satisfy all the three conditions of the lemma is $O(n^{-11/21})$. Since $|\mathcal{U} \times \mathcal{V} \times \mathcal{W}| = \Theta(n^3)$, we thus obtain the statement of the lemma. ◀

The following simple lemma will be crucial for our analysis.

▶ **Lemma 8.** *The number of triples* $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times \mathcal{W}$ *such that the three lightcones* $L(x_u)$, $L(x_v)$ *and* $L(x_w)$ *are not pairwise disjoint is* $O(n^{2+2/7})$.

Proof. Let $t \in \overline{V}_d$ be any vertex of $\overline{\mathcal{G}}_d$. When choosing (u, v) uniformly at random in $\mathcal{U} \times \mathcal{V}$, the probability that t is in $L(x_u) \cap L(x_v)$ is $O((n^{1/7}/n)^2) = O(n^{-12/7})$. By the union bound this implies that when choosing a triple (u, v, w) uniformly at random in $\mathcal{U} \times \mathcal{V} \times \mathcal{W}$, the probability that x is in more than one of the three lightcones $L(x_u)$, $L(x_v)$ and $L(x_w)$ is $O(n^{-12/7})$ as well. By the union bound again, we conclude that when choosing (u, v, w) uniformly at random in $\mathcal{U} \times \mathcal{V} \times \mathcal{W}$, the probability that the three lightcones $L(x_u)$, $L(x_v)$ and $L(x_w)$ are not pairwise disjoint is $O(n^{-5/7})$. ◀

The following proposition is the main result of this subsection.

▶ **Proposition 9.** *There exists a triple of vertices* $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times \mathcal{W}$ *such that all the following conditions hold:*
(i) *the lightcones* $L(x_u)$, $L(x_v)$ *and* $L(x_w)$ *are pairwise disjoint;*
(ii) *there exists a cycle* $\mathcal{C}$ *containing* u, v *and* w *such that*
    (ii-a) $\mathcal{C}$ *does not use any edge from the external border* $\partial(\overline{\mathcal{G}}_d)$;
    (ii-b) $\mathcal{C} \cap V_d^2 = \{u, v, w\}$;
    (ii-c) $q_1 \cap L(x_w) = \emptyset$, $q_2 \cap L(x_u) = \emptyset$ *and* $q_3 \cap L(x_v) = \emptyset$, *where* q_1 *denotes the direct path*[4] *from* v *to* w *in the cycle* $\mathcal{C}$, q_2 *denotes the direct path from* u *to* w *in* $\mathcal{C}$ *and let* q_3 *denote the direct path from* u *to* v *in* $\mathcal{C}$.

---

[4] There are two paths from v to w in the cycle $\mathcal{C}$: one path going via u and one path not using u. The direct path is the latter.

Proof. Lemmas 7 and 8 imply that among the $\Theta(n^3)$ triples $(u, v, w) \in \mathcal{U} \times \mathcal{V} \times \mathcal{W}$ there exists one triple such that Condition (i) and the three conditions of Lemma 7 simultaneously hold. Let us fix such a triple.

Let $u_1, \ldots, u_{d-2}$ denote the vertices on the right border[5] of Box(u) and $u'_1, \ldots, u'_{d-2}$ denote the vertices on the bottom border of Box(u). Similarly, let $v_1, \ldots, v_{d-2}$ denote the vertices on the left border of Box(v) and $v'_1, \ldots, v'_{d-2}$ denote the vertices on the bottom border of Box(v). Finally, let $w_1, \ldots, w_{d-2}$ denote the vertices on the top border of Box(w) and $w'_1, \ldots, w'_{d-2}$ denote the vertices on the left border of Box(w). We refer to Figure 6 for an illustration.

