Multilinear Formulas and Skepticism of Quantum Computing*

Scott Aaronson†

ABSTRACT

Several researchers, including Leonid Levin, Gerard 't Hooft, and Stephen Wolfram, have argued that quantum mechanics will break down before the factoring of large numbers becomes possible. If this is true, then there should be a natural set of quantum states that can account for all quantum computing experiments performed to date, but *not* for Shor's factoring algorithm. We investigate as a candidate the set of states expressible by a polynomial number of additions and tensor products. Using a recent lower bound on multilinear formula size due to Raz, we then show that states arising in quantum error-correction require $n^{\Omega(\log n)}$ additions and tensor products even to approximate, which incidentally yields the first superpolynomial gap between general and multilinear formula size of functions. More broadly, we introduce a complexity classification of pure quantum states, and prove many basic facts about this classification. Our goal is to refine vague ideas about a breakdown of quantum mechanics into specific hypotheses that might be experimentally testable in the near future.

Categories and Subject Descriptors

F.0 [**Theory of Computation**]: General

General Terms

Theory

Keywords

quantum computing, multilinear formula size, matrix rank, error-correcting codes

*Due to space limitations, most proofs have been omitted from this abstract. Proofs can be found in an earlier version at www.arxiv.org/abs/quant-ph/0311039.

†University of California, Berkeley. Email: aaronson@cs.berkeley.edu. Supported by an NSF Graduate Fellowship, by NSF ITR Grant CCR-0121555, and by the Defense Advanced Research Projects Agency (DARPA).

STOC'04, June 13–15, 2004, Chicago, Illinois, USA.

1. INTRODUCTION

> QC of the sort that factors long numbers seems firmly rooted in science fiction ... The present attitude would be analogous to, say, Maxwell selling the Daemon of his famous thought experiment as a path to cheaper electricity from heat.
> —Leonid Levin [30]

Quantum computing presents a dilemma: is it reasonable to study a type of computer that has never been built, and might never be built in one's lifetime? Some researchers strongly believe the answer is 'no.' Their objections generally fall into four categories:

(A) There is a fundamental physical reason why large quantum computers can never be built.

(B) Even if (A) fails, large quantum computers will never be built in practice.

(C) Even if (A) and (B) fail, the speedup offered by quantum computers is of limited theoretical interest.

(D) Even if (A), (B), and (C) fail, the speedup is of limited practical value.[1]

The objections can be classified along two axes:

	Theoretical	**Practical**
Physical	(A)	(B)
Algorithmic	(C)	(D)

This paper focuses on objection (A). Its goal is not to win a debate about this objection, but to lay the groundwork for a rigorous discussion, and thus hopefully lead to new science. Section 2 provides the philosophical motivation for our paper, by examining the arguments of several quantum computing skeptics, including Leonid Levin, Gerard 't Hooft, and Stephen Wolfram. It concludes that a key weakness of their arguments is their failure to answer the following question: *Exactly what property separates the quantum states we are sure we can create, from those that suffice for Shor's*

[1]Because of the 'even if' clauses, the objections seem to us logically independent, so that there are 16 possible positions regarding them (or 15 if one is against quantum computing). We ignore the possibility that no speedup exists, in other words that $\mathsf{BPP} = \mathsf{BQP}$. By 'large quantum computer' we mean any computer much faster than its best classical simulation, as a result of asymptotic complexity rather than the speed of elementary operations. Such a computer need not be universal; it might be specialized for (say) factoring.

factoring algorithm? We call such a property a *Sure/Shor separator*. Section 3 develops a complexity theory of pure quantum states, that studies possible Sure/Shor separators. In particular, it introduces *tree states*, which informally are those states $|\psi\rangle \in \mathcal{H}_2^{\otimes n}$ expressible by a polynomial-size 'tree' of addition and tensor product gates. For example, $\alpha|0\rangle^{\otimes n} + \beta|1\rangle^{\otimes n}$ and $(\alpha|0\rangle + \beta|1\rangle)^{\otimes n}$ are both tree states.

Our main results, proved in Section 5, are lower bounds on tree size for several families of quantum states. Specifically, we show in Section 5.1 that if C is a coset in $\mathbb{Z}_2^n$, then a uniform superposition over the elements of C cannot be represented by a tree of size $n^{o(\log n)}$, with high probability if C is chosen at random.[2] Indeed, with high probability such states are not even *approximated* by trees of size $n^{o(\log n)}$. These 'coset states' are exactly what arise in stabilizer codes, a type of quantum error-correcting code.

Originally, we had hoped to show a tree size lower bound for states that arise in Shor's factoring algorithm—for example, a uniform superposition over all multiples of a fixed positive integer p, written in binary. However, we were only able to show such a bound assuming a number-theoretic conjecture, which is stated in Section 5.2.

Our lower bounds use a sophisticated recent technique of Raz [35], which was introduced to show that the permanent and determinant of a matrix require superpolynomial-size multilinear formulas. Currently, Raz's technique is only able to show lower bounds of the form $n^{\Omega(\log n)}$, but we conjecture that $2^{\Omega(n)}$ lower bounds hold in all of the cases discussed above.

The full version of this paper goes on to address the following question. If the state of a quantum computer at every time step is a tree state, then can the computer be simulated classically? In other words, letting $\mathsf{TreeBQP}$ be the class of languages accepted by such a machine, does $\mathsf{TreeBQP} = \mathsf{BPP}$? A positive answer would make tree states more attractive as a Sure/Shor separator. For once we admit any states incompatible with the polynomial-time Church-Turing thesis, it seems like we might as well go all the way, and admit *all* states preparable by polynomial-size quantum circuits! Although we leave this question open, we do show that $\mathsf{TreeBQP} \subseteq \mathsf{\Sigma}_3^{\mathsf{P}} \cap \mathsf{\Pi}_3^{\mathsf{P}}$, where $\mathsf{\Sigma}_3^{\mathsf{P}} \cap \mathsf{\Pi}_3^{\mathsf{P}}$ is the third level of the polynomial hierarchy PH. By contrast, it is conjectured that $\mathsf{BQP} \not\subset \mathsf{PH}$, though admittedly not on strong evidence.

We conclude in Section 6 with some open problems.

1.1 The Experimental Situation

An earlier version of this paper advanced the thesis that *all* quantum states prepared to date are best seen as tree states. It also proposed an experiment whose goals would be to (1) prepare coset states that provably have large tree size, and (2) demonstrate by tomography that these states were indeed prepared. We argued that such an experiment would do more than test the feasibility of quantum error-correction—it would provide an important new test of quantum mechanics itself. We have not changed this opinion. However, we have since learned that *there exist condensed-matter systems that have already been experimentally studied, and whose states very likely have superpolynomial tree size.* An example is the magnetic salt $\mathrm{LiHo}_x\mathrm{Y}_{1-x}\mathrm{F}_4$ considered by Ghosh et al. [18], which, like the cluster states of Briegel and Raussendorf [9], basically consists of a lattice of spins subject to pairwise nearest-neighbor Hamiltonians. So, in evaluating an experimental claim that a system's state has superpolynomial tree size, we now believe there are three crucial issues:

(1) How much experimental control is available? It is one thing to infer a system's state from bulk properties such as magnetic susceptibility and specific heat, and quite another to *prepare* a system in that state by (say) applying a known pulse sequence.

(2) How explicitly can we write down the hypothesized state? Do we know all pairwise interaction strengths to within some accuracy? If not, can we at least specify a probability distribution from which they were drawn? Also, for proving lower bounds, knowing a system's Hamiltonian is not enough; we need to be able to *solve* to obtain an explicit formula for the amplitudes at a particular time t.

(3) Does the state contain localized subsystems that can be interpreted as qubits? Or is the state a "soup" of free-wandering fermions or bosons? If the latter, it makes no sense to talk about the state's tree size; a different complexity measure would be needed.

Let us make two further points. First, since tree size is an asymptotic notion, when we say that an n-qubit state $|\psi_n\rangle$ was "prepared," what we really mean is that (say) $|\psi_{50}\rangle$ or $|\psi_{100}\rangle$ was prepared, and that we have no reason to suppose that preparing $|\psi_{10000}\rangle$ or $|\psi_{10000000}\rangle$ would be fundamentally different. Second, for simplicity we consider only pure states, but one can imagine several ways of extending our formalism to mixed states. For example, given a mixed state ρ, we could minimize tree size over all purifications of ρ, or minimize the expected (or maximum) tree size over all decompositions $\rho = \sum_i \alpha_i |\psi_i\rangle\langle\psi_i|$.

