Exponential Quantum Speed-ups are Generic

Fernando G.S.L. Brandão
Departamento de Física, Universidade Federal de Minas Gerais,
Belo Horizonte, Caixa Postal 702, 30123-970, MG, Brazil

Michał Horodecki
Institute for Theoretical Physics and Astrophysics,
University of Gdańsk, 80-952 Gdańsk, Poland

A central problem in quantum computation is to understand which quantum circuits are useful for exponential speed-ups over classical computation. We address this question in the setting of query complexity and show that for almost any sufficiently long quantum circuit one can construct a black-box problem which is solved by the circuit with a constant number of quantum queries, but which requires exponentially many classical queries, even if the classical machine has the ability to postselect.

We prove the result in two steps. In the first, we show that almost any element of an approximate unitary 3-design is useful to solve a certain black-box problem efficiently. The problem is based on a recent oracle construction of Aaronson and gives an exponential separation between quantum and classical post-selected bounded-error query complexities.

In the second step, which may be of independent interest, we prove that linear-sized random quantum circuits give an approximate unitary 3-design. The key ingredient in the proof is a technique from quantum many-body theory to lower bound the spectral gap of local quantum Hamiltonians.

I. INTRODUCTION

Quantum computation holds the promise of solving certain problems substantially faster than classical computation. The most famous example is arguably Shor’s polynomial-time quantum algorithm for factoring, a task which is believed to require exponential time in a classical computer. Other problems for which quantum algorithms appear to be give exponential speed-ups include simulating quantum systems, solving Pell’s equation, approximating the Jones polynomial, and estimating certain properties of sparse systems of linear equations. Unfortunately, the apparent computational superiority of quantum mechanics is presently only conjectural. In fact, one cannot hope to separate the class of problems solved in polynomial time by quantum and classical computation without settling major open questions in computational complexity theory.

A setting for which quantum computation is provably superior to classical is the one of query complexity (also known as decision tree complexity or black-box complexity). There one is given the ability to query a black-box function and the goal is to determine a certain property of the function. The complexity of the problem is measured by the minimum number of queries needed to determine such property. In the quantum case, one is able to query the black-box in superposition, a feature which potentially renders it more powerful than the classical one.

The first example of a black-box problem exhibiting a superpolynomial separation of quantum
and randomized classical query complexities was the recursive fourier sampling (RFS) problem of Bernstein and Vazirani \[8\]. Soon after it, Simon presented a black-box problem with an exponential quantum-classical separation \[9\]; Simon’s problem is also a good example of the usefulness of the query complexity model for the development of new algorithms: its quantum solution was both a motivation for and an important element in Shor’s quantum algorithm for factoring \[2\]. Many other oracle separations have since been found, see e.g. \[10 –15\]. In terms of complexity classes, these query complexity results show the existence of an oracle \(U\) for which \(\text{BQP}_U \neq \text{BPP}_U\).2

Having collected evidence that quantum computation is superior to randomized classical computation, it is interesting to get insight about where exactly does \(\text{BQP}\) sit in the zoo of classical complexity classes. For example, are there problems that a quantum computer can solve efficiently, but which a classical computer cannot even check a potential solution in reasonable time? This is the question whether \(\text{BQP} \subseteq \text{NP}\) and already in the seminal paper \[8\], the RFS problem was used to build an oracle \(U\) such that \(\text{BQP}_U \not\subseteq \text{NP}_U\). One can go even further and ask for an oracle for which \(\text{BQP}\) is not contained in the entire polynomial hierarchy \((\text{PH})\). In \[8\] it was conjectured that the RFS problem also gives an oracle relative to which \(\text{BQP} \not\subseteq \text{PH}\), but whether this is indeed the case remains an open question.

Recently, Aaronson \[1\] proposed an interesting new oracle problem as a candidate to put \(\text{BQP}\) outside \(\text{PH}\).3 Although the usefulness of this oracle for the \(\text{BQP} vs. \text{PH}\) question still has to be elucidated, the problem was shown to have a huge separation of quantum and classical query complexities: it can be solved by a constant number of quantum queries, while it requires exponentially many queries by a classical machine, even if we give the classical machine the – extremely powerful – ability to postselect on a given result of the computation. This is the strongest separation of quantum and classical query complexities to date. It also implies oracles relative to which \(\text{BQP} \not\subseteq \text{BPP}_\text{path}\)4 and \(\text{BQP} \not\subseteq \text{SZK}\), which supersedes all previous oracle separations for \(\text{BQP}\).

Aaronson’s problem, named \textsc{Fourier Checking}, is the following: We are given two boolean functions \(f, g : \{0, 1\}^n \rightarrow \{-1, 1\}\) with the promise that either

- \(f \) and \(g\) are chosen uniformly at random, or
- for a vector \(v \in \mathbb{R}^{2^n}\) with entries \(v_x\) drawn independently from a normal distribution of mean 0 and variance 1, the functions are chosen as \(f(x) = \text{sgn}(v_x)\) and \(g(x) = \text{sgn}(\hat{v}_x)\)5.

Here the vector \(\hat{v}\) is the Fourier transform over \(\mathbb{Z}_2^n\) of \(v\) and is given by

\[
\hat{v}_x = \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} v_y.
\]  

The task is to decide which is the case. In words, we should determine if the two functions are not correlated at all or if one of them is well correlated with the Fourier transform of the other.

The quantum algorithm proposed in \[1\] to solve the problem is particularly simple. One prepares the uniform superposition over the computational basis, queries \(f\), applies the quantum

---

2 see the complexity zoo (http://qwiki.stanford.edu/wiki/ComplexityZoo) for definitions of the standard complexity classes.
3 In \[1\] it was shown that the separation would follow from a certain generalization of the Linial-Nisan conjecture\[16\] recently settled by Braverman \[17\]. However this generalization was later falsified in \[18\].
4 Here \(\text{BPP}_\text{path}\) is defined as the class of problems which can be solved in polynomial time, with high probability, by a randomized classical computer which can postselect on given outcomes of the computation \[19\].
5 The sign function is defined as: \(\text{sgn}(x) = 1\) for \(x \geq 0\), and \(\text{sgn}(x) = -1\) otherwise.
Fourier transform (QFT), queries $g$, and checks if the final state is again in a uniform superposition over the computational basis. If the functions are independent, then there is only an exponential small chance of getting the right outcome in the final measurement, while in the case where they are correlated, this happens with constant probability.

Considering how well this problem fleshes out the superiority of quantum computation to classical, it is worthwhile to try to understand what exactly gives its strength. For instance, what is the role played by the Fourier transform, both the the definition of the problem and in the quantum algorithm solving it? Can we replace it by some other transformation? One of the goals of this paper is to shed light on these questions.

From a broader perspective, we will be concerned with the following question, central to our understanding of the computational capabilities offered by quantum mechanics: What is the set of quantum circuits which provide large quantum speed-ups? More precisely, for which quantum circuits can we construct black-box problems which are solved by the circuit with only a few queries to the black-box, but which require a large number of queries for randomized classical computation? This question is in a sense a converse to the well-studied problem of characterizing the class of black-box functions allowing for significant quantum speed-ups (see e.g. [20, 21]). While the latter deals with the determination of which computational problems are suited for quantum computing, the former contributes to the classification of which quantum algorithmic techniques are useful for solving problems efficiently.

For instance, all the early examples of quantum algorithms offering superpolynomial speed-ups [8, 9, 22, 23] were based on the quantum Fourier transform and this led to the speculation that it could be the defining aspect of quantum computation behind quantum speed-ups. Subsequently, other black-box problems showing a quantum advantage were found having no relation to the QFT [6, 15, 24], hence extending the scope of techniques for constructing quantum algorithms.

Of particular note in this context, and for this paper, is the work of Hallgren and Harrow [24] on generalizations of Bernstein and Vazirani’s RFS classical–quantum separation is built in two steps: first one construct a black-box problem requiring a constant number of quantum queries, but $\Omega(n)$ classical queries. Then one uses recursion to boost the separation to a $n^{O(1)}$ quantum versus $n^{\log(n)}$ classical queries. The oracle problem in the first part is based on the Fourier transform and solved by the QFT. In [24] it was shown that this problem could be modified to have almost any quantum circuit (from a natural measure on circuits) in the place of the Fourier transform and still achieve the constant versus linear separation, as in the original formulation. Moreover, any such problem could also be boosted by recursion to provide a black-box problem with a superpolynomial quantum-to-classical gap in query complexity.

