A Complete Characterization of Unitary Quantum Space

Bill Fefferman[*1] and Cedric Yen-Yu Lin[†1]

[1]Joint Center for Quantum Information and Computer Science (QuICS), University of Maryland

November 22, 2016

Abstract

Motivated by understanding the power of quantum computation with restricted number of qubits, we give two complete characterizations of unitary quantum space bounded computation. First we show that approximating an element of the inverse of a well-conditioned efficiently encoded $2^{k(n)} \times 2^{k(n)}$ matrix is complete for the class of problems solvable by quantum circuits acting on $\mathcal{O}(k(n))$ qubits with all measurements at the end of the computation. Similarly, estimating the minimum eigenvalue of an efficiently encoded Hermitian $2^{k(n)} \times 2^{k(n)}$ matrix is also complete for this class. In the logspace case, our results improve on previous results of Ta-Shma [32] by giving new space-efficient quantum algorithms that avoid intermediate measurements, as well as showing matching hardness results.

Additionally, as a consequence we show that PreciseQMA, the version of QMA with exponentially small completeness-soundess gap, is equal to PSPACE. Thus, the problem of estimating the minimum eigenvalue of a *local* Hamiltonian to inverse exponential precision is PSPACE-complete, which we show holds even in the frustration-free case. Finally, we can use this characterization to give a provable setting in which the ability to prepare the ground state of a local Hamiltonian is more powerful than the ability to prepare PEPS states.

Interestingly, by suitably changing the parameterization of either of these problems we can completely characterize the power of quantum computation with *simultaneously* bounded time and space.

[*]wjf@umd.edu

[†]cedricl@umiacs.umd.edu

1 Introduction

How powerful is quantum computation with a restricted number of qubits? In this work we will study unitary quantum space-bounded classes - those problems solvable using a given amount of (quantum and classical) space, with all quantum measurements performed at the end of the computation. We give two sets of complete problems for these classes; to the best of our knowledge, these are the first natural complete problems proposed for quantum space-bounded classes.

The first problem we consider, the $k(n)$-*Well-conditioned Matrix Inversion* problem, is a well-conditioned version of the ubiquitous matrix inversion problem. The second problem we consider, the $k(n)$-*Minimum Eigenvalue* problem, asks us to compute the minimum eigenvalue of a Hermitian matrix to high precision – in the context of quantum complexity, this is a natural generalization of the familiar local Hamiltonian problem [25]. Interestingly enough, the first (resp. second) problem is the space-bounded variant of a BQP-complete [20] (resp. QMA-complete [25]) problem; their complexities coincide in the space-bounded regime. For the sake of readability, we defer precise definitions of these problems and statements of our results until Sections 3 and 4.

We now proceed to give some justification for the importance of our results. In the following discussion, $\mathsf{BQ_USPACE}[k(n)]$ refers to the class of problems solvable with bounded error by a quantum algorithm running in $\mathcal{O}(k(n))$; the subscript **U** indicates that the algorithm is unitary, i.e. employs no intermediate measurements.

1.1 Background and Motivation

The Matrix Inversion problem is of central importance in computational complexity theory. Matrix inversion is known to be complete for DET, the class of functions as hard as computing the determinant of an integer matrix, which can be solved in classical $\mathcal{O}(\log^2(n))$ space [6, 13]. It is a major open problem to determine if Matrix Inversion can be solved in classical logarithmic space, which would imply $\mathsf{L} = \mathsf{NL} = \mathsf{DET}$.

Recently, Ta-Shma [32], building on work of Harrow, Hassidim, and Lloyd [20], showed that a well-conditioned $n \times n$ matrix can be inverted (up to $1/\mathrm{poly}(n)$ error) by a quantum $O(\log n)$ space algorithm using intermediate measurements. Similarly, Ta-Shma also gives an algorithm for computing eigenvalues of a Hermitian matrix with similar space. These algorithms achieve a quadratic advantage in space over the best known classical algorithms, which require $\Omega(\log^2 n)$ space. This is the maximum quantum advantage possible, since Watrous has shown $\mathsf{BQSPACE}[k(n)] \subseteq \mathsf{SPACE}[\mathcal{O}(k(n)^2)]$ [37, 38] even for quantum algorithms with intermediate measurements.