1.2 Recent Developments

Since this paper was first written, there have been three exciting developments of purely mathematical nature. First, we managed to 'derandomize' our lower bounds to show that certain *explicit* coset states have tree size $n^{\Omega(\log n)}$. Second, we showed *exponential* lower bounds on the "manifestly orthogonal" tree size of coset states, a notion defined in Section 3. The main ideas of these two developments are given in Section 5.1.

The third development is an $n^{\Omega(\log n)}$ tree size lower bound on 2-dimensional *cluster states* as proposed by Briegel and Raussendorf [9]. These states have the form

$$\frac{1}{2^{n/2}} \sum_x \left(\prod_{i,j} (-1)^{x_{ij}x_{i(j+1)} + x_{ij}x_{(i+1)j}} \right) |x\rangle$$

where $x = (x_{ij})$ is a $\sqrt{n} \times \sqrt{n}$ array of bits and $i, j \in \{1, \ldots, \sqrt{n}\}$ are indices that wrap around. Intriguingly, the 1-dimensional analogues of cluster states (called *spin chains*) have polynomially-bounded tree size.

Cluster states have attracted a great deal of attention recently, mostly because of their application to quantum computing via 1-qubit measurements only [34]. However, Dür and Briegel [13] gave another interesting property of cluster states: they are "persistently entangled," in the sense that one can distill n-partite entanglement from them even after each qubit has interacted with a heat bath for an amount of

[2] This result has a corollary of independent complexity-theoretic interest—the first superpolynomial gap between formula size and multilinear formula size of functions $f : \{0,1\}^n \to \mathbb{R}$.

time independent of n. Persistence of entanglement turns out to be closely related to how we show tree size lower bounds using Raz's technique. In physical terms, Raz's technique involves measuring most of a state's qubits, then partitioning the unmeasured qubits into two subsystems of equal size, and arguing that with high probability those two subsystems are still almost maximally entangled. In light of this connection between large tree size and robustness to decoherence, it is not so surprising that the first states for which we obtained an $n^{\Omega(\log n)}$ tree size lower bound are the states arising in quantum error-correction.[3]

2. HOW QUANTUM MECHANICS COULD FAIL

This section discusses objection (A), that quantum computing is impossible for a fundamental physical reason. Although this objection has been raised by several physicists, including Gerard 't Hooft [23] and Stephen Wolfram [39], we will begin with the arguments of Leonid Levin [30, 31], since those are the best known to computer scientists.[4] The following is a sample of points made by Levin that we were able to understand. We should mention that Levin does not consider our sample to be an accurate summary of his views; thus, readers are encouraged to consult [30, 31] where Levin makes further points, for example about a distinction between topological and metric approximation.

First, Levin draws an analogy between quantum computing and the unit-cost arithmetic model, suggesting that if we reject the latter as extravagant, then we should also reject the former. "[Shamir] proved ... that factoring (on infeasibility of which RSA depends) can be done in polynomial number of arithmetic operations. This result uses a so-called 'unit-cost model,' which charges one unit for each arithmetic operation, however long the operands ... The closed-minded cryptographers, however, were not convinced and this result brought a dismissal of the unit-cost model, not RSA" [30]. Levin then says about quantum computing: "Another, not dissimilar, attack is raging this very moment."

Second, in a newsgroup discussion [31] involving Levin, Daniel Gottesman, and others, Gottesman began a defense of quantum error-correction as follows: "We know linearity and all other laws of quantum mechanics are at least approximately true. Let us fix, for the sake of convenience, some degree of accuracy to which this approximation is correct—say, 20 digits." Levin interjected: "To this accuracy all these amplitudes are 0." Later Levin again said: "Rounded to 10^{-4} (if not to 10^{-10^4} :-), all amplitudes in your algorithm would be 0." To us, the most natural interpretation of these remarks is that Levin wishes to subject amplitudes to *additive* rather than multiplicative error. That is, he imagines an error process that corrupts the amplitude α_x of each basis state $|x\rangle$ to $\alpha_x \pm \varepsilon$, rather than to $\alpha_x (1 \pm \varepsilon)$ as is assumed in results on quantum fault-tolerance due to Aharonov and Ben-Or [3] among others.[5] In the additive case, clearly only classical computation is possible, since an adversary could corrupt all but $O(1/\varepsilon)$ amplitudes to 0.

Third, Levin sees no reason even to hypothesize that quantum mechanics remains valid to the accuracy needed for quantum computing. "We have never seen a physical law valid to over a dozen decimals. Typically, every few new decimal places require major rethinking of most basic concepts. Are quantum amplitudes still complex numbers to such accuracies or do they become quaternions, colored graphs, or sick-humored gremlins?" [30]

Fourth, Levin rejects the idea that quantum computing research "wins either way"—either by building quantum computers, or by discovering that our current understanding of quantum mechanics is incomplete. In his words [30]: "[Consider] this scenario. With few q-bits, QC is eventually made to work. The progress stops, though, long before QC factoring starts competing with pencils. The QC people then demand some noble [sic] prize for the correction to the Quantum Mechanics. But the committee wants more specifics than simply a nonworking machine, so something like observing the state of the QC is needed. Then they find the Universe too small for observing individual states of the needed dimensions and accuracy. (Raising sufficient funds to compete with pencil factoring may justify a Nobel Prize in Economics.)"

Levin points out that, by a simple counting argument, a 'generic' state $|\psi\rangle \in \mathcal{H}_2^{\otimes n}$ is indistinguishable from the set of states $|\varphi\rangle$ such that $|\langle\psi|\varphi\rangle| \leq \varepsilon$ by quantum circuits of subexponential size. "So, what thought experiments can probe the QC to be in the state described with the accuracy needed? I would allow to use the resources of the entire Universe, but *not more*!"

A few responses to Levin's arguments can be offered immediately. First, even classically, one can flip a coin a thousand times to produce probabilities of order 2^{-1000}. Should one dismiss such probabilities as unphysical, or subject them to additive rather than multiplicative noise? At the very least, it is not obvious that amplitudes should behave differently than probabilities with respect to error—since both evolve linearly, and neither is directly observable.

Second, if Levin believes that quantum mechanics will fail, but is agnostic about what will replace it, then his argument can be turned around. How do we know that the successor to quantum mechanics will limit us to BPP, rather than letting us solve (say) PSPACE-complete problems? This is more than a logical point. Abrams and Lloyd [2] argue that a wide class of nonlinear variants of the Schrödinger equation would allow NP-complete and even $\#\mathsf{P}$-complete problems to be solved in polynomial time. And Penrose [33], who proposed a model for 'objective collapse' of the wavefunction, believes that his proposal takes us outside the Kleene hierarchy!

Third, to falsify quantum mechanics, it would suffice to show that a quantum computer evolved to *some* state far from the state that quantum mechanics predicts. Measur-

[3] The connection is not exact: Dür and Briegel [13] showed that even spin chains are persistently entangled, whereas these have polynomial tree size as mentioned previously. So it would be interesting to study *formally* the relation between tree size and persistence of entanglement.

[4] Since this paper was written, Oded Goldreich [19] has also put forward an argument against quantum computing. Compared to Levin's arguments, Goldreich's is easily understood: he believes that Shor states have exponential "non-degeneracy" and therefore take exponential time to prepare, and that there is no burden on those who hold this view to suggest a definition of non-degeneracy.

[5] In personal correspondence, Levin denied this interpretation, claiming that it makes no sense to discuss *any* equations governing a quantum computer—whether subject to additive, multiplicative, or any other kind of error.

ing the exact state is unnecessary. Nobel prizes have been awarded in the past 'merely' for falsifying a previously held theory, rather than replacing it by a new one. An example is the physics Nobel awarded to Fitch [14] and Cronin [12] in 1980 for discovering CP symmetry violation.