A. Our results

In this paper we generalize Aaronson’s Fourier Checking problem [11] and show that the Fourier transform, both in the definition of the problem and in the quantum algorithm solving it, can be replaced by a large class of quantum circuits. These include both the Fourier transform over any (possibly non-abelian) finite group and almost any sufficiently long quantum circuit from a natural distribution on the set of quantum circuits, which we discuss later on. We obtain exponential separations of quantum and postselected classical query complexities for all such circuits.

Flat circuits imply exponential separation: In more detail, we first introduce a simple measure of flatness, or dispersiveness, of a unitary $U$ on $n$ qubits, denoted $C(U)$. It is defined as the minimal
min-entropy\(^6\) (over \(j \in \{0, .., 2^n – 1\}\)) of the outcome probability distribution of a computational basis measurement applied to \(U|j\). For \(N := 2^n\),

\[
C(U) := \min_{j \in [N]} \log \left( \sum_{|j|} \left( |\langle 0 | U | j \rangle|^2, |\langle N - 1 | U | j \rangle|^2 \right) \right),
\]

with \([N] := \{0, ..., N – 1\}\). It thus measures the worst-case dispersiveness of states obtained by applying \(U\) to computational basis states.

In section [II] we define, for a unitary \(U\), the black-box problem U-CIRCUIT CHECKING, a variant of FOURIER CHECKING in which the Fourier transform in the definition of the vector \(\hat{v}\) (given by Eq. (1)) is replaced by \(U\). The problem is constructed so that a quantum computer can easily solve it with access to a few realizations of the unitary \(U\), while it is classically hard for any \(U\) with large \(C(U)\).

In detail, on one hand we prove a lower bound of \(2^{Ω(C(U))}\) on the classical query complexity with postselection of U-CIRCUIT CHECKING (see section [IV]). Following the ideas of [I], we do so by showing that the discretized version of the random vector \((v, Uv)\) – for a vector \(v\) composed of independent elements \(\varepsilon_x\) each drawn from a normal distribution of mean 0 and variance \(1 – k\) is \(k^{O(1)}2^{Ω(C(U))}\)-almost \(k\)-wise independent (a fact which was shown to imply the previous exponential lower bound on the postselected classical query complexity [I]).

On the other hand, on a quantum computer we can solve U-CIRCUIT CHECKING by the following simple modification of Aaraoon’s algorithm: we prepare each qubit in the \(|+\rangle := (|0\rangle + |1\rangle)/\sqrt{2}\) state, forming the uniform superposition over the computational basis. Then we query the \(f\) function, apply the circuit \(U\), query the \(g\) function, and measure each qubit in the Hadamard basis, accepting if all of them are found in the \(|+\rangle\) state. Therefore we obtain:

**Theorem I** For any circuit \(U\) acting on \(n\) qubits with \(C(U) = Ω(n)\), the problem U-CIRCUIT CHECKING shows an exponential separation of quantum and postselected classical query complexities.

We then proceed by giving two classes of unitaries with \(C(U) = Ω(n)\).

**Theorem II**

(i) Let \(U_{QFT}(G)\) be the quantum Fourier transform over the finite group \(G\). Then \(C(U_{QFT}(G)) \geq \frac{1}{2} \log |G|\).

(ii) Given any \(2^{-3n}\)-approximate unitary \(t\)-design on \(n\) qubits, all but a \(2^{-((t-1)β-2)n+1}\) fraction of its elements have \(C(U) \geq βn\).

In particular, we find that for \(2^{-9n}\)-approximate unitary 3-designs, all but a \(2^{-n/2+1}\) fraction of its element have \(C(U) \geq n/6\). We note that the result of the theorem does not appear to hold true for unitary 2-designs and thus we seem to have the first application of unitary \(t\)-designs for \(t > 2\).

The proofs of both statements of Theorem II are elementary and are given in section [III].

**Random circuits are unitary 3-designs:** A unitary \(t\)-design is an ensemble of unitaries \(\{µ(dU), U\}\), for a measure \(µ\) on the set of unitaries, such that the average (over \(µ\)) of any \(t\)-degree polynomial on the entries of \(U\) and their complex conjugates is equal to the average over the Haar measure.

\(^6\) For a probability distribution \(p(x)\) we define its min-entropy as \(h_{min}(p) := – \log \max_x p(x)\)
An approximate unitary $t$-design is a relaxed version of the previous definition, in which we only require that the averages are close to each other (see section II C for a precise definition) [23, 26].

In a series of papers [27–31] it was established that polynomially long random quantum circuits constitute an approximate unitary 2-design. The random quantum circuit model used is the following: in each step a random pair of qubits is chosen and a gate from a universal set of gates, also chosen at random, is applied to them. Although there is evidence that random quantum circuits of polynomial length are unitary $t$-design for every $t = \text{poly}(n)$ [32, 33], this has not been rigorously proved so far, even for the 3-design case.

Here we prove that random quantum circuits are indeed approximate unitary 3-designs. We show it both for the random circuit model of the previous paragraph and for a different one, introduced in [34] as a toy model for the evolution of black holes, which is more suited for the methods we employ. In this model, which we call local random quantum circuit model, the qubits are arranged in a circle and in each step a random two-qubit gate is applied to two neighbouring qubits.

**Theorem III** 5$n$ log$(1/\varepsilon)$-size local random quantum circuits form an $\varepsilon$-approximate unitary 3-design.

The proof of Theorem III is based on a reduction, first put forward by Brown and Viola [33], connecting the convergence rate of moments of the random quantum circuit to the spectral gap (the difference of the lowest and second lowest eigenvalues) of a quantum local Hamiltonian. Our main contribution is to show in section VII that we can obtain a lower bound on this spectral gap employing a technique from quantum many-body theory used e.g. in [35–38].

In particular, using this technique we are able to reduce the problem of bounding the spectral gap of the random walk on $n$ qubits induced by the random circuit, to bounding the spectral gap of the same random walk, but now defined only on three neighbouring qubits. Then it suffices to bound the convergence time of the second and third moments of the latter random walk in order to prove that the random circuit constitute a 3-design. We believe our approach is promising also for higher values of $t$ and might pave the way to a proof that random quantum circuits are approximate unitary $t$-designs for all $t = \text{poly}(n)$. We however leave such possibility as an open problem for future work.

Combining Theorems III and II we obtain our main result that almost any polynomial quantum circuit is useful for exponential quantum speed-ups.

**Theorem IV** For the distribution induced by the local random quantum circuit model, all but a $2^{-\Omega(n)}$ fraction of quantum circuits $U$ with more than $O(n^2)$ gates are such that U-Circuit Checking shows an exponential gap in the quantum and the postselected classical query complexities.

The role of $C(U)$ and classical efficient solution for sparse unitaries: We have seen that dispersive unitaries $U$ with large $C(U)$ give an exponential speed-up in U-Circuit Checking. Is a large $C(U)$ always required for a speed-up? We present two results indicating that this is indeed the case.

First we show that with a modified notion of oracle access (we call it the independent query model), in which a different independent realization of the random parameters of the oracle is chosen in each query, a linear $C(U)$ is necessary for an exponential speed-up.

**Theorem V** In the independent query model of oracle access, the randomized classical query complexity of U-Circuit Checking is equal to $2^{\Theta(C(U))}$. 
Second we consider the circuit checking problem for \textit{approximately-sparse} $U$, defined as unitaries which can be approximated (in operator norm) by a sparse matrix with only polynomially many non-zero entries in each row and column. Then we show the following.

\textbf{Theorem VI} For \textit{approximately-sparse} $U$ the randomized classical query complexity of \textsc{U-Circuit Checking} is polynomial.

We prove Theorem VI by showing how a recent result of Van den Nest \cite{39} on the classical simulability of certain quantum states and operations implies that the quantum algorithm for \textsc{U-Circuit Checking} with a sparse $U$ can be efficiently simulated with only polynomial many classical queries to $f$ and $g$.

\textbf{B. Related Work}

This paper has a similar flavor to Hallgren and Harrow’s work on the RFS problem \cite{24}. The idea of considering the dispersiveness of quantum circuits as a resource for oracle speed-ups also first appeared in \cite{24}, where a different, but related, notion of dispersive circuits was proposed and a constant versus linear separation in query complexity was shown for all such dispersive circuits; in section II B we discuss it in more detail and show that our definition of a dispersive circuit is somewhat more demanding than theirs (although not completely comparable). In \cite{24} it was shown that both the Fourier transform over any finite group and almost any sufficiently long quantum circuit are dispersive. Implicit in their work is also the statement that most elements of an approximate unitary 2-design are dispersive. Although their definition of dispersiveness is weaker than ours and therefore broader, the separations we obtain are much stronger. While we get an exponential separation of quantum and postselected classical query complexities, they get a superpolynomial versus polynomial separation of quantum and classical query complexities, and only by using recursion (which itself can be seen as the responsible for the superpolynomial speed-up).