Our completeness result for matrix inversion, along with observing our algorithm for matrix inversion (Theorem 15) actually gives a high-precision approximation, gives the following corollary in the logspace case (see Remark 11).

Corollary 1. *The problem of approximating, to constant precision, an entry of the inverse of an $n \times n$ positive semidefinite matrix with condition number at most $\mathrm{poly}(n)$ is $\mathsf{BQ_UL}$-complete under L-reductions, where $\mathsf{BQ_UL}$ is the set of problems solvable in unitary quantum logspace. This problem remains in $\mathsf{BQ_UL}$ even if $1/\mathrm{poly}(n)$ precision is required.*

Similarly, restricting Thereom 4 to the logspace case gives the following corollary.

Corollary 2. *The problem of approximating, to $1/\mathrm{poly}(n)$ precision, the minimum eigenvalue of an $n \times n$ positive semidefinite matrix is $\mathsf{BQ_UL}$-complete under L-reductions.*

These corollaries improve upon Ta-Shma's results [32] in two ways. First, our algorithms solve these problems without needing intermediate measurements. Unlike in time complexity, where the

"Principle of safe storage" gives a time-efficient procedure to defer intermediate measurements, these methods may incur an exponential blow-up in space.

One might wonder why we care so much about avoiding intermediate measurements. The main reason is that removing intermediate measurements from the computation allows us to give matching hardness results, showing the optimality of our algorithms. This is the second way our results improve on those of Ta-Shma. In particular, our proofs crucially use space-efficient methods for the amplification of unitary quantum computations, which are not known in the non-unitary model. This is because the techniques require applying the inverse of the circuit, which of course is impossible if the circuit contains intermediate measurements. We will also rely on ideas from Kitaev's clock construction, which constructs a local Hamiltonian from a unitary quantum circuit.

Specifically, we will show that the problems of inverting well-conditioned matrices and computing minimum eigenvalues of Hermitian matrices are hard for unitary quantum logspace under L-reductions. In the case of our algorithm for Matrix Inversion, this means that the upper bound on the condition number bound is unlikely to be improved upon. Likewise, this gives some of the strongest evidence that even well-conditioned matrices cannot be inverted in deterministic logspace, since otherwise our results would immediately imply $L = BQ_UL$, which seems quite unlikely.

Interestingly, although our algorithms for both problems use different techniques from those of Ta-Shma, our algorithm for computing the minimum eigenvalue is completely different. In particular, our algorithm crucially relies on new methods for space efficient QMA amplification, together with some of the most powerful recent results in Hamiltonian simulation [7, 9].

Concurrently with our work, Doron, Sarid, and Ta-Shma have shown that analogous problems for stochastic matrices (e.g. computing the eigenvalue gap) are complete for classical randomized logspace, or BPL [16, 15]. In addition, Le Gall has shown that analogous problems for Laplacian matrices can be solved in BPL [18]. Since it is straightforward to see that Well-conditioned Matrix Inversion reduces to Integer Matrix Inversion, we obtain a direct proof that $BQ_UL \subseteq DET$, which was previously known indirectly via the containments $BQ_UL \subseteq PL \subseteq DET$ [38, 10].

Therefore the power of classical and quantum space-bounded classes are characterized by the ability to approximate solutions of different problems in DET (stochastic matrices for the former, and Hermitian matrices for the latter). This could shed light on the differences between deterministic, randomized, and quantum space complexity. An open question is to find a class of interesting matrices whose inverse (or eigenvalues) can be computed in *deterministic* logspace.