We can construct a path p_i^1 from u_i to v_i, a path p_i^2 from v'_i to w_i and a path p_i^3 from w'_i to u'_i, for each $i \in \{1, \ldots, d-2\}$, so that the $3(d-2)$ paths constructed are disjoint, do not use any edge on the border $\partial(\overline{\mathcal{G}}_d)$, do not go through any vertex in V_d^2, and do not contain any vertex in Box(u) $\cup$ Box(v) $\cup$ Box(w) except their endpoint. From Inequality (14) and since the three lightcones $L(x_u)$, $L(x_v)$ and $L(x_w)$ do not have size larger than $n^{1/7}$, there necessarily exist three indices $i_1, i_2, i_3 \in \{1, \ldots, d-2\}$ such that the three paths $p_{i_1}^1$, $p_{i_2}^2$ and $p_{i_3}^3$ do not contain any vertex in $L(x_u) \cup L(x_v) \cup L(x_v)$. Finally, observe that these three paths can be completed (avoiding all vertices in $V_d^2 \setminus \{u, v, w\}$) to obtain a cycle

$$u \to u_{i_1} \xrightarrow{p_{i_1}^1} v_{i_1} \to v \to v'_{i_2} \xrightarrow{p_{i_2}^2} w_{i_2} \to w \to w'_{i_3} \xrightarrow{p_{i_3}^3} u'_{i_3} \to u$$

that satisfies Conditions (ii-a), (ii-b) and (ii-c). See Figure 6 for an illustration. Note that Condition (ii-c) can be guaranteed due to the fact that (u, v, w) satisfies the three conditions from Lemma 7.   ◀

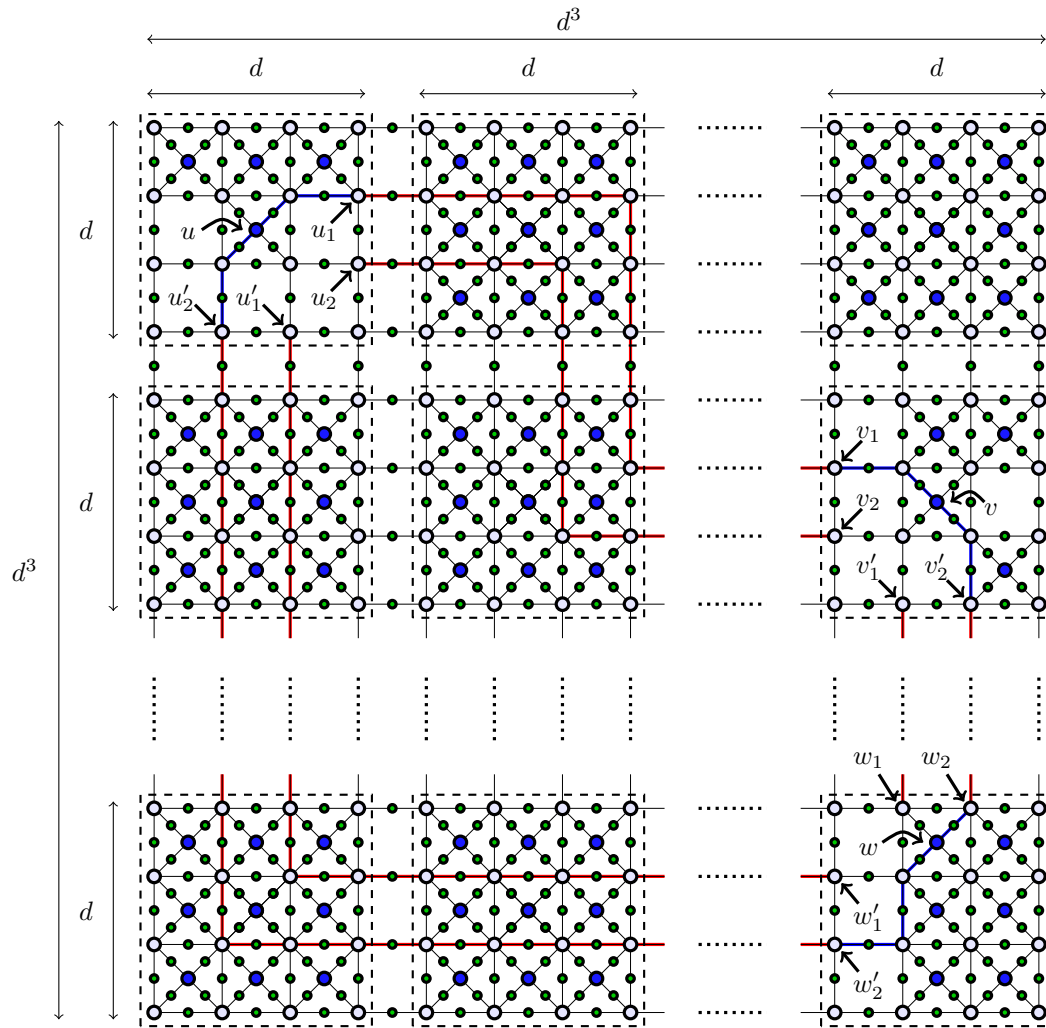## 5.3   Upper bound on the success probability