There has been a series of work \cite{25, 26, 30, 40, 42} on unitary $t$-designs (and on the closely related quantum expanders \cite{43–48}) and on their connection to random quantum circuits \cite{25, 41}. An important problem in this area is to derive efficient constructions on a quantum computer of approximate unitary $t$-designs. While there are several efficient constructions for 2-designs \cite{25, 41, 46, 47}, there is only a single one (based on the QFT) for unitary $t$-design on $n$ qubits with $t > 2$ (going up to $t = \Omega(n/\log(n))$) \cite{41}. Our proof that random quantum circuits constitute a 3-design gives an alternative efficient construction for the $t = 3$ case.

Recently Brown and Viola \cite{33} proposed an interesting approach to the problem of random quantum circuits as unitary $t$-designs, based on mapping the convergence time of moments of the random circuit to the spectral gap of a mean-field quantum Hamiltonian. Conditioned on an unproven, but reasonable, conjecture about the low-lying eigenstates of the Hamiltonian, they showed that random quantum circuits of linear length are $t$-designs for every fixed $t$ and sufficiently large $n$. Our approach also starts by a reduction of the problem to lower bounding the spectral gap of quantum Hamiltonians. However, in our case, we find a local quantum Hamiltonian, consisting of nearest-neighbor terms only. We are also able to rigorously lower bound such spectral gap for $t \leq 3$, therefore obtaining a complete proof in this case.

After the completion of our work, we learned about a recent paper by Fefferman and Umans \cite{49}, in which the problem \textsc{U-Circuit Checking} is also considered. Their focus is to study the usefulness of this problem in constructing an oracle separation of \textsc{BQP} and \textsc{PH}, by relating
such possibility to a conjecture \cite{50} about the capacity of the Nisan-Wigderson pseudorandom generator \cite{51} to fool \( AC_0 \). To this aim only unitaries of a very special structure are considered in U-CIRCUIT CHECKING. Our approach has the advantage that we can show an exponential gap of quantum and postselected-classical query complexities for a generic polynomial quantum circuit (a task not considered in \cite{49}), but has the drawback that we fail to give evidence that there are circuits providing a separation of BQP to PH.

II. PRELIMINARIES

A. The oracle problem and its quantum solution

Given a unitary \( U \in \mathbb{U}(N) \) (with \( \mathbb{U}(N) \) the group of \( N \times N \) unitary matrices) we consider the following extension of the FOURIER CHECKING problem \cite{1}:

**U-CIRCUIT CHECKING**: We are given access to two black-box functions\(^7\) \( f, g : \{0, 1\}^n \to \{1, -1\} \) with the promise that either

- (independent and random) \( f \) and \( g \) are chosen independently and uniformly at random, with each of their entries drawn from a random unbiased coin, or

- (\( U \)-correlated) for a vector \( v \in \mathbb{C}^N \) with entries \( v_x \) drawn independently from a complex normal distribution \( v_x = v_{x,r} + iv_{x,i} \) with \( v_{x,r} \) and \( v_{x,i} \) normal real variables of mean 0 and variance 1, the functions are chosen as \( f(x) = \text{sgn}(\text{Re}(v_x)) \) and \( g(x) = \text{sgn}(\text{Re}((Uv)_x)) \) \(^8\). The vector \( Uv \) is given explicitly by

\[
(Uv)_x = \sum_{y \in [N]} U_{xy} v_y. \tag{3}
\]

The problem is to decide which is the case.

Consider the following quantum algorithm for solving U-CIRCUIT CHECKING, where \( U \) acts on \( n \) qubits:

**Quantum Algorithm for U-CIRCUIT CHECKING**:
(i) Prepare each of the \( n \) qubits in the \( |+\rangle \) state.
(ii) Query the \( f \) oracle.
(iii) Apply the unitary \( U \).
(iv) Query the \( g \) oracle.
(v) Measure each qubit in the Hadamard basis \( \{|+\rangle, |-\rangle\} \) and accept if all qubits are in the \( |+\rangle \) state.

Let

\[
|f\rangle := \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)} |x\rangle \tag{4}
\]

\(^7\) In this work we consider the phase-oracle model for quantum queries. Namely, let \( U_f \) be the oracle unitary of \( f \) and \( |x\rangle \) a computational basis state. Then \( U_f |x\rangle = (-1)^{f(x)} |x\rangle \) (and likewise for \( g \)).

\(^8\) here \( \text{Re}(z) \) is the real part of the complex number \( z \).
and

\[ |g\rangle := \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} (-1)^{g(x)} |x\rangle \]  

Then it follows that the acceptance probability of the algorithm is given by

\[ p_U(f, g) := |\langle g | U | f \rangle|^2. \]  

The next proposition shows the quantum algorithm above can distinguish the cases of correlated (by the action of \( U \)) and independent \( f \) and \( g \).

**Proposition II.1** If \( f \) and \( g \) are drawn independently and uniformly at random,

\[ \mathbb{E}(p_U(f, g)) = \frac{1}{2^n}. \]  

while if \( f \) and \( g \) are \( U \)-correlated,

\[ \mathbb{E}(p_U(f, g)) \geq 0.07. \]  

**Proof** In [1] Aaronson proved the proposition for the case in which \( U \) is the quantum Fourier transform over \( \mathbb{Z}_2^n \), which appeared as Theorem 9 in [1]. A closer inspection at his proof shows that the only property of the quantum Fourier used is the fact that it is a unitary. Therefore the reasoning of [1] can be applied here without any modification. We omit reproducing the full argument and instead refer the reader to [1]. \( \Box \)

### B. Dispersing Circuits

We now define a notion of dispersive, or flat, circuits which will play a central role in this work. Let \( h_{\text{min}} \) be the min-entropy defined as

\[ h_{\text{min}}(p) = -\log \max_x p(x). \]  

**Definition II.2** For a unitary \( U \) we define:

\[ C(U) := \min_{j \in [N]} h_{\text{min}} \left( \{|0\rangle|U|j\rangle|^2, ..., |N - 1\rangle|U|j\rangle|^2 \right) \]  

\[ = -\log \left( \max_{i,j \in [N]} |U_{i,j}|^2 \right). \]

It is interesting to compare this definition of a dispersive circuit with Harrow and Hallgren’s [24]:

**Definition II.3** (HH-dispersiveness [24]) A unitary \( U \in \text{U}(2^n) \) is \((\alpha, \beta)\)-dispersing if there exists a set \( A \subseteq \{0,1\}^n \) with \(|A| \geq 2^{\alpha n} \) and

\[ \sum_{x \in \{0,1\}^n} |\langle a | U | x \rangle| \geq \beta 2^{\frac{n}{2}} \]  

for all \( a \in A \).
Thus while Def. II.2 looks at the infinity norm of the outcome probability distribution of measurements in the computational basis, maximized over all initial computational basis states, HH-dispersiveness (Def. II.3) is concerned with the 1-norm of such probability distribution, and the maximum taken only over a constant-size fraction of all the computational basis states.

In [24] it was shown that \((\alpha, \beta)\)-dispersing circuits, for \(\alpha, \beta = O(1)\), are useful for speed-ups in the variant of the RFS problem there defined. If we allow for lower values of \(\beta\) than a constant (but still requiring the circuit to be fairly flat), then a dispersive unitary according to Def. II.2 is also HH-dispersive. Indeed, a simple calculation (which we omit here) shows that if \(U\) is such that \(C(U) \geq \gamma n\), then \(U\) is also \((1, 2^{(\gamma-1)n/2})\)-dispersive according to Def. II.3.

C. Approximate Unitary Designs

We start defining a norm on quantum operations which we will use to compare two superoperators. For \(X \in B(\mathbb{C}^d)\) we define the \(p\)-Schatten norms \(\|X\|_p := \text{tr}(|X|^p)^{1/p}\). Then for a superoperator \(\Lambda : B(\mathbb{C}^d) \rightarrow B(\mathbb{C}^{d'})\) we define the \(p \rightarrow q\) induced Schatten norm as

\[
\|\Lambda\|_{p \rightarrow q} := \sup_{X \neq 0} \frac{\|\Lambda(X)\|_p}{\|X\|_q}.
\]

Finally, the diamond norm is defined as the CB-completion of the \(1 \rightarrow 1\) norm,

\[
\|\Lambda\|_\diamond := \sup_{d} \|I_d \otimes \Lambda\|_{1 \rightarrow 1}.
\]

There are several different definitions of \(\varepsilon\)-approximate unitary \(t\)-designs. A convenient one for us is the following.