Perhaps the key to understanding Levin's unease about quantum computing lies in his remark that "we have never seen a physical law valid to over a dozen decimals." Here he touches on a serious epistemological question: *How far should we extrapolate from today's experiments to where quantum mechanics has never been tested?* We will try to address this question by reviewing the evidence for quantum mechanics. For our purposes it will not suffice to declare the predictions of quantum mechanics "verified to one part in a trillion," because we need to distinguish at least three different *types* of prediction: *interference*, *entanglement*, and *Schrödinger cats*. Let us consider these in turn.

(1) Interference. If the different paths that an electron could take in its orbit around a nucleus did not interfere destructively, canceling each other out, then electrons would not have quantized energy levels. So being accelerating electric charges, they would lose energy and spiral into their respective nuclei, and all matter would disintegrate. That this has not happened—together with the results of (for example) single-photon double-slit experiments—is compelling evidence for the reality of quantum interference.

(2) Entanglement. One might accept that a single particle's position is described by a wave in three-dimensional phase space, but deny that two particles are described by a wave in *six*-dimensional phase space. However, the Bell inequality experiments of Aspect et al. [7] and successors have convinced all but a few physicists that quantum entanglement exists, can be maintained over large distances, and cannot be explained by local hidden-variable theories.

(3) Schrödinger Cats. Accepting two- and three-particle entanglement is not the same as accepting that whole molecules, cats, humans, and galaxies can be in coherent superposition states. However, recently Arndt et al. [6] have performed the double-slit interference experiment using C_{60} molecules (buckyballs) instead of photons; while Friedman et al. [15] have found evidence that a superconducting current, consisting of billions of electrons, can enter a coherent superposition of flowing clockwise around a coil and flowing counterclockwise (see Leggett [29] for a survey of such experiments). Though short of cats, these experiments at least allow us to say the following: *if we could build a general-purpose quantum computer with as many components as have already been placed into coherent superposition, then on certain problems, that computer would outperform any computer in the world today.*

Having reviewed some of the evidence for quantum mechanics, we must now ask what alternatives have been proposed that might also explain the evidence. The simplest alternatives are those in which quantum states "spontaneously collapse" with some probability, as in the GRW (Ghirardi-Rimini-Weber) theory [17]. (Penrose [33] has proposed another such theory, but as mentioned earlier, his suggests that the quantum computing model is *too* restrictive.) The drawbacks of the GRW theory include violations of energy conservation, and parameters that must be fine-tuned to avoid conflicting with experiments. More relevant for us, though, is that even if the GRW theory were true, fairly large quantum computers could still be built.

A second class of alternatives includes those of 't Hooft [23] and Wolfram [39], in which something like a deterministic cellular automaton underlies quantum mechanics. On the basis of his theory, 't Hooft predicts that "[i]t will never be possible to construct a 'quantum computer' that can factor a large number faster, and within a smaller region of space, than a classical machine would do, if the latter could be built out of parts at least as large and as slow as the Planckian dimensions" [23]. Similarly, Wolfram states that "[i]ndeed within the usual formalism [of quantum mechanics] one can construct quantum computers that may be able to solve at least a few specific problems exponentially faster than ordinary Turing machines. But particularly after my discoveries ... I strongly suspect that even if this is formally the case, it will still not turn out to be a true representation of ultimate physical reality, but will instead just be found to reflect various idealizations made in the models used so far" [39, p.771].

The obvious question then is how these theories account for Bell inequality violations. We confess to being unable to understand 't Hooft's answer to this question, except that he believes that the usual notions of causality and locality might no longer apply in quantum gravity. As for Wolfram's theory, which involves "long-range threads" to account for Bell inequality violations, we argued in [1] that it fails Wolfram's own desiderata of causal and relativistic invariance.

So the challenge for quantum computing skeptics is clear. Ideally, come up with an alternative to quantum mechanics—even an idealized toy theory—that can account for all present-day experiments, yet would not allow large-scale quantum computation. Failing that, *at least say what you take quantum mechanics' domain of validity to be.* More concretely, propose a natural set S of quantum states that you believe corresponds to possible physical states of affairs.[6] The set S must contain all "Sure states" (informally, the states that have already been demonstrated in the lab), but no "Shor states" (again informally, the states that can be shown to suffice for factoring, say, 500-digit numbers). If S satisfies both of these constraints, then we call S a *Sure/Shor separator* (see Figure 1).

Of course, an alternative theory need not involve a sharp cutoff between possible and impossible states. So it is perfectly acceptable for a skeptic to define a "complexity measure" $C(|\psi\rangle)$ for quantum states, and then say something like the following: *If $|\psi_n\rangle$ is a state of n spins, and $C(|\psi_n\rangle)$ is at most, say, n^2, then I predict that $|\psi_n\rangle$ can be prepared using only "polynomial effort." Also, once prepared, $|\psi_n\rangle$ will be governed by standard quantum mechanics to extremely high precision. All states created to date have had small values of $C(|\psi_n\rangle)$. However, if $C(|\psi_n\rangle)$ grows as, say, 2^n, then I predict that $|\psi_n\rangle$ requires "exponential effort" to prepare, or else is not even approximately governed by quantum mechanics. The states that arise in Shor's factoring algorithm have exponential values of $C(|\psi_n\rangle)$. So as my Sure/Shor separator, I propose the set of all infinite families of states $\{|\psi_n\rangle\}_{n\geq 1}$, where $|\psi_n\rangle$ has n qubits, such that $C(|\psi_n\rangle) \leq p(n)$ for some polynomial p.*

To understand the importance of Sure/Shor separators, it is helpful to think through some examples. A major theme of Levin's arguments was that exponentially small ampli-

[6] A skeptic might also specify what happens if a state $|\psi\rangle \in S$ is acted on by a unitary U such that $U|\psi\rangle \notin S$, but this will not be insisted upon.

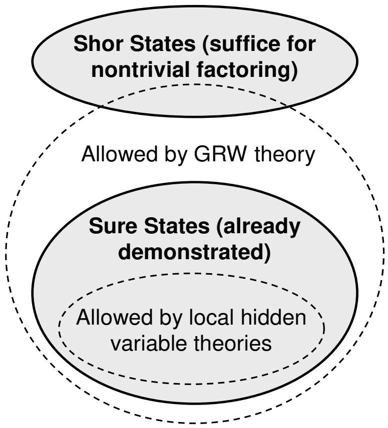

Figure 1: A Sure/Shor separator must contain all Sure states but no Shor states. That is why neither local hidden variables nor the GRW theory yields a Sure/Shor separator.

tudes are somehow unphysical. However, clearly we cannot reject *all* states with tiny amplitudes—for would anyone dispute that the state $2^{-5000}(|0\rangle + |1\rangle)^{\otimes 10000}$ is formed whenever $10,000$ photons are each polarized at $45°$? Indeed, once we accept $|\psi\rangle$ and $|\varphi\rangle$ as Sure states, we are almost *forced* to accept $|\psi\rangle \otimes |\varphi\rangle$ as well—since we can imagine, if we like, that $|\psi\rangle$ and $|\varphi\rangle$ are prepared in two separate laboratories. So considering a Shor state such as

$$|\Phi\rangle = \frac{1}{2^{n/2}} \sum_{r=0}^{2^n-1} |r\rangle \, |x^r \bmod N\rangle \, ,$$

what property of this state could quantum computing skeptics latch onto as being physically extravagant? They might complain that $|\Phi\rangle$ involves entanglement across hundreds or thousands of particles; but as mentioned earlier, there are other states with that same property, namely the "Schrödinger cats" $\left(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}\right)/\sqrt{2}$, that should be regarded as Sure states. Alternatively, the skeptics might object to the *combination* of exponentially small amplitudes with entanglement across hundreds of particles. However, simply viewing a Schrödinger cat state in the Hadamard basis produces an equal superposition over all strings of even parity, which has both properties. We seem to be on a slippery slope leading to all of quantum mechanics! Is there any defensible place to draw a line?