Let (u, v, w) denote the triple from $\mathcal{U} \times \mathcal{V} \times \mathcal{W}$ whose existence is guaranteed by Proposition 9. Let $\mathcal{C}$, q_1, q_2 and q_3 denote the cycle and the three paths of Condition (ii) of the proposition.

Remember that each entry A_{ij} of the input matrix A specifies the basis in which the qubit of vertex u_{ij} in the graph state $|\overline{\mathcal{G}}_d\rangle$ is measured. We will say that the vertex u_{ij} is marked if $A_{ij} = 1$. The input matrix $A \in \{0, 1\}^{k \times k}$ can then be constructed by first considering the $k^2 - 3$ entries corresponding to all vertices in $V_d^2 \setminus \{u, v, w\}$, and then specifying the entries of the three vertices u, v and w. This means that A can be represented as a pair of strings (a, b) where $a \in \{0, 1\}^{k^2 - 3}$ and $b = (b_u, b_v, b_w) \in \{0, 1\}^3$.

The randomized classical circuit C_d can be seen as a deterministic circuit receiving a random string r. Let us fix the value of this random string. Let us also fix the string $a \in \{0, 1\}^{k^2 - 3}$ and assume that the Hamming weight $|a|$ is even (note that $|a|$ corresponds to the number of marked vertices in $V_d^2 \setminus \{u, v, w\}$). The only remaining variables are thus the three bits b_u, b_v and b_w.

Observe that the graph $\overline{\mathcal{G}}_d$ remains connected when removing all the vertices on the cycle $\mathcal{C}$, due to Conditions (ii-a) and (ii-b) of Proposition 9. No vertex from $V_d^2 \setminus \{u, v, w\}$ appears in $\mathcal{C}$, from Condition (ii-b) of Proposition 9. This implies that there exists a set of $|a|/2$ paths $\{p_1, \ldots, p_{|a|/2}\}$ such that $p_i \cap \mathcal{C} = \emptyset$ for each $i \in \{1, \ldots, |a|/2\}$, and each marked vertex in $V_d^2 \setminus \{u, v, w\}$ appears once as an endpoint of one of these paths. Define

---

[5] To simplify the presentation we exclude the two corners at the extremities of each border.

Figure 6 The paths considered to construct the cycle $\mathcal{C}$ in the proof of Proposition 9 are depicted in red. The blue line show how the paths are completed to construct the cycle $\mathcal{C}$ in the case $i_1 = 1$, $i_2 = 2$ and $i_3 = 2$. Note that some vertices in V_d^2 are omitted in order to make the figure clearer.

the three bits

$$\lambda_1 = \bigoplus_{\ell \in V_d} z_\ell,$$

$$\lambda_2 = \bigoplus_{i=1}^{|a|/2} \bigoplus_{\ell \in p_i \cap V_d^*} z_\ell,$$

$$y = \begin{cases} \lambda_1 \oplus \lambda_2 & \text{if } |a| \bmod 4 = 0, \\ \lambda_1 \oplus \lambda_2 \oplus 1 & \text{if } |a| \bmod 4 = 2. \end{cases}$$

A crucial observation is that y is an affine function of b_u, b_v and b_w, due to Condition (i) of Proposition 9.