**Definition II.4 (Approximate unitary \(t\)-design)** Let \(\{\mu, U\}\) be an ensemble of unitary operators from \(\mathbb{U}(d)\). Define

\[
G_{\mu,t}(\rho) = \int_{\mathbb{U}(d)} U^{\otimes t} \rho(U^{\dagger})^{\otimes t} \mu(dU)
\]

and

\[
G_{H,t}(\rho) = \int_{\mathbb{U}(d)} U^{\otimes t} \rho(U^{\dagger})^{\otimes t} \mu_{H}(dU),
\]

where \(\mu_{H}\) is the Haar measure. Then the ensemble is a \(\varepsilon\)-approximate unitary \(t\)-design if

\[
\|G_{\mu,t} - G_{H,t}\|_{2 \rightarrow 2} \leq \varepsilon.
\]

The following Lemma from [52] shows that the previous notion of an approximate unitary design implies two others, which will also be relevant in this work.

**Lemma II.5** (Lemma 2.2.14 of [52]). Let \(\{\mu, U\}\) be an \(\varepsilon\)-approximate unitary \(t\)-design on \(\mathbb{U}(d)\) according to Def. II.4. Then

(a) For \(G_{\mu,t}\) and \(G_{H,t}\) given by Def. II.4

\[
\|G_{\mu,t} - G_{H,t}\|_{\diamond} \leq d^k \varepsilon.
\]

(b) For every balanced monomial \(M = U_{p_1 q_1} \ldots U_{p_k q_k} U_{r_1 s_1}^* \ldots U_{r_k s_k}^*\) of degree \(k \leq t\),

\[
|E_{U \sim \mu}(M(U)) - E_{U \sim \mu_{H}}(M(U))| \leq d^t \varepsilon.
\]
III. FAMILIES OF FLAT UNITARIES

In this section we prove Theorem III showing two examples of families of unitaries which are highly dispersing.

Quantum Fourier Transforms: Let $G$ be a finite group with irreducible unitary representations \( \{ V_\lambda \}_{\lambda \in \hat{G}} \), where $r_\lambda(g)$ is the unitary matrix representation of the group element $g \in G$ in the irrep $V_\lambda$ and $\hat{G}$ labels all inequivalent irreps of $G$. Let also \( \{|g\rangle\}_g \in G \) be an orthogonal basis for $\mathbb{C}^{|G|}$. The quantum Fourier transform over $G$ is given by

\[
U_{\text{QFT}}(G) = \sqrt{\frac{\dim V_\lambda}{|G|}} \sum_{g \in G} \sum_{\lambda \in \hat{G}} \dim V_\lambda \sum_{i,j=1}^d r_\lambda(g)_{ij} |\lambda, i, j\rangle \langle g|.
\]  

(20)

We now prove the first part of Theorem III.

Proof (Theorem III part (i)) The statement is a simple application of the following basic relation, valid for any finite group [53]:

\[
\sum_{\lambda \in \hat{G}} \dim(V_\lambda)^2 = |G|.
\]  

(21)

Indeed, Eq. (21) implies $\dim V_\lambda \leq |G|^{1/2}$ for every $\lambda \in \hat{G}$ and thus

\[
|\langle \lambda, i, j | U_{\text{QFT}}(G) | g \rangle|^2 = \frac{\dim V_\lambda}{|G|} |r_\lambda(g)_{ij}|^2 \leq |G|^{-1/2},
\]  

which implies $C(U_{\text{QFT}}(G)) \geq \log_2 |G|/2$. \[\square\]

Unitary 3-designs:

Lemma III.1 For every $\varepsilon$-approximate t-design \( \{ \mu(dU), U \} \) on $U(d)$,

\[
\Pr_{U \sim \mu} (C(U) \leq \nu) \leq d^2 2^{\nu t} \left( d^{-t!} + d^{2t} \varepsilon \right).
\]  

(23)

Proof By Markov’s inequality

\[
\Pr_{U \sim \mu} (|\langle i | U_j | j \rangle|^2 \geq \lambda) = \Pr_{U \sim \mu} (|\langle i | U_j | j \rangle|^{2t} \geq \lambda^t) \leq \frac{\mathbb{E}_{U \sim \mu} (|\langle i | U_j | j \rangle|^{2t})}{\lambda^t}.
\]  

(24)

Since \( \{ \mu(dU), U \} \) is an $\varepsilon$-approximate unitary $t$-design, Lemma II.5 gives

\[
\mathbb{E}_{U \sim \mu} (|\langle i | U_j | j \rangle|^{2t}) \leq \mathbb{E}_{U \sim \mu_H} (|\langle i | U_j | j \rangle|^{2t}) + d^{2t} \varepsilon.
\]  

(25)

We have

\[
\mathbb{E}_{U \sim \mu_H} (|\langle i | U_j | j \rangle|^{2t}) = \mathbb{E}_{U \sim \mu_H} \left( \langle (\langle i | U_j | j \rangle | U^\dagger | i \rangle)^{\otimes t} \right) = \text{tr} \left( \int_U \mu_H(dU) U^{\otimes t} |j\rangle \langle j|^\otimes(U^\dagger)^{\otimes t} (|i\rangle \langle i|)^{\otimes t} \right) = \left( \frac{d+t-1}{t} \right)^{-1} \text{tr} \left( P_{\text{sym},t} (|i\rangle \langle i|)^{\otimes t} \right) = \left( \frac{d+t-1}{t} \right)^{-1},
\]  

(26)
for $P_{\text{sym},t}$ the projector onto the $(d^{t-1})$-dimensional symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$.

Then, from Eqs. (24), (25), (26) and the union bound,

$$\Pr_{U \sim \mu} \left( \max_{i,j \in [d]} |\langle i|U|j \rangle|^2 \geq \lambda \right) \leq \frac{d^2}{\lambda^t} \left( \frac{(d + t - 1)}{t} \right)^{-1} + d^2 \varepsilon. \quad (27)$$

Now we set $\lambda = 2^{-\nu}$, use the bound

$$\left( \frac{d + t - 1}{t} \right) = \frac{(d + t - 1) \ldots (d + 1) d}{t!} \geq \frac{d^t}{t!}, \quad (28)$$

and we are done. $\square$

**Proof** (Theorem II part (ii))

Let $\{\mu(dU), U\}$ be a $2^{-3n}$-approximate unitary $t$-design on $\mathbb{U}(2^n)$. Then applying Lemma III.1 with $\nu = \beta n$, $\Pr_{U \sim \mu} (C(U) \leq \beta n) \leq 2^{-\beta nt} (1 + t!)$. $\square$

### IV. CLASSICAL LOWER BOUNDS

In this section we prove an exponential lower bound on the postselected classical query complexity of U-CIRCUIT CHECKING for dispersive circuits. Following [1], our strategy will be to show that the distribution in the U-CIRCUIT CHECKING problem in the case of U-correlated strings is approximately $k$-wise independent. Then the result follows from the following proposition from [1], relating this property to bounds on the postselected query complexity of distinguishing such distribution from the uniform one:

**Proposition IV.1** (Lemma 20 of [1]) Suppose a probability distribution $\mathcal{D}$ over oracle strings is $\delta$-almost $k$-wise independent. Then no bounded-error postselected classical machine running in less than $k$ steps can distinguish $\mathcal{D}$ from the uniform distribution with bias larger than $2\delta$.

**Proof** (Theorem II) Propositions IV.3 and IV.1 give a lower bound of $2^{C(U)/7}$ on the classical query complexity with postselection of U-CIRCUIT CHECKING. Together with the $O(1)$ queries quantum algorithm for the problem from section IIIA implies Theorem II. $\square$

For a string $\{x_1, \ldots, x_M\} \subseteq \{-1, 1\}^M$ we call a term of the form $\frac{1 + x_i}{2}$ a literal and define a $k$-term as a product of $k$-literals, which equals 1 if all the literals are 1 and 0 otherwise. Then an approximate $k$-wise independent distribution is defined as follows.