Interestingly, if we change the scaling of the parameters in our Matrix Inversion and Minimum Eigenvalue problems suitably, then we obtain problems that are known to be complete for BQP [20] and QMA [25, 2]. Thus by appropriately bounding the dimension of the matrix and either the condition number or the promise gap, we can give problems complete for quantum time or quantum space. In fact we can strengthen these results to settings with a simultaneously bounded amount of space and time; see Section 5.

1.2 Relationship with Matchgates

Matchgates are a subclass of quantum gates introduced by Valiant [34], who also showed that nearest neighbor matchgate circuits (which we will just call matchgate computations) are classically simulable. Matchgate computations were further shown to be equivalent to a one-dimensional model of noninteracting fermions by Terhal and DiVincenzo [33]; and equivalent to unitary quantum logspace by Jozsa, Kraus, Miyake, and Watrous [22]. Our complete problems therefore elucidate the computational power of noninteracting fermions.

We know that sampling from output distributions of matchgate computations gives us the power of BPL; but what is the computational power of computing exactly the output probabilities of

matchgate computations? We conjecture that this computational power corresponds to DET, since amplitudes of noninteracting fermion circuits are related to determinants (and see also the discussion in the previous subsection). It is known that output probabilities of matchgate computations can be exactly calculated by an efficient classical algorithm [23], which is consistent with our conjecture because DET $\in$ P.

1.3 Quantum Merlin-Arthur with Small Gap

A consequence of our proof of completeness for the $k(n)$-*Minimum Eigenvalue* problem is an equivalence between space-bounded quantum computations and quantum Merlin-Arthur proof systems. Here we give this equivalence for the polynomial space case: let PreciseQMA be the variant of QMA with exponentially small completeness-soundness gap. Then we show the following:

Corollary 3. PreciseQMA $=$ BQ$_U$PSPACE $=$ PSPACE.

The second equality is due to Watrous [37, 38]. We give similar equivalences for space-bounded quantum computations with and without a witness for other space bounds as well (Theorem 19).

We note that PreciseQMA is likely far more powerful than its classical counterpart. The analogous classical complexity class is contained in NP^{PP}: given a classical witness, the verifier runs a classical computation that in the YES case accepts with probability at least c, or in the NO case accepts with probability at most s, where $c > s$. Note that in the classical case $c - s > \exp(-\text{poly})$ is automatically satisfied. Since NP^{PP} is in the counting hierarchy, the entirety of which is contained in PSPACE (see e.g., [3]), we see that the quantum proof protocol is strictly stronger than the classical one, unless the counting hierarchy collapses to the second level.

We also show that the *local* Hamiltonian problem is PSPACE-complete when the promise gap is exponentially small (for details see Appendix G). This is in contrast to the usual case when the gap is polynomially small, where the problem is QMA-complete. Perhaps more surprisingly, PreciseQMA $=$ PSPACE is more powerful than PostBQP $=$ PP, the class of problems solvable with postselected quantum computation [1].

Another consequence concerns Projected Entangled Pair States, or PEPS, a natural extension of matrix product states to two and higher spatial dimensions, which can be described as the ground state of certain frustration-free local Hamiltonians [36]. A characterization of the computational power of PEPS was given in [30], and can be summarized as follows: let O_{PEPS} be a quantum oracle that, given the description of a PEPS, outputs the PEPS (so the output is quantum). Then BQP$_{\|,\text{classical}}^{O_{PEPS}} = $ PostBQP $=$ PP, where (following Aaronson [1]) the subscript denotes that only classical nonadaptive queries to the oracle are allowed. Moreover, let PQP stand for the set of problems solvable by a quantum computer with *unbounded error*; then it can be straightforwardly shown that PQP$_{\|,\text{classical}}^{O_{PEPS}} = $ PP as well (see Appendix I).