Define

$$y_1 = \bigoplus_{\ell \in q_1 \cap V_d^*} z_\ell, \qquad y_2 = \bigoplus_{\ell \in q_2 \cap V_d^*} z_\ell, \qquad y_3 = \bigoplus_{\ell \in q_3 \cap V_d^*} z_\ell.$$

Condition (i) of Proposition 9 again guarantees that y_1, y_2 and y_3 are affine functions of the three bits b_u, b_v, b_w. Moreover, Condition (ii-c) implies that y_1 does not depend on b_u, y_2 does not depend on b_v and y_3 does not depend on b_w.

Theorem 5 implies that if the output of the circuit is in the set $\Lambda_d(A)$ (i.e., the output corresponds to a valid outcome arising from the corresponding measurement of the graph state $|\overline{\mathcal{G}}_d\rangle$), then the following condition should hold:

$$y_1 \oplus y_2 \oplus y_3 = 0 \quad \text{for all } (b_u, b_v, b_w) \in \{0,1\}^3. \tag{15}$$

Theorem 6 additionally implies that if the output of the circuit is in the set $\Lambda_d(A)$ then the following condition should hold:

$$\begin{cases} y = 0 & \text{if } (b_u, b_v, b_w) = (0,0,0), \\ y \oplus y_1 = 1 & \text{if } (b_u, b_v, b_w) = (0,1,1), \\ y \oplus y_2 = 1 & \text{if } (b_u, b_v, b_w) = (1,0,1), \\ y \oplus y_3 = 1 & \text{if } (b_u, b_v, b_w) = (1,1,0). \end{cases} \tag{16}$$

Lemma 7 implies that there is at least one value for the triple (b_u, b_v, b_w) for which these conditions are not satisfied.

We have just shown that for any value of r and any value of a such that $|a|$ is even, the output of the circuit C_d is incorrect for at least a fraction $1/8$ of the strings $b = (b_u, b_v, b_w) \in \{0,1\}^3$. Since $|a|$ is even with probability $1/2$ when choosing the matrix A uniformly at random, we conclude that for any value of r the output of the circuit is incorrect for at least a fraction $1/16$ of the matrices $A \in \{0,1\}^{k \times k}$. This implies the inequality

$$\sum_{A \in \{0,1\}^{k \times k}} \Pr_r[C_d(A) \notin \Lambda_d(A)] \geq \frac{2^{k^2}}{16}.$$

and thus

$$\frac{1}{2^{k^2}} \sum_{A \in \{0,1\}^{k \times k}} \Pr_r[C_d(A) \in \Lambda_d(A)] < 15/16.$$

This concludes the proof of Theorem 2.

6   Soundness Amplification for Small-Depth Circuits

In this section we show how to obtain Theorem 1 from Theorem 2. In Section 6.1 we first present a general soundness amplification result that holds for any relation. Then in Section 6.2 we apply this result to the relation R_d of Theorem 2 in order to obtain Theorem 1.

6.1   General result

Consider any relation $\mathfrak{R} \subseteq \{0,1\}^m \times \{0,1\}^n$ for some positive integers m and n. As usual, this relation is interpreted as the following computational problem: given as input a string $x \in \{0,1\}^m$, output one string from the set $\mathfrak{R}(x) = \{z \in \{0,1\}^n \mid (x,z) \in \mathfrak{R}\}$. For any integer $t \geq 1$, now consider the following computational problem: given as input t strings $x_1, \ldots, x_t \in \{0,1\}^m$, output one element from the set $\mathfrak{R}(x_1) \times \cdots \times \mathfrak{R}(x_t)$. This computational problem corresponds to the direct product of t copies of the relation $\mathfrak{R}$. We will write this relation $\mathfrak{R}^{\times t}$ and interpret it as the subset

$$\mathfrak{R}^{\times t} \subseteq \{0,1\}^{mt} \times \{0,1\}^{nt}$$

by associating $\{0,1\}^{mt}$ with the t copies of $\{0,1\}^m$ and $\{0,1\}^{nt}$ with the t copies of $\{0,1\}^n$.