**Definition IV.2** A distribution $\mathcal{D}$ over $\{-1, 1\}^M$ is $\varepsilon$-almost $k$-wise independent if for every $k$-term $C$,

$$\frac{1 - \varepsilon}{2^k} \leq \Pr_{C \sim \mathcal{D}} (C) \leq \frac{1 + \varepsilon}{2^k}. \quad (29)$$

In words, $\mathcal{D}$ is $\varepsilon$-almost $k$-wise independent if the probability of every $k$-term is $\varepsilon$-multiplicatively close to its value on the uniform distribution (which is simply $2^{-k}$).

Consider the vector $\omega_{v,U} \subseteq \{-1, 1\}^{2N}$ given by

$$\omega_{v,U} := (\text{sgn}(v_1), \ldots, \text{sgn}(v_N), \text{sgn}(\text{Re}(Uv)_1), \ldots, \text{sgn}(\text{Re}(Uv)_N)) \quad (30)$$

and let $\mathcal{D}_U$ be the distribution over $\omega_{v,U}$ when the vector $v := (v_1, \ldots, v_N)$ is composed of independent entries $v_k$, each drawn from a complex normal distribution $v_k = u_{k,r} + iu_{k,i}$ with $u_{k,r}$, $u_{k,i}$ real normal variables of mean 0 and variance 1. Then we have
Proposition IV.3. \( D_U \) is \((6k^32^{-C(U)/2})\)-almost \(k\)-wise independent.

Proof. Define

\[
    z_j = \begin{cases} 
        v_{s(j)} & \text{if } 0 \leq j \leq m \\
        (Uv)_{r(j)} & \text{if } m < j \leq k.
    \end{cases}
\]

(31)

for injective functions \( s: [m] \to [N] \) and \( r: [k - m] \to [N] \). Consider the following probability

\[
    P := \Pr (\text{sgn}(\text{Re}(z_1)) = a_1, \text{sgn}(\text{Re}(z_2)) = a_2, \ldots, \text{sgn}(\text{Re}(z_k)) = a_k),
\]

(32)

for a tuple \( a := (a_1, \ldots, a_k) \in \{-1, 1\}^k \). In the remainder of the proof we show that the probability in Eq. (32) is \((6k^32^{-C(U)/2})\)-multiplicatively close to \(2^{-k}\) for every choice of the tuple \( a \), functions \( s, r, \) and integer \( m \leq k \), which readily implies the statement of the proposition.

First we note that the probability of Eq. (32) is equal to

\[
    P = \Pr (a_1 \text{Re}(z_1) \geq 0, a_2 \text{Re}(z_1) \geq 0, \ldots, a_k \text{Re}(z_k) \geq 0)
\]

(33)

and that \( y := (a_1z_1, \ldots, a_kz_k) \) is a (circular symmetric complex) multivariable normal distribution with mean zero. That is, there is a matrix \( M \in \mathbb{C}^{k \times N} \) such that \( y := Mv \), where \( v \in \mathbb{R}^N \) is a vector of independent normal variables of mean 0 and variance 1.

It is a standard fact of multivariate normal distributions that they are completely specified by the mean vector \( \mu \) and the covariance matrix \( \Sigma \) of the distribution, i.e. the probability density function of \( y \) is given by

\[
    \frac{1}{\pi^{k/2} |\Sigma|^{k/2}} \exp \left(-y^\dagger \Sigma^{-1} y + \mu^\dagger y\right).
\]

(34)

In our case \( \mu = 0 \), while the covariance matrix \( \Sigma_{i,j} = \mathbb{E}(y_i^*y_j) \) is given by

\[
    \Sigma = \begin{pmatrix} I_m & Q \\ Q^\dagger & I_{k-m} \end{pmatrix},
\]

(35)

with

\[
    Q_{i,j} = a_ia_j \mathbb{E}(v_{s(i)}(Uv)_{r(j)}) = a_ia_jU_{s(i)r(j)}.
\]

(36)

Thus \( \|\Sigma - I\|_\infty \leq k^2 \max_{i,j \in [N]} |U_{i,j}| \leq k^22^{-\frac{1}{2}C(U)} \), which implies

\[
    (1 - k^22^{-\frac{1}{2}C(U)})I \leq \Sigma \leq (1 + k^22^{-\frac{1}{2}C(U)})I
\]

(37)

and

\[
    (1 + k^22^{-\frac{1}{2}C(U)})^{-1}I \leq \Sigma^{-1} \leq (1 - k^22^{-\frac{1}{2}C(U)})^{-1}I.
\]

(38)

For \( k \) much smaller than \( 2^{C(U)} \) we thus see that the covariance matrix \( \Sigma \) is close to the identity, which means that the distribution over \( y \) is close to the uniform. In the rest of the proof we make this observation quantitative.
We have
\[
P = \Pr \left( \Re(y_1) \geq 0, \Re(y_2) \geq 0, \ldots, \Re(y_k) \geq 0 \right)
\]
\[
= |\Sigma|^{-k} \pi^{-k} \int_{\Re(y_1) \geq 0, \ldots, \Re(y_k) \geq 0} \exp \left( -y^\dagger \Sigma^{-1} y \right) dy_1 \ldots dy_k
\]
\[
\leq |\Sigma|^{-k} \pi^{-k} \int_{\Re(y_1) \geq 0, \ldots, \Re(y_k) \geq 0} \exp \left( -(1 + k^2 2^{-\frac{1}{2} C(U)})^{-1} y^\dagger y \right) dy_1 \ldots dy_k
\]
\[
\leq \left( \frac{1 + k^2 2^{-\frac{1}{2} C(U)}}{1 - k^2 2^{-\frac{1}{2} C(U)}} \right)^k 2^{-k},
\]
where the two inequalities in the Eq. (39) above follow from the two sides of Eq. (38). Then using the bound
\[
\left( \frac{1 + a}{1 - a} \right)^k \leq 1 + 6ka,
\]
valid for all \(0 \leq a \leq 1\), we get
\[
P \leq (1 + 6k^3 2^{-C(U)/2}) 2^{-k}.
\]

By a completely similar argument (using again the two sides of Eq. (38) and Eq. (40)) we also find
\[
P \geq \left( \frac{1 - k^2 2^{-\frac{1}{2} C(U)}}{1 + k^2 2^{-\frac{1}{2} C(U)}} \right)^k 2^{-k} \geq (1 - 6k^3 2^{-C(U)/2}) 2^{-k},
\]
and we are done.

## V. CLASSICAL UPPER BOUNDS

In this section we prove Theorems \[\text{V}\] and \[\text{VI}\].

**Proof (Theorem \[\text{V}\])**

Let \((i, j)\) be such that \(|U_{i,j}| = 2^{-C(U)/2}\). Assuming that \(\Re(U_{i,j}) \geq \Im(U_{i,j})\) \(^9\), we get \(\Re(U_{i,j}) \geq 2^{-C(U)/2 + 1}\).

The algorithm for U-CIRCUIT CHECKING in the independent query model works as follows: One queries \(f_r(i)\) and \(g_r(j)\) over \(N = \lceil \Re(U_{i,j})^{-2} \log(1/\epsilon) \rceil\) independent realizations of the oracle (labelled by \(r\)) and computes
\[
E_N := \frac{1}{N} \sum_{r=1}^{N} f_r(i) g_r(j),
\]
deciding that the functions are U-correlated if \(E_N \geq \Re(U_{i,j})/4\).

\(^9\) Otherwise we consider instead the unitary \(\sqrt{-1} U\).
In the case of independent $f$ and $g$, we have $\mathbb{E}(f(i)g(j)) = 0$. In the remainder of the proof we show that for $U$-correlated $f$ and $g$, $\mathbb{E}(f(i)g(j)) \geq \text{Re}(U_{i,j})/2$. Then Chernoff bound gives that the algorithm fails with probability at most $\epsilon$.

Let us turn to the lower bound on $\mathbb{E}(f(i)g(j))$ in the case of $U$-correlated $f$ and $g$. We have

$$\mathbb{E}(f(i)g(j)) = \mathbb{E}(\text{sgn}(\text{Re}(v_i))\text{sgn}(\text{Re}(U_v)_{j})) .$$

(44)

Note that $\mathbb{E}(\text{Re}(v_i)\text{Re}(U_v)_{j}) = \text{Re}(U_{i,j}) \geq 2^{-(C(U)/2+1)}$. So all we have to do is to check that the discretized version given by Eq. (44) has a similar expectation value.