On the other hand, suppose we have an oracle O_{LH} that given the description of a local Hamiltonian, outputs a ground state of the Hamiltonian. Then our results show that PreciseQMA $=$ PSPACE $\subseteq$ PQP$_{\|,\text{classical}}^{O_{LH}}$. This shows that in the setting of unbounded-error quantum computation, PEPS do not capture the full computational complexity of general local Hamiltonian ground states unless PP $=$ PSPACE. We leave open the problem of determining the complexity of BQP$_{\|,\text{classical}}^{O_{LH}}$.

Lastly, we are able to strengthen our characterization to show that PreciseQMA contains PSPACE (see Appendix F), even when restricted to having perfect completeness. This allows us to prove that testing if a local Hamiltonian is frustration-free is a PSPACE-complete problem (Appendix G). We note that if the local Hamiltonian is promised to have a ground state energy of at least $1/\text{poly}$ if it is frustrated, then this is the Quantum Satisfiability problem defined by Bravyi, which

3

is known to be QMA_1 complete [11, 19]. Our results show that if the promise gap is removed then we instead get PSPACE-completeness.

2 Preliminaries

2.1 Quantum circuits

We will assume a working knowledge of quantum information. For an introduction, see [28].

A *quantum circuit* consists of a series of quantum gates each taken from some universal gateset, such as the gateset consisting of Hadamard and Toffoli gates [31]. For functions $f, g : \mathbb{N} \to \mathbb{N}$, we say a family of quantum circuits $\{Q_x\}_{x \in \{0,1\}^*}$ is *f-time g-space uniformly generated* if there exists a deterministic classical Turing machine that on input $x \in \{0,1\}^n$ and $i > 0$ outputs the i-th gate of Q_x within time $f(n)$ and workspace $g(n)$ [28].

Our restriction to a specific gateset is without loss of generality, even for logarithmic space algorithms: there exists a deterministic algorithm that given any unitary quantum gate U and a parameter ϵ, outputs a sequence of at most $\mathrm{polylog}(1/\epsilon)$ gates from any universal quantum gateset that approximates U to precision ϵ in space $\mathcal{O}(\log(1/\epsilon))$ and time $\mathrm{polylog}(1/\epsilon)$ [35]. This improves the Solovay-Kitaev theorem, which guarantees a space bound of $\mathrm{polylog}(1/\epsilon)$; see e.g., [28].

2.2 Space-bounded computation

For our model of unitary quantum space-bounded computation, we consider a quantum system with purely classical control, because there are no intermediate quantum measurements to condition future operations on. Specifically, we use the following definition (see Appendix A for more details):

Definition 4. Let $k(n)$ be a function satisfying $\Omega(\log(n)) \leq k(n) \leq \mathrm{poly}(n)$. A promise problem $L = (L_{yes}, L_{no})$ is in $\mathsf{Q_USPACE}[k(n)](c, s)$ if there exists a $\mathrm{poly}(|x|)$-time $\mathcal{O}(k)$-space uniformly generated family of quantum circuits $\{Q_x\}_{x \in \{0,1\}^*}$, where each circuit $Q_x = U_{x,T} U_{x,T-1} \cdots U_{x,1}$ has $T = 2^{\mathcal{O}(k)}$ gates, and acts on $\mathcal{O}(k(|x|))$ qubits, such that:
If $x \in L_{yes}$:
$$\left\langle 0^k \middle| Q_x^\dagger |1\rangle\langle 1|_{out} Q_x \middle| 0^k \right\rangle \geq c. \tag{1}$$
Whereas if $x \in L_{no}$:
$$\left\langle 0^k \middle| Q_x^\dagger |1\rangle\langle 1|_{out} Q_x \middle| 0^k \right\rangle \leq s. \tag{2}$$
Here *out* denotes a single qubit we measure at the end of the computation; no intermediate measurements are allowed. Furthermore, we require c and s to be computable in classical $\mathcal{O}(k(n))$-space.

For the rest of the paper we will always assume that $\Omega(\log(n)) \leq k(n) \leq \mathrm{poly}(n)$.