The main result of this section is the following repetition theorem, which shows that if $\mathfrak{R}$ cannot be computed with average success probability larger than $1 - \alpha$ using small-depth classical circuits, then $\mathfrak{R}^{\times t}$ cannot be computed with average success probability larger than $(1 - \alpha)^{t'}$ for some $t' \approx t$ by circuits of the same depth. The idea is to show how to extract, from the t copies of $\mathfrak{R}$ making $\mathfrak{R}^{\times t}$, at least t' copies on which the circuit acts independently.

▶ **Theorem 10.** *Let $\mathfrak{R} \subseteq \{0,1\}^m \times \{0,1\}^n$ be a relation for which the following assertion holds for some real numbers $c \geq 0$ and $\alpha \in [0, 1]$: any m-input n-output randomized circuit C with bounded-fanin gates and depth at most $c \log_2 m$ satisfies the inequality*

$$\frac{1}{2^m} \sum_{x \in \{0,1\}^m} \Pr[C(x) \in \mathfrak{R}(x)] < 1 - \alpha.$$

Let t be any integer such that $t \geq 6nm^c + 2$. Then any (mt)-input (nt)-output randomized circuit C' with bounded-fanin gates and depth at most $c \log_2 m$ satisfies

$$\frac{1}{2^{mt}} \sum_{x' \in \{0,1\}^{mt}} \Pr[C'(x') \in \mathfrak{R}^{\times t}(x')] < (1 - \alpha)^{t/(6m^c n + 2)}.$$

Proof. Consider any (mt)-input (nt)-output randomized circuit C' with gates of fanin at most 2 and depth at most $c \log_2 m$ for the relation $\mathfrak{R}^{\times t}$. For each $i \in \{1, \ldots, t\}$, let S_i denote the set of wires corresponding to the inputs of the i-th copy of $\mathfrak{R}$ in $\mathfrak{R}^{\times t}$ and T_i denote the set of wires corresponding to the outputs of the i-th copy of $\mathfrak{R}$ in $\mathfrak{R}^{\times t}$. The following claim is the crucial part of the proof.

▷ **Claim 11.** There exists a subset of indices $I \subseteq \{1, \ldots, t\}$ of size $|I| \geq \frac{t}{6nm^c+2}$ such that $L(S_i) \cap T_j = \emptyset$ for all distinct $i, j \in I$.

Proof. Define the set

$$J = \left\{ i \in \{1, \ldots, t\} \mid \sum_{x \in S_i} |L(x)| \leq 2m^c n \right\}.$$

Since the circuit C' has depth $c \log_2 m$ and its gates have fanin at most 2, we have $|L(z)| \leq m^c$ for any output wire z. Since the total number of output wires is nt, a simple counting argument shows that $|J| \geq t/2$.

Let us now construct a graph on the vertex set J as follows: two distinct vertices $i, j \in J$ are connected by an edge if and only if at least one of $L(S_i) \cap T_j \neq \emptyset$ and $L(S_j) \cap T_i \neq \emptyset$ holds. In this graph each vertex has degree at most $2m^c n + m^c n = 3m^c n$. There thus exists[6] an independent set $I \subseteq J$ of G of size

$$|I| \geq \frac{t/2}{3m^c n + 1} = \frac{t}{6m^c n + 2}.$$

This independent set is precisely the set of indices we wanted to construct. ◀

To lighten the notation we will assume that the set I from Claim 11 is $I = \{1, \ldots, \ell\}$ for some integer ℓ (with $\ell \geq \frac{t}{6m^c n+2}$). This assumption can be made without loss of generality. Claim 11 implies that when the values of all the input wires in $S_{\ell+1} \cup \cdots \cup S_t$ are fixed, then

---

[6] Here we are using a trivial result from graph theory that states that a graph of maximum degree Δ on N vertices has an independent set of size at least $N/(\Delta + 1)$.