First we write

$$\mathbb{E}(\text{sgn}(\text{Re}(v_i))\text{sgn}(\text{Re}(U_v)_{j})) = \Pr(\text{Re}(v_i)\text{Re}(U_v)_{j} \geq 0) - \Pr(\text{Re}(v_i)\text{Re}(U_v)_{j} < 0)$$

$$= 2 \Pr(\text{Re}(v_i)\text{Re}(U_v)_{j} \geq 0) - 1$$

$$= 4 \Pr(\text{Re}(v_i) \geq 0 \text{ and } \text{Re}(U_v)_{j} \geq 0) - 1 .$$

(45)

Now define the bivariate normal variable $w := (w_1, w_2)$ with $w_1 = v_i$ and $w_2 := (U_v)_{j}$. The probability distribution of $w$ is completely characterized by the covariance matrix with entries $\Sigma_{k,l} := \mathbb{E}(w_k w_l^*)$, which reads

$$\Sigma = \begin{pmatrix} 1 & U_{i,j} \\ U_{i,j}^* & 1 \end{pmatrix} .$$

(46)

Then,

$$\Pr(\text{Re}(v_i) \geq 0 \text{ and } \text{Re}(U_v)_{j} \geq 0)$$

$$= |\Sigma|^{-2} \pi^{-2} \int_{\text{Re}(w_1) \geq 0, \text{Re}(w_2) \geq 0} \exp\left(-w^\dagger \Sigma^{-1} w\right) dw_1 dw_2$$

$$= |\Sigma|^{-2} \pi^{-2} \int_{\text{Re}(w_1) \geq 0, \text{Re}(w_2) \geq 0} \exp\left(-\pi^{-1} |U_{i,j}|^{-1} w^\dagger w + 2 \text{Re}(w_1^* w_2 U_{i,j})\right) dw_1 dw_2$$

$$\geq |\Sigma|^{-2} \pi^{-2} \int_{\text{Re}(w_1) \geq 0, \text{Re}(w_2) \geq 0} \exp\left(-\pi^{-1} |U_{i,j}|^{-1} w^\dagger w\right) \left(1 + 2 \text{Re}(w_1^* w_2 U_{i,j})\right) dw_1 dw_2$$

$$= \pi^{-2} \int_{\text{Re}(w_1) \geq 0, \text{Re}(w_2) \geq 0} \exp\left(-w^\dagger w\right) \left(1 + 2 \text{Re}(w_1^* w_2 U_{i,j})(1 - |U_{i,j}|^2)\right) dw_1 dw_2$$

$$= \frac{1}{4} \left(1 + \frac{\text{Re}(U_{i,j})(1 - |U_{i,j}|^2)}{8}\right) .$$

(47)

where we used basic facts of Gaussian integrals and that $e^x \geq 1 + x$, for $x \geq 0$.

We now turn to the proof of Theorem [VI] which is largely based on recent techniques of Van den Nest for the efficient classical simulation of certain types of quantum states and operations [39]. We will make use of the notion of computational tractable states, which are defined below in a slightly more general way than in [39], in order to accommodate for oracle queries.

**Definition V.1** [39] A state on $n$ qubits $|\psi\rangle$ is $f$-computational tractable given access to the oracle function $f_\psi : \{0, 1\}^m \rightarrow \{0, 1\}$ (with $m = \text{poly}(n)$) if the following holds:

(a) it is possible to sample from the probability distribution $\text{Pr}(x) := |\langle x | \psi \rangle|^2$ on the set of $n$-bit strings in $\text{poly}(n)$ time in a classical computer with $\text{poly}(n)$ many queries to $f_\psi$, and
(b) upon input of any bit of \( x \), the coefficient \( \langle x | \psi \rangle \) can be computed in \( \text{poly}(n) \) time on a classical computer with \( \text{poly}(n) \) queries to \( f_{\psi} \).

Then we have:

**Proposition V.2** \[33\] Let \( |\psi\rangle \) and \( |\phi\rangle \) be \( f \)-computational tractable states (given access to oracle \( f \)) of \( n \) qubits each and let \( A \) be an efficiently computable sparse \( n \)-qubit operation with \( \|A\|_\infty \leq 1 \). Then there exists an efficient classical algorithm to approximate \( \langle \phi | A | \psi \rangle \) with polynomial accuracy in \( n \), given access to the oracle \( f \).

In \[39\] Van den Nest proved Proposition V.2 in the non-oracular case. However, it is easy to check that his proof carries through without any modification to cover the statement of Proposition V.2.

**Proof (Theorem VI)**

Let \( \tilde{U} \) be the sparse approximation of \( U \) with only \( \text{poly}(n) \) many non-zero entries in each row and column and such that \( \|U - \tilde{U}\|_\infty \leq 0.03 \). Let \( |f\rangle \) and \( |g\rangle \) be the states given by Eqs. (4) and (5). Note that \( |f\rangle \) and \( |g\rangle \) are \( f \)- and \( g \)-computational tractable, respectively. Indeed any of their coefficients can be read directly from the oracles \( f \) and \( g \), while the probability distribution \( \text{Pr}(x) := |\langle x | f \rangle|^2 \) (and analogously \( |\langle x | g \rangle|^2 \)) is uniform and therefore easily samplable.

From Proposition V.1 we see it suffices to calculate \( |\langle g | U | f \rangle|^2 \) to accuracy \( < 0.07 \) in order to solve \( \text{U-CIRCUIT CHECKING} \) with high probability and, indeed,

\[
|\langle g | U | f \rangle|^2 - |\langle g | \tilde{U} | f \rangle|^2 | \\
\leq |\langle f | \left( U | g \rangle \langle g | U^\dagger - \tilde{U} | g \rangle \langle g | \tilde{U}^\dagger \right) | f \rangle| \\
\leq \left\| U | g \rangle \langle g | U^\dagger - \tilde{U} | g \rangle \langle g | \tilde{U}^\dagger \right\|_1 \\
= \left\| U | g \rangle - \tilde{U} | g \rangle \right\|_1 \cdot \left\| U | g \rangle + \tilde{U} | g \rangle \right\|_1 \\
\leq 2 \left\| U - \tilde{U} \right\|_\infty \leq 0.06. \tag{48}
\]

\( \square \)

**VI. RANDOM CIRCUITS ARE APPROXIMATE 3-DESIGNS**

In this section we prove that random quantum circuits of linear length form an approximate unitary 3-design, which is the main technical contribution of the paper.

We consider two classes of random quantum circuits, both defined as random walks on \( \mathbb{U}(2^n) \):

- **uniform random circuit**: in each step two indices \( i \neq j \) are chosen uniformly at random from \([n]\) and a two-qubit unitary gate \( U_{i,j} \) drawn from the Haar measure on \( \mathbb{U}(4) \) is applied to qubits \( i \) and \( j \).

- **local random circuit**: in each step of the walk an index \( i \) is chosen uniformly at random from \([n]\) and a two-qubit gate \( U_{i,i+1} \) drawn from the Haar measure on \( \mathbb{U}(4) \) is applied to the two neighbouring qubits \( i \) and \( i + 1 \) (we arrange the qubits on a circle, so we identify the \((n + 1)\)-th qubit with the first).

Throughout this section we will focus on local random circuits and then show how our results can be extended to uniform random circuits.
We will make use the following well-known correspondence of superoperators and operators, which allow us to evaluate the eigenvalues of the former by computing the eigenvalues of the latter. Given a superoperator $G$ given by

$$G(X) := \sum_k A_k X B_k^\dagger,$$

we define the operator

$$G := \sum_k A_k \otimes B_k,$$

with $\overline{X}$ the complex conjugate of $X$.

Let $X$ be such that $\text{tr}(X X^\dagger) = 1$ and $G(X) = \lambda X$, for a complex number $\lambda$, i.e. $X$ is an eigenoperator of $G$ with eigenvalue $\lambda$. Then defining $|X\rangle := X \otimes I |\Phi\rangle$, with

$$|\Phi\rangle := \sum_k |k\rangle \otimes |k\rangle,$$

it holds that $G|X\rangle = \lambda |X\rangle$, i.e. $|X\rangle$ is an eigenvector of $G$ with eingenvalue $\lambda$.