The dilemma above is what led us to propose *tree states* as a candidate Sure/Shor separator. The idea, which might seem more natural to logicians than to physicists, is this. Once we accept the linear combination and tensor product rules of quantum mechanics—allowing $\alpha|\psi\rangle + \beta|\varphi\rangle$ and $|\psi\rangle \otimes |\varphi\rangle$ into our set S of possible states whenever $|\psi\rangle, |\varphi\rangle \in S$—one of our few remaining hopes for keeping S a proper subset of the set of *all* states is to impose some restriction on how those two rules can be iteratively applied. In particular, we could let S be the closure of $\{|0\rangle, |1\rangle\}$ under a *polynomial number* of linear combinations and tensor products. That is, S is the set of all infinite families of states $\{|\psi_n\rangle\}_{n\geq 1}$ with $|\psi_n\rangle \in \mathcal{H}_2^{\otimes n}$, such that $|\psi_n\rangle$ can be expressed as a

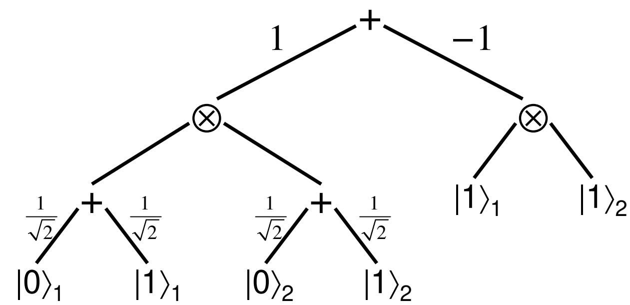

Figure 2: Expressing $(|00\rangle + |01\rangle + |10\rangle - |11\rangle)/2$ by a tree of linear combination and tensor product gates, with scalar multiplication along edges. Subscripts denote the identity of a qubit.

"tree" involving at most $p(n)$ addition, tensor product, $|0\rangle$, and $|1\rangle$ gates for some polynomial p (see Figure 2).

One can check that S so defined is rich enough to include Schrödinger cats, collections of Bell pairs, and many other examples of Sure states. Indeed, it is not obvious that there are *any* Sure states not in S; whether there are hinges on considerations such as those in Section 1.1. For the reasons discussed in that section, we would not defend the idea that "all states in Nature are tree states" as a serious physical hypothesis. Our point is simply that to debate objection (A), we need a *foil*—a way the world could be such that (i) large-scale quantum computing is impossible, but (ii) no experiment has yet detected any deviation from quantum mechanics. Several of the obvious ideas for such a foil are nonstarters. Limiting the class of quantum states to those with a certain kind of polynomial-size representation is the simplest example of a foil we could come up with. Our goal in this paper is to investigate where that idea leads.

3. CLASSIFYING QUANTUM STATES

In both quantum and classical complexity theory, the objects studied are usually sets of languages or Boolean functions. However, a generic n-qubit quantum state requires exponentially many classical bits to describe, and this suggests looking at *the complexity of quantum states themselves*. That is, which states have polynomial-size classical descriptions of various kinds? This question has been studied from several angles by Aharonov and Ta-Shma [4]; Janzing, Wocjan, and Beth [24]; Vidal [38]; and Green et al. [22]. Here we propose a unified framework for the question. For simplicity, we limit ourselves to pure states $|\psi_n\rangle \in \mathcal{H}_2^{\otimes n}$ with the fixed orthogonal basis $\{|x\rangle : x \in \{0,1\}^n\}$. Also, by 'states' we mean infinite families of states $\{|\psi_n\rangle\}_{n\geq 1}$.

Like complexity classes, pure quantum states can be organized into a hierarchy (see Figure 3). At the bottom are the classical basis states, which have the form $|x\rangle$ for some $x \in \{0,1\}^n$. We can generalize classical states in two directions: to the class $\otimes_1$ of separable states, which have the form $(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes \cdots \otimes (\alpha_n|0\rangle + \beta_n|1\rangle)$; and to the class Σ_1, which consists of all states $|\psi_n\rangle$ that are superpositions of at most $p(n)$ classical states, where p is a polynomial. At the next level, $\otimes_2$ contains the states that can be written as a tensor product of Σ_1 states, with qubits permuted arbitrar-

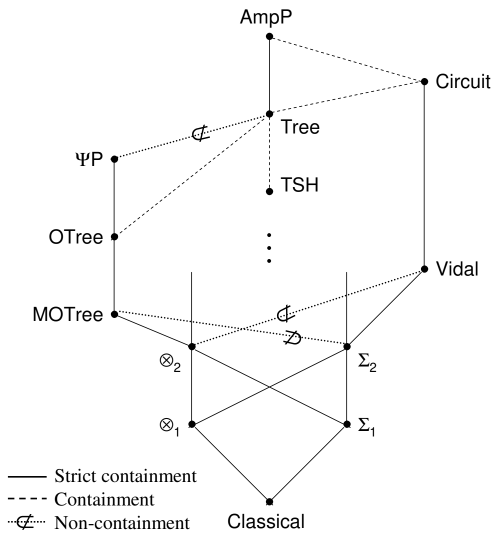

Figure 3: Relations among quantum state classes.

ily. Likewise, $\mathbf{\Sigma}_2$ contains the states that can be written as a linear combination of a polynomial number of $\otimes_1$ states. We can continue indefinitely to $\mathbf{\Sigma}_3$, $\otimes_3$, etc. Containing the whole 'tensor-sum hierarchy' $\cup_k \mathbf{\Sigma}_k = \cup_k \otimes_k$ is the class Tree, of all states expressible by a polynomial-size tree of additions and tensor products nested arbitrarily. (Formally, Tree consists of all states $|\psi_n\rangle$ such that $\mathrm{TS}(|\psi_n\rangle) \leq p(n)$ for some polynomial p, where the *tree size* $\mathrm{TS}(|\psi_n\rangle)$ will be defined shortly.) Four other classes deserve mention:

Circuit, a circuit analog of Tree, contains the states $|\psi_n\rangle = \sum_x \alpha_x |x\rangle$ such that for all n, there exists a multilinear algebraic circuit of size $p(n)$ over the complex numbers that outputs α_x given x as input, for some polynomial p.

AmpP contains the states $|\psi_n\rangle = \sum_x \alpha_x |x\rangle$ such that for all n, b, there exists a classical circuit of size $p(n+b)$ that outputs α_x to b bits of precision given x as input, for some polynomial p.

Vidal contains the states that are 'polynomially entangled' in the sense of Vidal [38]. Given a partition of $\{1,\ldots,n\}$ into A and B, let $\chi_A(|\psi_n\rangle)$ be the minimum k for which $|\psi_n\rangle$ can be written as $\sum_{i=1}^k \alpha_i |\varphi_i^A\rangle \otimes |\varphi_i^B\rangle$, where $|\varphi_i^A\rangle$ and $|\varphi_i^B\rangle$ are states of qubits in A and B respectively. ($\chi_A(|\psi_n\rangle)$ is known as the *Schmidt rank*.) Let $\chi(|\psi_n\rangle) = \max_A \chi_A(|\psi_n\rangle)$. Then $|\psi_n\rangle \in$ Vidal if and only if $\chi(|\psi_n\rangle) \leq p(n)$ for some polynomial p.

ΨP contains the states $|\psi_n\rangle$ such that for all n and $\varepsilon > 0$, there exists a quantum circuit of size $p(n + \log(1/\varepsilon))$ that maps the all-0 state to a state some part of which has trace distance at most $1-\varepsilon$ from $|\psi_n\rangle$, for some polynomial p. Because of the Solovay-Kitaev Theorem [25, 32], ΨP is invariant under the choice of universal gate set.

We now formalize the notion of *tree size* of a quantum state, which will be used throughout this paper.

Definition 1. *A quantum state tree over $\mathcal{H}_2^{\otimes n}$ is a rooted tree where each leaf vertex is labeled with $\alpha|0\rangle + \beta|1\rangle$ for some $\alpha, \beta \in \mathbb{C}$, and each non-leaf vertex (called a gate) is labeled with either $+$ or $\otimes$. Each vertex v is also labeled with a set $S(v) \subseteq \{1,\ldots,n\}$, such that*

(i) If v is a leaf then $|S(v)| = 1$,

(ii) If v is the root then $S(v) = \{1,\ldots,n\}$,

(iii) If v is a $+$ gate and w is a child of v, then $S(w) = S(v)$,

(iv) If v is a $\otimes$ gate and $w_1,\ldots,w_k$ are the children of v, then $S(w_1),\ldots,S(w_k)$ are pairwise disjoint and form a partition of $S(v)$.