for each $i \in \{1, \ldots, \ell\}$ the values of the output wires in T_i only depend on the values of the input wires in S_i. This implies that for any $(x_{\ell+1}, \ldots, x_t) \in \{0,1\}^{(t-\ell)m}$ the inequality

$$\frac{1}{2^{m\ell}} \sum_{x_1, \ldots, x_\ell \in \{0,1\}^m} \Pr[C'(x_1, \ldots, x_\ell, x_{\ell+1}, \ldots, x_t) \in \mathfrak{R}^{\times t}(x_1, \ldots, x_\ell, x_{\ell+1}, \ldots, x_t)]$$

$$< (1-\alpha)^\ell,$$

holds, from our assumption on the relation $\mathfrak{R}$ (since the depth of C' is at most $c \log_2 m$). Thus

$$\frac{1}{2^{mt}} \sum_{x_1, \ldots, x_t \in \{0,1\}^m} \Pr[C'(x_1, \ldots, x_t) \in \mathfrak{R}^{\times t}(x_1, \ldots, x_t)] < (1-\alpha)^\ell \leq (1-\alpha)^{t/(6m^c n + 2)},$$

as claimed. This concludes the proof of Theorem 10. ◀

6.2 Application: proof of Theorem 1

We are now able to give the proof of Theorem 1.

Proof of Theorem 1. We consider the relation $R_d \subseteq \{0,1\}^m \times \{0,1\}^n$ defined in Section 4.2 and used in Theorem 2. Remember that for this relation we have $m = \Theta(d^6)$ and $n = \Theta(d^6)$. Take the integer $t = \lceil (6nm^{1/8} + 2)^3 \rceil$ and observe that the inequality $t \geq m^{27/8}$ holds. Define $\mathcal{R} = R_d^{\times t}$. The sizes of the inputs and outputs in $\mathcal{R}$ are $M = mt$ and $N = nt$, respectively. Observe that $t \geq m^{27/8}$ implies $t \geq M^{27/35}$. Theorem 2 and then Theorem 10 with $\mathfrak{R} = R_d$ imply that there exist constants $c > 0$ and $\alpha > 0$ such that any M-input N-output randomized circuit C' with bounded-fanin gates and depth at most $c \log_2 m$ satisfies

$$\frac{1}{2^M} \sum_{x' \in \{0,1\}^M} \Pr[C'(x') \in \mathcal{R}(x')] < (1-\alpha)^{t^{2/3}} \leq (1-\alpha)^{M^{54/105}} < (1-\alpha)^{\sqrt{M}},$$

which leads to the claimed statement. ◀

── **References** ──

1   Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the 43rd ACM Symposium on Theory of Computing*, pages 333–342, 2011. `doi:10.1145/1993636.1993682`.

2   Scott Aaronson and Alex Arkhipov. BosonSampling is far from uniform. *Quantum Information & Computation*, 14(15-16):1383–1423, 2014. URL: `http://www.rintonpress.com/xxqic14/qic-14-1516/1383-1423.pdf`.

3   Scott Aaronson and Lijie Chen. Complexity-Theoretic Foundations of Quantum Supremacy Experiments. In *Proceedings of the 32nd Computational Complexity Conference*, pages 22:1–22:67, 2017. `doi:10.4230/LIPIcs.CCC.2017.22`.

4   Andris Ambainis. Understanding quantum algorithms via query complexity. In *Proceedings of the 2018 International Congress of Mathematicians*, volume 3, pages 3249–3270, 2018.

5   Jonathan Barrett, Carlton M. Caves, Bryan Eastin, Matthew B. Elliott, and Stefano Pironio. Modeling Pauli measurements on graph states with nearest-neighbor classical communication. *Physical Review A*, 75:012103, 2007. `doi:10.1103/PhysRevA.75.012103`.

6   Adam Bene Watts, Robin Kothari, Luke Schaeffer, and Avishay Tal. Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits. In *Proceedings of the 43rd ACM Symposium on Theory of Computing*, 2019. To appear.