A direct implication of this correspondence is that

$$\|G\|_{2 \to 2} = \|G\|_{\infty}.$$

Proof (Theorem III)

Let $\{\mu(dU), U\}$ be the distribution of unitaries after one step of the random walk according to the local random circuit model. Following Eq. (49), define

$$G_{\mu^k,t} := \int_{U(4)} \mu^k(dU) U^{\otimes t} \otimes \overline{U}^{\otimes t}.$$

where $\mu^k$ is the $k$-fold convolution of $\mu$ with itself, i.e.

$$\mu^k := \int_{U_1 \cdots U_k} \mu(dU_1) \cdots \mu(dU_k).$$

In the sequel we show that for $t = 2, 3$,

$$\|G_{\mu^k,t} - G_{\mu_H,t}\|_{\infty} \leq \left(1 - \frac{1}{5n}\right)^k,$$

with $\mu_H$ is the Haar measure on $U(2^n)$. Then by Def. [11A] and Eq. (52) we find that $\{\mu^{5n \log(1/\epsilon)}(dU), U\}$ is an $\epsilon$-approximate unitary 3-design, which is the statement of the theorem.

Following [33], we have $G_{\mu^k,t} = (M_{t,n})^k$, with

$$M_{t,n} := \frac{1}{n} \sum_{i=1}^n P_{i,i+1}$$

and

$$P_{i,i+1} := \int_{U(4)} \mu_H(dU) U_{i,i+1}^{\otimes t} \otimes \overline{U}_{i,i+1}^{\otimes t}.$$
Moreover the projector onto the eigenvalue-one subspace of $M_t$ is equal to $G_{\mu,t}$, since \[30, 33\]
\[
\lim_{k \to \infty} \|G_{\mu^k,t} - G_{\mu,t}\|_{\infty} = 0.
\]
(58)

Therefore
\[
\|G_{\mu^k,t} - G_{\mu,t}\|_{\infty} = (\lambda_2(M_{t,n}))^k,
\]
(59)

with $\lambda_k(X)$ the $k$-th largest eigenvalue of $X$.

Lemma VI.1 and Lemma VI.4 gives that $\lambda_2(M_{t,n}) \leq 1 - \frac{1}{50}$. Then Eq. (55) is a consequence of this bound and Eq. (59). 

The next proposition is the key part of the argument. It shows that in order to upper bound $\lambda_2(M_{t,n})$ it is enough to obtain a sufficiently strong upper bound on $\lambda_2(M_{t,3})$. The latter is associated to the convergence time of a random walk on only three qubits and, therefore, can be more easily analysed.

**Lemma VI.1**

\[
\lambda_2(M_{t,n}) \leq 1 - \frac{3 - 4\lambda_2(M_{t,3})}{n}.
\]
(60)

**Proof** Define $H_{i,i+1} := I - P_{i,i+1}$ and
\[
H := \sum_{i=1}^{n} H_{i,i+1} = n(I - M_{t,n}).
\]
(61)

The operator $H$ is a quantum local Hamiltonian (i.e. a sum of terms which act non-trivially only on neighbouring sites), composed of local projectors $H_{i,i+1}$, with the following properties:

- **periodic boundary conditions**: the $(i + 1)$-th site is identified with the first.
- **zero ground-state energy**: $\lambda_{\min}(H) = 0$, with $\lambda_{\min}(H)$ the minimum eigenvalue of $H$.
- **frustration-freeness**: every state $|\psi\rangle$ in the groundstate manifold, composed of all eigenvectors with eigenvalue zero, is such that $H_{i,i+1}|\psi\rangle = 0$, for all $i$.

Let $\Delta(X)$ be the **spectral gap** of $X$, i.e. the difference of the second lowest to the lowest eigenvalues. Then from Lemma VI.2 we get
\[
\lambda_2(M_{t,n}) = 1 - \frac{\Delta(H)}{n} \\
\leq 1 - \frac{2\Delta(H_{1,2} + H_{2,3}) - 1}{n} \\
= 1 - \frac{4\Delta(I - M_{t,3}) - 1}{n} \\
= 1 - \frac{3 - 4\lambda_2(M_{t,3})}{n}.
\]
(62)

The next lemma appeared e.g. in \[35\].
Lemma VI.2 Let $H = \sum_{i=1}^{N} H_{i,i+1}$ be a local Hamiltonian with periodic boundary conditions, with $H_{i,i+1}$ projectors and $\lambda_{\min}(H) = 0$. Then the spectral gap of $H$ satisfies

$$\Delta(H) \geq 2 \min_{i,\ldots,N} \Delta (H_{i,i+1} + H_{i+1,i+2}) - 1$$

(63)

Proof Let $\gamma := 2 \min_{i,\ldots,N} \Delta (H_{i,i+1} + H_{i+1,i+2}) - 1$. Then by Lemma VI.3

$$(H_{i,i+1} + H_{i+1,i+2})^2 \geq \gamma + \frac{1}{2} (H_{i,i+1} + H_{i+1,i+2}).$$

(64)

Rearranging terms in the equation above we get

$$\frac{1}{2}H_{i,i+1} + H_{i,i+1}H_{i+1,i+2} + H_{i+1,i+2}H_{i,i+1} + \frac{1}{2}H_{i+1,i+2} \geq \frac{\gamma}{2} (H_{i,i+1} + H_{i+1,i+2}).$$

(65)

We have

$$H^2 \geq \sum_{i=1}^{N} \left( \frac{1}{2}H_{i,i+1} + H_{i,i+1}H_{i+1,i+2} + H_{i+1,i+2}H_{i,i+1} + \frac{1}{2}H_{i+1,i+2} \right)$$

$$\geq \gamma \sum_{i=1}^{N} \left( \frac{1}{2}H_{i,i+1} + \frac{1}{2}H_{i+1,i+2} \right) = \gamma H,$$

(66)

where the first inequality follows from

$$H^2 = \sum_{i=1}^{N} \left( \frac{1}{2}H_{i,i+1} + H_{i,i+1}H_{i+1,i+2} + H_{i+1,i+2}H_{i,i+1} + \frac{1}{2}H_{i+1,i+2} \right)$$

$$+ \sum_{|k-l|>1} H_{k,k+1} \otimes H_{l,l+1}.$$  

(67)

and the positivity of $H_{i,i+1}$, and the second one from Eq. 66. Then, by Lemma VI.3, $\Delta(H) \geq \gamma$. □

Lemma VI.3 For a positive semi-definite matrix $M$ with $\lambda_1(M) = 0$,

$$\lambda_2(M) = \max_{\gamma \in \mathbb{R}} \gamma : M^2 \geq \gamma M$$

(68)

Proof Let $(0, \lambda_2, \lambda_3, \ldots)$ be the vector of ordered eigenvalues of $M$. Then $(0, \lambda_2^2, \lambda_3^2, \ldots)$ are the ordered eigenvalues of $M^2$ and $M^2 \geq \gamma M$ holds true if, and only if, $\lambda_k^2 \geq \gamma \lambda_k$ for all $k$, which is equivalent to $\lambda_2 \geq \gamma$. □

The second result we need in the proof of Theorem III is the following upper bound on $\lambda_2(M_t,3)$, valid for $t = 2, 3$:

Lemma VI.4 For $t \in \{2, 3\}$,

$$\lambda_2(M_t,3) = \frac{7}{10}.$$  

(69)
Proof

The operator $M_{t,3}$ has the following form

$$M_{t,3} = \frac{1}{2}(P_{12} \otimes I_3 + I_1 \otimes P_{23}) \quad (70)$$

In analogy to the definition of $P_{i,i+1}$, we define the projectors $P_i$ as follows

$$P_i := \int_{U(2)} \mu_H(dU) U_i^\otimes \otimes U_i^\otimes. \quad (71)$$

We also consider the associate superoperator

$$P_i(X) := \int_{U(2)} \mu_H(dU) U_i^\otimes X(U_i^\dagger)^\otimes. \quad (72)$$

From Schur duality we find that all operators $X$ invariant under $P_i$ can be written as a sum of permutation operators. In more detail, consider the representation $V_{\pi,i}$ of the symmetric group $S_t$ acting on $H_i^\otimes$ given by

$$V_{\pi,i}|k_1 \otimes ... \otimes |k_t\rangle = |k_{\pi^{-1}(1)} \otimes ... \otimes k_{\pi^{-1}(t)}\rangle, \quad (73)$$

for every $\pi \in S_t$ and $|k_t\rangle \in H_i$. Then it follows that any $X$ satisfying $P_i(X) = X$ is such that

$$X = \sum_\pi c_\pi V_{\pi,i}, \quad (74)$$

for complex numbers $c_\pi$. Moreover, the subspace defined by $P_i$ is spanned by the (overcomplete) basis given by the non-normalized vectors $|V_{\pi,i}\rangle := V_{\pi,i} \otimes |\Phi\rangle$.