Finally, if v is a $+$ gate, then the outgoing edges of v are labeled with complex numbers. For each v, the subtree rooted at v represents a quantum state of the qubits in $S(v)$ in the obvious way. We require this state to be normalized for each v.[7]

We say a tree is *orthogonal* if it satisfies the further condition that if v is a $+$ gate, then any two children w_1, w_2 of v represent $|\psi_1\rangle, |\psi_2\rangle$ with $\langle\psi_1|\psi_2\rangle = 0$. If the condition $\langle\psi_1|\psi_2\rangle = 0$ can be replaced by the stronger condition that for all basis states $|x\rangle$, either $\langle\psi_1|x\rangle = 0$ or $\langle\psi_2|x\rangle = 0$, then we say the tree is *manifestly orthogonal*.

For reasons of convenience, we define the *size* $|T|$ of a tree T to be the number of leaf vertices. Then given a state $|\psi\rangle \in \mathcal{H}_2^{\otimes n}$, the *tree size* $\mathrm{TS}(|\psi\rangle)$ is the minimum size of a tree that represents $|\psi\rangle$. The *orthogonal tree size* $\mathrm{OTS}(|\psi\rangle)$ and *manifestly orthogonal tree size* $\mathrm{MOTS}(|\psi\rangle)$ are defined similarly. Then OTree is the class of $|\psi_n\rangle$ such that $\mathrm{OTS}(|\psi_n\rangle) \leq p(n)$ for some polynomial p, and MOTree is the class such that $\mathrm{MOTS}(|\psi_n\rangle) \leq p(n)$ for some p.

It is easy to see that

$$n \leq \mathrm{TS}(|\psi\rangle) \leq \mathrm{OTS}(|\psi\rangle) \leq \mathrm{MOTS}(|\psi\rangle) \leq n2^n$$

for every $|\psi\rangle$, and that the set of $|\psi\rangle$ such that $\mathrm{TS}(|\psi\rangle) < 2^n$ has measure 0 in $\mathcal{H}_2^{\otimes n}$. Two other important properties of TS and OTS are that they are invariant under local[8] basis changes; and that if $|\phi\rangle$ is obtained from $|\psi\rangle$ by applying a k-qubit unitary, then $\mathrm{TS}(|\phi\rangle) \leq 4^k\,\mathrm{TS}(|\psi\rangle)$ and $\mathrm{OTS}(|\phi\rangle) \leq 4^k\,\mathrm{OTS}(|\psi\rangle)$.

We can also define the *ε-approximate tree size* $\mathrm{TS}_\varepsilon(|\psi\rangle)$ to be the minimum size of a tree representing a state $|\varphi\rangle$ such that $|\langle\psi|\varphi\rangle|^2 \geq 1-\varepsilon$, and define $\mathrm{OTS}_\varepsilon(|\psi\rangle)$ and $\mathrm{MOTS}_\varepsilon(|\psi\rangle)$ similarly.

Definition 2. *An arithmetic formula (over the ring $\mathbb{C}$ and n variables) is a rooted binary tree where each leaf vertex is labeled with either a complex number or a variable in $\{x_1,\ldots,x_n\}$, and each non-leaf vertex is labeled with either $+$ or $\times$. Such a tree represents a polynomial $p(x_1,\ldots,x_n)$ in the obvious way. We call a polynomial multilinear if no variable appears raised to a higher power than 1, and an arithmetic formula multilinear if the polynomials computed by each of its subtrees are multilinear.*

The *size* $|\Phi|$ of a multilinear formula Φ is the number of leaf vertices. Given a multilinear polynomial p, the multilinear formula size $\mathrm{MFS}(p)$ is the minimum size of a multilinear formula that represents p. Then given a function $f: \{0,1\}^n \to \mathbb{C}$, we define

$$\mathrm{MFS}(f) = \min_{p\,:\,p(x)=f(x)\ \forall x\in\{0,1\}^n} \mathrm{MFS}(p).$$

[7]Requiring only the *whole* tree to represent a normalized state clearly yields no further generality.

[8]Several people told us that a reasonable complexity measure must be invariant under *all* basis changes. Alas, this would imply that all pure states have the same complexity!

(Actually p turns out to be unique.) We can also define the ε-approximate multilinear formula size of f,

$$\mathrm{MFS}_{\varepsilon}(f) = \min_{p\,:\,\|p-f\|_2^2 \leq \varepsilon} \mathrm{MFS}(p)$$

where $\|p-f\|_2^2 = \sum_{x\in\{0,1\}^n} |p(x) - f(x)|^2$. Now given a state $|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x |x\rangle$ in $\mathcal{H}_2^{\otimes n}$, let f_ψ be the function from $\{0,1\}^n$ to $\mathbb{C}$ defined by $f_\psi(x) = \alpha_x$.

THEOREM 3. *For all* $|\psi\rangle$*,*

(i) $\mathrm{MFS}(f_\psi) \leq \mathrm{TS}(|\psi\rangle)$.

(ii) $\mathrm{TS}(|\psi\rangle) = O(\mathrm{MFS}(f_\psi) + n)$.

(iii) $\mathrm{MFS}_\delta(f_\psi) \leq \mathrm{TS}_\varepsilon(|\psi\rangle)$ *where* $\delta = 2 - 2\sqrt{1-\varepsilon}$.

(iv) $\mathrm{TS}_{2\varepsilon}(|\psi\rangle) = O(\mathrm{MFS}_\varepsilon(f_\psi) + n)$.

We conclude this section with some results about the quantum state hierarchy in Figure 3: Proposition 4 shows simple inclusions and separations, while Proposition 5 shows that separations higher in the hierarchy would imply major complexity class separations (and vice versa).

PROPOSITION 4.

(i) $\mathsf{Tree} \cup \mathsf{Vidal} \subseteq \mathsf{Circuit} \subseteq \mathsf{AmpP}$.

(ii) All states in Vidal *have tree size* $n^{O(\log n)}$.

(iii) $\Sigma_2 \subseteq \mathsf{Vidal}$ *but* $\otimes_2 \not\subset \mathsf{Vidal}$.

(iv) $\otimes_2 \subsetneq \mathsf{MOTree}$.

(v) Σ_1*,* Σ_2*,* Σ_3*,* $\otimes_1$*,* $\otimes_2$*, and* $\otimes_3$ *are all distinct. Also,* $\otimes_3 \neq \Sigma_4 \cap \otimes_4$.

PROPOSITION 5.

(i) $\mathsf{BQP} = \mathsf{P}^{\#\mathsf{P}}$ *implies* $\mathsf{AmpP} \subseteq \Psi\mathsf{P}$.

(ii) $\mathsf{AmpP} \subseteq \Psi\mathsf{P}$ *implies* $\mathsf{NP} \subseteq \mathsf{BQP/poly}$.

(iii) $\mathsf{P} = \mathsf{P}^{\#\mathsf{P}}$ *implies* $\Psi\mathsf{P} \subseteq \mathsf{AmpP}$.

(iv) $\Psi\mathsf{P} \subseteq \mathsf{AmpP}$ *implies* $\mathsf{BQP} \subseteq \mathsf{P/poly}$.

4. BASIC RESULTS

Before studying the tree size of specific quantum states, we would like to know in general how tree size behaves as a complexity measure. In this section we state three rather nice properties of tree size (again, proofs are omitted from this abstract).

THEOREM 6. *For all* $\varepsilon > 0$*, there exists a tree representing* $|\psi\rangle$ *of size* $O\left(\mathrm{TS}(|\psi\rangle)^{1+\varepsilon}\right)$ *and depth* $O(\log \mathrm{TS}(|\psi\rangle))$*, and a manifestly orthogonal tree of size* $O\left(\mathrm{MOTS}(|\psi\rangle)^{1+\varepsilon}\right)$ *and depth* $O(\log \mathrm{MOTS}(|\psi\rangle))$.

THEOREM 7. *Any* $|\psi\rangle$ *can be prepared by a quantum circuit of size polynomial in* $\mathrm{OTS}(|\psi\rangle)$*. Thus* $\mathsf{OTree} \subseteq \Psi\mathsf{P}$.

THEOREM 8. *If* $|\psi\rangle \in \mathcal{H}_2^{\otimes n}$ *is chosen uniformly at random under the Haar measure, then* $\mathrm{TS}_{1/16}(|\psi\rangle) = 2^{\Omega(n)}$ *with probability* $1 - o(1)$.