7   Ethan Bernstein and Umesh V. Vazirani. Quantum Complexity Theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997. `doi:10.1137/S0097539796300921`.

8 Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. "Quantum Supremacy" and the Complexity of Random Circuit Sampling. In *Proceedings of the 10th Innovations in Theoretical Computer Science conference*, pages 15:1–15:2, 2019. arXiv:1803.04402. `doi:10.4230/LIPIcs.ITCS.2019.15`.

9 Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. arXiv:1704.00690 (preliminary version of [10]), 2017. `arXiv:1704.00690`.

10 Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018. `doi:10.1126/science.aar3106`.

11 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 of London A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, 2010. `doi:10.1098/rspa.2010.0301`.

12 Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations. *Physical Review Letters*, 117:080501, 2016. `doi:10.1103/PhysRevLett.117.080501`.

13 Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum circuits. *Quantum*, 1:8, 2017. `doi:10.22331/q-2017-04-25-8`.

14 Matthew Coudron, Jalex Stark, and Thomas Vidick. Trading locality for time: certifiable randomness from low-depth circuits, 2018. `arXiv:1810.04233`.

15 Edward Farhi and Aram W. Harrow. Quantum supremacy through the quantum approximate optimization algorithm, 2016. `arXiv:1602.07674`.

16 Keisuke Fujii, Hirotada Kobayashi, Tomoyuki Morimae, Harumichi Nishimura, Shuhei Tamate, and Seiichiro Tani. Impossibility of Classically Simulating One-Clean-Qubit Model with Multiplicative Error. *Physical Review Letters*, 120:200502, 2018. `doi:10.1103/PhysRevLett.120.200502`.

17 Keisuke Fujii and Shuhei Tamate. Computational quantum-classical boundary of noisy commuting quantum circuits. *Scientific Reports*, 6(25598), 2016. `doi:10.1038/srep25598`.

18 Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. Counting, fanout and the complexity of quantum ACC. *Quantum Information & Computation*, 2(1):35–65, 2002. URL: `http://portal.acm.org/citation.cfm?id=2011420`.

19 Marc Hein, Jens Eisert, and Hans J. Briegel. Multiparty entanglement in graph states. *Physical Review A*, 69:062311, 2004. `doi:10.1103/PhysRevA.69.062311`.

20 Peter Høyer and Robert Spalek. Quantum Fan-out is Powerful. *Theory of Computing*, 1(1):81–103, 2005. `doi:10.4086/toc.2005.v001a005`.

21 Attila Kondacs and John Watrous. On the Power of Quantum Finite State Automata. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, pages 66–75, 1997. `doi:10.1109/SFCS.1997.646094`.

22 François Le Gall, Harumichi Nishimura, and Ansis Rosmanis. Quantum Advantage for the LOCAL Model in Distributed Computing. In *Proceedings of the International Symposium on Theoretical Aspects of Computer Science*, pages 49:1–49:14, 2019. `doi:10.4230/LIPIcs.STACS.2019.49`.

23 Tomoyuki Morimae, Keisuke Fujii, and Joseph F. Fitzsimons. Hardness of Classically Simulating the One-Clean-Qubit Model. *Physical Review Letters*, 112:130502, 2014. `doi:10.1103/PhysRevLett.112.130502`.

24 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. `doi:10.1017/CBO9780511976667`.

25 Ran Raz and Avishay Tal. Oracle Separation of BQP and PH. *Proceedings of the 43rd ACM Symposium on Theory of Computing*, 2019. To appear.

26 Yasuhiro Takahashi and Seiichiro Tani. Collapse of the Hierarchy of Constant-Depth Exact Quantum Circuits. *Computational Complexity*, 25(4):849–881, 2016. `doi:10.1007/s00037-016-0140-0`.

27   Barbara M. Terhal and David P. DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. *Quantum Information & Computation*, 4(2):134–145, 2004. URL: `http://portal.acm.org/citation.cfm?id=2011582`.