Likewise we define

$$P_{i,i+1}(X) := \int_{U(4)} \mu_H(dU) U_{i,i+1}^\otimes X(U_{i,i+1}^\dagger)^\otimes, \quad (75)$$

and again we find that any $X$ invariant under $P_{i,i+1}$ can be written as a sum of permutation operators $V_{\pi,(i,i+1)}$, now permuting the $t$ copies of the Hilbert space $H_{i,i+1}^\otimes$.

Since $V_{\pi,(i,i+1)} = V_{\pi,i} \otimes V_{\pi,i+1}$ it follows that $P_{i,i+1} \subset P_i \otimes P_{i+1}$ and thus

$$P_{12} \otimes I_3 + I_1 \otimes P_{23} = (P_{12} \otimes P_3 + P_1 \otimes P_{23}) \oplus (P_{12} \otimes P_3^\perp) \oplus (P_1^\perp \otimes P_{23}) \quad (76)$$

Since the last two terms in the direct sum above have eigenvalues 0 and 1, it follows that $\lambda_2(M_{t,3}) = \lambda_2(X)/2$, with $X$ given by

$$X = P_{12} \otimes P_3 + P_1 \otimes P_{23}. \quad (77)$$

Note that the support of $X$ is contained in $P_1 \otimes P_2 \otimes P_3$.

Now, for any Hilbert space $H$, in a subspace spanned by swaps acting on $H^\otimes$ one can construct a basis of operators which is orthonormal in Hilbert-Schmidt scalar product:

$$R_k = \sum_\pi b_{k\pi} V_\pi \quad (78)$$
where the coefficients $b_{k\pi}$ do not depend on the dimension of $\mathcal{H}$, but only on $t$ (we do not write this dependence explicitly). Note that depending on the dimension of $\mathcal{H}$, some $R_k$ may vanish. Using the operators $R_k$ we can write:

$$P_{i,i+1} = \sum_k |R_{k}^{(i,i+1)}\rangle\langle R_{k}^{(i,i+1)}|, \quad P_{i} = \sum_k |R_{k}^{(i)}\rangle\langle R_{k}^{(i)}|.$$  \hfill (79)

To diagonalize the operator $X$ we need to represent $P_{1,2}$ and $P_{2,3}$ in terms of product basis $R_{k}^{(1)} \otimes R_{k}^{(2)}$ and $R_{k}^{(2)} \otimes R_{k}^{(3)}$, respectively. To this end we use (78), which in our case read

$$R_{k}^{(1,2)} = \sum_{\pi} b_{k\pi} V_{\pi,1} \otimes V_{\pi,2}$$

$$R_{k}^{(1)} = \sum_{\pi} b_{k\pi} V_{\pi,1}$$

$$R_{k}^{(2)} = \sum_{\pi} b_{k\pi} V_{\pi,2}. \quad \hfill (80)$$

A simple calculation gives

$$R_{k}^{1,2} = \sum_{s,u} r_{s,u}^{(k)} R_{s} \otimes R_{u}$$ \hfill (82)

where the coefficients $r_{s,u}^{(k)}$ form a matrix given by

$$r_{(k)} = (B^{-1})^T A^{(k)} B^{-1} \quad \hfill (83)$$

with $B$ defined as the matrix with entries $b_{kl}$ and $A^{(k)}$ the diagonal matrices

$$A_{ij}^{(k)} = \delta_{ij} b_{ki}. \quad \hfill (84)$$

In this way, from the matrix $B$ we can obtain the matrix elements of the projectors $P_{1,2}$ and $P_{2,3}$ in the product bases $R_{k}^{(1)} \otimes R_{l}^{(2)}$ and $R_{k}^{(2)} \otimes R_{l}^{(3)}$, respectively. These, in turn, allow us to obtain the matrix elements of $X$ and to compute its eigenvalues. In the following we perform the analysis for $t = 2$ and $t = 3$.

2-design. For $t = 2$, the basis is given by (suitably normalized) projectors onto the symmetric and the antisymmetric subspace. The matrix $B$ is then given by

$$B = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2} \end{bmatrix} \quad \hfill (85)$$

where the basis of swaps was ordered as $\{(), (12)\}^{10}$, while the matrices $r^{(k)}$ are given by

$$r^{(1)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad r^{(2)} = \frac{1}{2} \begin{bmatrix} \alpha & 0 & 0 & \sqrt{\alpha \beta} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\alpha \beta} \\ \sqrt{\alpha \beta} & 0 & 0 & \beta \end{bmatrix}. \quad \hfill (86)$$

\[\text{\footnote{10 Here () labels the trivial permutation and (12) the swap of systems 1 and 2.}}\]
with \( \alpha = \frac{9}{5}, \beta = \frac{1}{5} \). Diagonalizing \( X \), we obtain that it has the second largest eigenvalue \( 7/5 \), which gives \( \lambda_2(M_{3,t}) = 7/10 \) for \( t = 2 \).

3-design. Here we exploit the orthonormal basis constructed in [54]:

\[
R_+ = \frac{1}{12}(\mathbb{I} + V_{(12)} + V_{(13)} + V_{(23)} + V_{(123)} + V_{(132)}),
\]

\[
R_- = \frac{1}{12}(\mathbb{I} - V_{(12)} - V_{(13)} - V_{(23)} + V_{(123)} + V_{(132)}),
\]

\[
R_0 = \frac{1}{2}(R_+ - R_-),
\]

\[
R_1 = \frac{1}{6}(2V_{(23)} - V_{(13)} - V_{(12)}),
\]

\[
R_2 = \frac{1}{2\sqrt{3}}(V_{(12)} - V_{(13)}),
\]

\[
R_3 = \frac{i}{2\sqrt{3}}(V_{(123)} - V_{(132)}).
\]

(87)

where permutations are written by means of cycles. The related matrix \( B \) is given by

\[
B = \frac{1}{2}
\begin{bmatrix}
\frac{1}{3} & 0 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} \\
0 & \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\
0 & \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{3}} & 0 & 0 \\
0 & 0 & 0 & 0 & i & -i \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{bmatrix}
\]

(88)

where the basis of swaps is ordered as follows \( \{(), (12), (23), (13), (123), (132)\} \). Since we work with qubits, \( R_3 = 0 \) and hence we have 5 basis elements and \( X \) is a \( 125 \times 125 \) matrix. We have computed its matrix elements and then its eigenvalues, and found that again the second largest eigenvalue is equal to \( 7/5 \).

Lastly, let us discuss the case of uniform random circuits. We have:

**Proposition VI.5** 5\( n^2 \log(1/\epsilon) \)-sized uniform random circuits form an \( \epsilon \)-approximate unitary 3-design.

**Proof** Following the proof of Theorem III is suffices to show that for \( t = 3 \),

\[
\lambda_2(N_{t,n}) \leq 1 - \frac{1}{5n^2}
\]

(89)

with

\[
N_{t,n} := \frac{1}{n^2} \sum_{i<j} P_{i,j}
\]

(90)

and

\[
P_{i,j} := \int_{U(4)} \mu_H(dU) U_{i,j}^{\otimes t} \otimes \overline{U}_{i,j}^{\otimes t}.
\]

(91)

\(^{11}\) We have diagonalized the matrix \( X \) using the symbolic manipulation program Mathematica.
Defining
\[
\bar{H} := \sum_{i,j} H_{i,j} = n^2 (\mathbb{I} - N_{t,n})
\]  
(92)

with \(H_{i,j} = \mathbb{I} - P_{i,j}\), we find Eq. (89) to be equivalent to \(\Delta(\bar{H}) \geq 1/5\).

Both \(\bar{H}\) and \(H\) (given by Eq. (61)) have the same groundspace \(S_0\). Furthermore, \(\bar{H} \geq H\). Then
\[
\Delta(\bar{H}) \geq \min_{|\psi\rangle \perp S_0} \langle \psi | \bar{H} | \psi \rangle \geq \min_{|\psi\rangle \perp S_0} \langle \psi | H | \psi \rangle = \Delta(H) \geq \frac{1}{5},
\]  
(93)

where the last inequality follows from Lemmas VI.1 and VI.4.

\[\Box\]

VII. ACKNOWLEDGMENTS

We would like to thank Salman Beigi for useful correspondence. FB is supported by a “Conhecimento Novo” fellowship from the Brazilian agency Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG). MH is supported by EU grant QESSENCE and by Polish Ministry of Science and Higher Education grant NN202 231937. FB would like to thank the hospitality of the National Quantum Information Center of Gdansk where part of this work was done. FB and MH thank the Institute Mittag Leffler for their hospitality, where (another) part of this work was done.