A corollary of Theorem 8 is the following 'nonamplification' property: there exist states that can be approximated to within, say, 1% by trees of polynomial size, but that require exponentially large trees to approximate to within a smaller margin (say 0.01%).

COROLLARY 9. *For all* $\delta \in (0,1]$*, there exists a state* $|\psi\rangle$ *such that* $\mathrm{TS}_\delta(|\psi\rangle) = n$ *but* $\mathrm{TS}_\varepsilon(|\psi\rangle) = 2^{\Omega(n)}$ *where* $\varepsilon = \delta/32 - \delta^2/4096$.

5. LOWER BOUNDS

We want to show that certain quantum states of interest to us are not represented by trees of polynomial size. At first this seems like a hopeless task. Proving superpolynomial formula-size lower bounds for 'explicit' functions is a notorious open problem, as it would imply complexity class separations such as $\mathsf{NC}^1 \neq \mathsf{P}$.

Here, though, we are only concerned with *multilinear* formulas. Could this make it easier to prove a lower bound? The answer is not obvious, but very recently, for reasons unrelated to quantum computing, Raz [35] showed the first superpolynomial lower bounds on multilinear formula size. In particular, he showed that multilinear formulas computing the permanent or determinant of an $n \times n$ matrix over any field have size $n^{\Omega(\log n)}$.

Raz's technique is a beautiful combination of the Furst-Saxe-Sipser method of random restrictions [16], with matrix rank arguments as used in communication complexity. We now outline the method. Given a function $f : \{0,1\}^n \to \mathbb{C}$, let a *k-restriction* R (for $0 \leq k \leq n/2$) set $n - 2k$ of the variables of f to either 0 or 1, and partition the remaining $2k$ variables into two collections $y = (y_1, \ldots, y_k)$ and $z = (z_1, \ldots, z_k)$. This yields a restricted function $f_{|R}(y,z) : \{0,1\}^k \times \{0,1\}^k \to \mathbb{C}$. Then let $M_{f|R}$ be a $2^k \times 2^k$ matrix whose rows are labeled by assignments $y \in \{0,1\}^k$, and whose columns are labeled by assignments $z \in \{0,1\}^k$. The (y,z) entry of $M_{f|R}$ equals $f_{|R}(y,z)$. Let $\mathrm{rank}\left(M_{f|R}\right)$ be the rank of $M_{f|R}$ over the complex numbers. The following is a special case[9] of Raz's main theorem [35]; recall that $\mathrm{MFS}(f)$ is the minimum size of a multilinear formula for f.

THEOREM 10 (RAZ). *Let* $\mathcal{D}_k$ *be the uniform distribution over k-restrictions of* f*, meaning that* $y_1, \ldots, y_k$ *and* $z_1, \ldots, z_k$ *are chosen uniformly at random, and each of the remaining* $n - 2k$ *variables is set to* 1 *with independent probability* 1/2 *and to* 0 *otherwise. Set* $k = n^\delta$*, and suppose that for some constants* $\delta \in (0, 1/3]$ *and* $c > 0$*,*

$$\Pr_{R \in \mathcal{D}_k}\left[\mathrm{rank}\left(M_{f|R}\right) \geq c2^k\right] = \Omega(1).$$

Then $\mathrm{MFS}(f) = n^{\Omega(\log n)}$.

A simple extension of Theorem 10 yields lower bounds on approximate tree size. Given an $N \times N$ matrix $M = (m_{ij})$, let $\mathrm{rank}_\varepsilon(M) = \min_{L\,:\,\|L-M\|_2^2 \leq \varepsilon} \mathrm{rank}(L)$ where $\|L - M\|_2^2 = \sum_{i,j=1}^N |l_{ij} - m_{ij}|^2$.

COROLLARY 11. *Letting* $\mathcal{D}_k$ *be as before, suppose that for some* κ*,*

$$\Pr_{R \in \mathcal{D}_k}\left[\mathrm{rank}_\delta\left(M_{f|R}\right) \geq c2^k\right] = \frac{1}{\kappa} + \Omega(1)$$

where $\|f\|_2^2 = 1$ *and* $\delta = \kappa\varepsilon 2^{2k}/2^n$*. Then* $\mathrm{MFS}_\varepsilon(f) = n^{\Omega(\log n)}$.

We will apply Raz's technique to obtain $n^{\Omega(\log n)}$ tree size lower bounds for two classes of quantum states: states arising in quantum error-correction in Section 5.1, and (assuming a number-theoretic conjecture) states arising in Shor's factoring algorithm in Section 5.2.

[9] Raz uses a distribution over restrictions that is more tailored to the permanent and determinant functions, but examining his proof, it is easy to see that our distribution works equally well.

5.1 Coset States

Let the elements of $\mathbb{Z}_2^n$ be labeled by n-bit strings. Given a coset C in $\mathbb{Z}_2^n$, we define the *coset state* $|C\rangle$ as follows:

$$|C\rangle = \frac{1}{\sqrt{|C|}} \sum_{x \in C} |x\rangle .$$

Coset states arise as codewords in the class of quantum error-correcting codes known as stabilizer codes [11, 20, 37]. Our interest in these states, however, arises from their large tree size rather than their error-correcting properties.

For an integer $k \geq 0$, let $\mathcal{E}_{k,n}$ be the following distribution over cosets C. Choose a $k \times n$ matrix A and $k \times 1$ vector v by setting each entry to 0 or 1 uniformly and independently. Then let $C = \{x \mid Ax \equiv v\}$ (here all congruences are mod 2). By Theorem 3, it suffices to consider the multilinear formula size of the function $f_C(x)$, which is 1 if $x \in C$ and 0 otherwise. Throughout this subsection we set $k = n^{1/3}$.

Theorem 12. *If C is drawn from $\mathcal{E}_{k,n}$, then* $\operatorname{MFS}(f_C) = n^{\Omega(\log n)}$ *(and hence* $\operatorname{TS}(|C\rangle) = n^{\Omega(\log n)}$*), with probability* $\Omega(1)$ *over* C.

Proof. Let R be a random k-restriction of f_C: that is, it renames $2k$ randomly chosen inputs $y_1, \ldots, y_k, z_1, \ldots, z_k$, and restricts the remaining $n - 2k$ inputs to 0 or 1 each with independent probability $1/2$. Let $M_{C|R}$ be the $2^k \times 2^k$ matrix whose (y,z) entry is $f_{C|R}(y,z)$; then we need to show that $\operatorname{rank}(M_{C|R})$ is large with high probability. Let A_y be the $k \times k$ submatrix of the $k \times n$ matrix A consisting of all rows that correspond to y_i for some $i \in \{1, \ldots, k\}$. Similarly, let A_z be the $k \times k$ submatrix consisting of all rows that correspond to z_i for some $i \in \{1, \ldots, k\}$. Then it is easy to see that, so long as A_y and A_z are both invertible, for all 2^k settings of y there exists a *unique* setting of z for which $f_{C|R}(y,z) = 1$. This then implies that $M_{C|R}$ is a permutation of the identity matrix, and hence that $\operatorname{rank}(M_{C|R}) = 2^k$. Now, the probability that a random $k \times k$ matrix over $\mathbb{Z}_2$ is invertible is

$$\frac{1}{2} \cdot \frac{3}{4} \cdot \cdots \cdot \frac{2^k - 1}{2^k} > 0.288.$$

So the probability that A_y and A_z are both invertible is at least 0.288^2. By Markov's inequality, it follows that for at least an 0.04 fraction of C's, $\operatorname{rank}(M_{C|R}) = 2^k$ for at least an 0.04 fraction of R's. Theorem 10 then yields the desired result. □

Since coset states are easily prepared by polynomial-size quantum circuits, a corollary of Theorem 12 is that $\mathsf{\Psi P} \not\subset \mathsf{Tree}$. Since f_C clearly has a (non-multilinear) arithmetic formula of size $O(nk)$, a second corollary is the following.

Corollary 13. *There exists a family of functions* $g_n : \{0,1\}^n \to \mathbb{R}$ *that has polynomial-size arithmetic formulas, but no polynomial-size multilinear formulas.*

The reason Corollary 13 does not follow from Raz's results is that polynomial-size formulas for the permanent and determinant are not known; the smallest known formulas for the determinant have size $n^{O(\log n)}$ (see [10]).

We have shown that not all coset states are tree states, but it is still conceivable that all coset states are extremely well *approximated* by tree states. Let us now rule out the latter possibility. We first need a lemma about matrix rank, which follows from the Hoffman-Wielandt inequality.

Lemma 14. *Let M be an $N \times N$ complex matrix, and let I_N be the $N \times N$ identity matrix. Then* $\|M - I_N\|_2^2 \geq N - \operatorname{rank}(M)$.

Let $\widehat{f}_C(x)$ be $f_C(x)$ normalized to have $\left\| \widehat{f}_C \right\|_2^2 = 1$.

Theorem 15. *For $\varepsilon < 0.02$, if C is drawn from $\mathcal{E}_{k,n}$, then* $\operatorname{MFS}_\varepsilon\left(\widehat{f}_C\right) = n^{\Omega(\log n)}$ *with probability* $\Omega(1)$ *over* C.

Proof. As in Theorem 12, we look at the matrix $M_{C|R}$ induced by a random k-restriction R of $\widehat{f}_C$. We have already seen that for at least an 0.04 fraction of C's, $M_{C|R}$ is a permutation of $I_{2^k}/\sqrt{|C|}$ for at least an 0.04 fraction of R's, where I_{2^k} is the identity. In this case $\operatorname{rank}_\delta(M_{C|R}) \geq 2^k - \delta|C|$ by Lemma 14. Furthermore, since for these C's there exists an R that makes the matrices A_y and A_z from Theorem 12 invertible, it follows that the k equations that define C are linearly independent and solvable. Therefore $|C| = 2^{n-k}$. So taking $\delta = \kappa\varepsilon 2^{2k}/2^n$ with $\kappa = 1/(2\varepsilon)$, we have

$$\Pr_{R \in \mathcal{D}_k}\left[\operatorname{rank}_\delta(M_{C|R}) \geq 2^{k-1}\right] \geq 0.04 > 2\varepsilon = \frac{1}{\kappa},$$

and Corollary 11 yields the desired result. □

A corollary of Theorem 15 and of Theorem 3, part (iii), is that $\operatorname{TS}_\varepsilon(|C\rangle) = n^{\Omega(\log n)}$ with probability $\Omega(1)$ over C, for $\varepsilon < 0.0199$.

Let us say a little about how to derandomize the lower bound for coset states. In the proof of Theorem 12, all we used about the matrix A was that a random $k \times k$ submatrix has full rank with $\Omega(1)$ probability. If we switch from the field $\mathbb{F}_2$ to $\mathbb{F}_{2^d}$ for some $d \geq \log_2 n$, then it is easy to construct explicit $k \times n$ matrices with this same property. For example, let

$$V = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1^2 & 2^2 & \cdots & n^2 \\ \vdots & \vdots & & \vdots \\ 1^k & 2^k & \cdots & n^k \end{pmatrix}$$

be the transpose of the Vandermonde matrix, where $1, \ldots, n$ are labels of elements in $\mathbb{F}_{2^d}$. *Any* $k \times k$ submatrix of V has full rank, because the Reed-Solomon (RS) code that V represents is a perfect erasure code. Hence, there exists an explicit state of n "qupits" with $p = 2^d$ that has tree size $n^{\Omega(\log n)}$—namely the uniform superposition over all elements of the set $\{x \mid Vx = 0\}$.

To replace qupits by qubits, we can concatenate the RS and Hadamard codes to obtain a *binary* linear erasure code with parameters almost as good as those of the original RS code. Such a code yields an explicit $k \times n$ binary matrix V', with $k \geq n^\delta$ for some constant $\delta > 0$, such that a random $k \times k$ submatrix has full rank with $\Omega(1)$ probability. We thank Andrej Bogdanov for this observation.

An earlier version of this paper used Raz's techniques to show a separation between tree size and manifestly orthogonal tree size. Recently, using *ad hoc* techniques, we managed the following *tight* characterization of $\operatorname{MOTS}(|C\rangle)$:

Theorem 16. *For all cosets $C = \{x \mid Ax \equiv b\}$ in $\mathbb{Z}_2^n$, we have* $\operatorname{MOTS}(|C\rangle) = M(A)$ *where* $M(A)$ *equals*

$$\min\left(2^{\operatorname{rank}(A_I) + \operatorname{rank}(A_J) - \operatorname{rank}(A)}\left(M(A_I) + M(A_J)\right)\right).$$

Here the minimum is over all partitions (A_I, A_J) of the columns of A such that A_I and A_J are both nonempty. (If $n = 1$ then $M(A) = 1$.)

Theorem 16 has the following corollaries. First, if C is drawn from $\mathcal{E}_{k,n}$, then $\mathrm{MOTS}(|C\rangle) = (n/k)^{\Omega(k)}$ with probability $\Omega(1)$. We thus obtain exponential lower bounds on manifestly orthogonal tree size (this also works if C is one of the explicit cosets discussed above). Second, setting $k = \log n$, there exist orthogonal tree states $|C\rangle$ with $\mathrm{MOTS}(|C\rangle) = n^{\Omega(\log n)}$. Thus $\mathsf{OTree} \neq \mathsf{MOTree}$. Third, there exists an $O(3^n \,\mathrm{poly}(n))$-time algorithm that computes $\mathrm{MOTS}(|C\rangle)$ given C as input. (We do not know whether computing $\mathrm{MOTS}(|C\rangle)$ is NP-complete but suspect it is.)

5.2 Shor States

Since the motivation for our theory was to study possible Sure/Shor separators, an obvious question is, *do states arising in Shor's algorithm have superpolynoial tree size?* Unfortunately, we are only able to answer this question assuming a number-theoretic conjecture. To formalize the question, let p be a prime and a an integer with $0 \leq a < p < 2^n$. Then letting $w = \lfloor (2^n - a - 1)/p \rfloor$, define the 'Shor state' $|a + p\mathbb{Z}\rangle = w^{-1/2} \sum_{i=0}^{w} |a + pi\rangle$, where each integer is written as an n-bit string. This is a possible state of the first register in Shor's factoring algorithm, after the second register is measured but before the Fourier transform is applied.[10] So a lower bound on $\mathrm{TS}(|a + p\mathbb{Z}\rangle)$ would imply an equivalent lower bound for the joint state of the two registers. It is not hard to see that we can set $a = 0$ without loss of generality and consider $|p\mathbb{Z}\rangle$, and that $\mathrm{TS}(|p\mathbb{Z}\rangle) = O(\min\{np, n2^n/p\})$. We now state our number-theoretic conjecture.

CONJECTURE 17. *Let the set A consist of $5 + \log_2 \left(n^{1/3}\right)$ elements of $\left\{2^0, \ldots, 2^{n-1}\right\}$ chosen uniformly at random. For all $32n^{1/3}$ subsets $B \subseteq A$, let S contain the sum of the elements of B, and let $S \,(\mathrm{mod}\, p) = \{x \bmod p : x \in S\}$. If p is a prime chosen uniformly at random from $\left[n^{1/3}, 1.1n^{1/3}\right]$, then $\Pr_p \left[|S \,(\mathrm{mod}\, p)| \geq 3n^{1/3}/4\right] \geq 3/4$.*

PROPOSITION 18. *Conjecture 17 implies that, if we choose a prime p uniformly from the range $\left[2^{n^{1/c}}, 1.1 \cdot 2^{n^{1/c}}\right]$, then with $\Omega(1)$ probability, $\mathrm{TS}(|p\mathbb{Z}\rangle) = n^{\Omega(\log n)}$ and $\mathrm{TS}_\varepsilon(|p\mathbb{Z}\rangle) = n^{\Omega(\log n)}$ for some fixed $\varepsilon > 0$.*

In an earlier version of this paper, Conjecture 17 was stated without any restriction on how the set S is formed. The resulting conjecture was far more general than we needed, and indeed was falsified by Carl Pomerance (personal communication). On the other hand, Don Coppersmith (personal communication) has made partial progress toward proving our revised conjecture.

6. CONCLUSION AND OPEN PROBLEMS

A crucial step in quantum computing was to separate the question of whether quantum computers can be built from the question of what one could do with them. This separation allowed computer scientists to make great advances on the latter question, despite knowing nothing about the former. We have argued, however, that the tools of computational complexity theory are relevant to both questions. The claim that large-scale quantum computing is possible in principle is really a claim that certain *states* can exist—that quantum mechanics will not break down if we try to prepare those states. Furthermore, what distinguishes these states from states we have seen must be more than precision in amplitudes, or the number of qubits maintained coherently. The distinguishing property must instead be some sort of *complexity*. That is, Sure states must have succinct representations of a type that Shor states do not.

We have tried to show that, by adopting this viewpoint, we make the debate about whether quantum computing is possible less ideological and more scientific. By studying particular examples of Sure/Shor separators, quantum computing skeptics would strengthen their case—for they would then have a plausible research program aimed at identifying what, exactly, the barriers to quantum computation are. We hope, however, that the 'complexity theory of quantum states' initiated in this paper will be taken up by quantum computing proponents as well. This theory offers a new perspective on the transition from classical to quantum computing, and a new connection between quantum computing and the powerful circuit lower bound techniques of classical complexity theory.

We end with some open problems.

(1) Can Raz's technique be improved to show exponential tree size lower bounds?

(2) Can we prove Conjecture 17, implying an $n^{\Omega(\log n)}$ tree size lower bound for Shor states?

(3) Let $|\varphi\rangle$ be a uniform superposition over all n-bit strings of Hamming weight $n/2$. It is easy to show by divide-and-conquer that $\mathrm{TS}(|\varphi\rangle) = n^{O(\log n)}$. Is this upper bound tight? More generally, can we show a superpolynomial tree size lower bound for any state with permutation symmetry?

(4) Is $\mathsf{Tree} = \mathsf{OTree}$? That is, are there tree states that are not orthogonal tree states?

(5) Is the tensor-sum hierarchy of Section 3 infinite? That is, do we have $\Sigma_k \neq \Sigma_{k+1}$ for all k?

(6) Is $\mathsf{TreeBQP} = \mathsf{BPP}$? That is, can a quantum computer that is always in a tree state be simulated classically? The key question seems to be whether the concept class of multilinear formulas is efficiently learnable.

(7) Is there a practical method to compute the tree size of, say, 10-qubit states? Such a method would have great value in interpreting experimental results.

Acknowledgments

I thank Ran Raz for fruitful correspondence and for sharing an early version of his paper; Leonid Levin for clarifying his views; the anonymous reviewers for their comments; and Andrej Bogdanov, Don Coppersmith, Viatcheslav Dobrovitski, Ray Laflamme, Anthony Leggett, Mike Mosca, Ashwin Nayak, Carl Pomerance, John Preskill, Alexander Razborov, Barbara Terhal, Umesh Vazirani, Guifre Vidal, and Avi Wigderson for helpful discussions.

7. REFERENCES

[1] S. Aaronson. Book review on A New Kind of Science, *Quantum Information and Computation (QIC)* 2(5), 2002. quant-ph/0206089.

[10] In general there is no reason for p to be prime, but this seems like a convenient assumption.

[2] D. S. Abrams and S. Lloyd. Nonlinear quantum mechanics implies polynomial-time solution for NP-complete and #P problems, *Phys. Rev. Lett.* 81:3992–3995, 1998. quant-ph/9801041.

[3] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error, *Proc. ACM STOC*, pp. 176–188, 1997. quant-ph/9906129.

[4] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge, *Proc. ACM STOC*, pp. 20–29, 2003. quant-ph/0301023.

[5] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson. The quantum communication complexity of sampling, *SIAM J. Comput.* 32:1570–1585, 2003.

[6] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger. Wave-particle duality of C_{60} molecules, *Nature* 401:680–682, 1999.

[7] A. Aspect, P. Grangier, and G. Roger. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: a new violation of Bell's inequalities, *Phys. Rev. Lett.* 49:91–94, 1982.

[8] S. L. Braunstein, C. M. Caves, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing, *Phys. Rev. Lett.* 83:1054–1057, 1999. quant-ph/9811018.

[9] H. J. Briegel and R. Raussendorf. Persistent entanglement in arrays of interacting particles, *Phys. Rev. Lett.* 86:910–913, 2001. quant-ph/0004051.

[10] P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic Complexity Theory*, Springer-Verlag, 1997.

[11] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist, *Phys. Rev. A* 54:1098–1105, 1996.

[12] J. Cronin. CP symmetry violation—the search for its origin, Nobel Lecture, December 8, 1980.

[13] W. Dür and H. J. Briegel. Stability of macroscopic entanglement under decoherence, submitted, 2003. quant-ph/0307180.

[14] V. Fitch. The discovery of charge-conjugation parity asymmetry, Nobel Lecture, December 8, 1980.

[15] J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, and J. E. Lukens. Quantum superposition of distinct macroscopic states, *Nature* 406:43–46, 2000.

[16] M. Furst, J. B. Saxe, and M. Sipser. Parity, circuits, and the polynomial time hierarchy, *Math. Systems Theory* 17:13–27, 1984.

[17] G. C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems, *Phys. Rev. D*, 34:470, 1986.

[18] S. Ghosh, T. F. Rosenbaum, G. Aeppli, and S. N. Coppersmith. Entangled quantum state of magnetic dipoles, *Nature* 425:48–51, 2003. cond-mat/0402456.

[19] O. Goldreich. On quantum computing, 2004. www.wisdom.weizmann.ac.il/˜oded/on-qc.html

[20] D. Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound, *Phys. Rev. A.* 54:1862–1868, 1996. quant-ph/9604038.

[21] D. Gottesman. The Heisenberg representation of quantum computers, *Int. Conf. on Group Theoretic Methods in Physics*, 1998. quant-ph/9807006.

[22] F. Green, S. Homer, C. Moore, and C. Pollett. Counting, fanout, and the complexity of quantum ACC, *Quantum Information and Computation* 2(1):35–65, 2002. quant-ph/0106017.

[23] G. 't Hooft. Quantum gravity as a dissipative deterministic system, *Classical and Quantum Gravity* 16:3263–3279, 1999.

[24] D. Janzing, P. Wocjan, and T. Beth. Cooling and low energy state preparation for 3-local Hamiltonians are FQMA-complete, manuscript, 2003. quant-ph/0303186.

[25] A. Yu. Kitaev. Quantum computation: algorithms and error correction, *Russian Math. Surveys* 52(6):1191-1249, 1997.

[26] E. Knill, R. Laflamme, R. Martinez, and C. Negrevergne. Implementation of the five qubit error correction benchmark, *Phys. Rev. Lett.* 86:5811, 2001. quant-ph/0101034.

[27] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng. An algorithmic benchmark for quantum information processing, *Nature* 404:368–370, 2000. quant-ph/9908051.

[28] E. Knill, R. Laflamme, and W. Zurek. Resilient quantum computation, *Science* 279:342, 1998. quant-ph/9702058.

[29] A. J. Leggett. Testing the limits of quantum mechanics: motivation, state of play, prospects, *J. Phys. Condensed Matter* 14:R415–451, 2002.

[30] L. Levin. Polynomial time and extravagant models, in The tale of one-way functions, *Problems of Information Transmission* 39(1), 2003.

[31] L. Levin, D. Gottesman, P. Shor, et al. Discussion on sci.physics.research newsgroup, February-March 2000.

[32] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*, Cambridge Univ. Press, 2000.

[33] R. Penrose. *The Emperor's New Mind*, Oxford Univ. Press, 1989.

[34] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation on cluster states, *Phys. Rev. A* 68:022312, 2003. quant-ph/0301052.

[35] R. Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size, this Proceedings, 2004.

[36] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* 26(5):1484–1509, 1997. quant-ph/9508027.

[37] A. Steane. Multiple particle interference and quantum error correction, *Proc. Roy. Soc. Lond. A*, 452:2551–2577, 1996. quant-ph/9601029.

[38] G. Vidal. Efficient classical simulation of slightly entangled quantum computations, *Phys. Rev. Lett.* 91:147902, 2003. quant-ph/0301063.

[39] S. Wolfram. *A New Kind of Science*, Wolfram Media, 2002.