The bound $T = 2^{\mathcal{O}(k)}$ on the circuit size comes from that any classical Turing machine generating Q_x using space $\mathcal{O}(k(|x|))$ has at most $2^{\mathcal{O}(k)}$ configurations. We note that $2^{\mathcal{O}(k)}$ gates suffice to approximate any gate on $\mathcal{O}(k)$ qubits to high accuracy (see e.g. [28, Chapter 4]). The $\mathrm{poly}(|x|)$ time bound on the classical control can be assumed without loss of generality; see Appendix A.

Definition 5. $\mathsf{BQ_USPACE}[k] = \mathsf{Q_USPACE}[k](2/3, 1/3)$.

Theorem 6 (Watrous [37, 38]). $\mathsf{BQ_USPACE}[\mathrm{poly}] = \mathsf{PSPACE}$.

We now define space- and time-bounded analogues of QMA:

4

Definition 7. We say a promise problem $L = (L_{yes}, L_{no})$ is in (t, k)-*bounded* $\mathsf{QMA}_m(c, s)$ if there exists a t-time and $(k+m)$-space uniformly generated family of quantum circuits $\{V_x\}_{x \in \{0,1\}^*}$, each of size at most $t(|x|)$, acting on $k(|x|) + m(|x|)$ qubits, so that:

If $x \in L_{yes}$ there exists an m-qubit state $|\psi\rangle$ such that:

$$\left(\langle\psi| \otimes \langle 0^k| \right) V_x^\dagger |1\rangle\langle 1|_{out} V_x \left(|\psi\rangle \otimes |0^k\rangle \right) \geq c. \tag{3}$$

Whereas if $x \in L_{no}$, for all m-qubit states $|\psi\rangle$ we have:

$$\left(\langle\psi| \otimes \langle 0^k| \right) V_x^\dagger |1\rangle\langle 1|_{out} V_x \left(|\psi\rangle \otimes |0^k\rangle \right) \leq s. \tag{4}$$

out denotes a single qubit measured at the end of the computation; no intermediate measurements are allowed. Here c and s are computable in classical $\mathcal{O}(t(n))$-time and $\mathcal{O}(k(n) + m(n))$-space.

Definition 8. $\mathsf{QMA} = (\mathrm{poly}, \mathrm{poly})$-*bounded* $\mathsf{QMA}_{\mathrm{poly}}(2/3, 1/3)$.

Definition 9. $\mathsf{PreciseQMA} = \bigcup_{c \in (0,1]} (\mathrm{poly}, \mathrm{poly})$-*bounded* $\mathsf{QMA}_{\mathrm{poly}}(c, c - 2^{-\mathrm{poly}})$.

2.3 Other definitions and results

We use the following definition of *efficient encodings* of matrices:

Definition 10. Let M be a $2^k \times 2^k$ matrix, and $\mathcal{A}$ be a classical algorithm (e.g. a Turing machine) specified using n bits. We say that $\mathcal{A}$ is an *efficient encoding* of M if on input $i \in \{0,1\}^k$, $\mathcal{A}$ outputs the indices and contents of the non-zero entries of the i-th row, using at most $\mathrm{poly}(n)$ time and $\mathcal{O}(k)$ workspace (not including the output size). Note that as a consequence M has at most $\mathrm{poly}(n)$ nonzero entries in each row.

We will often specify a matrix M in the input by giving an efficient encoding of M. The size of the encoding is then the input size, which we will usually indicate by n.

Remark 11. It is not difficult to see that any $n \times n$ matrix has an efficient encoding of size $\mathcal{O}(n)$, since it is straightforward to construct a classical $\mathcal{O}(\log n)$-size circuit that on input i, j outputs the (i, j)-entry of the matrix.

In our results we will implicitly assume the existence of algorithms that compute some common functions on n-bit numbers, such as $\sin, \cos, \arcsin, \arccos$ and exponentiation, to within $1/\mathrm{poly}(n)$ accuracy in classical $\mathcal{O}(\log n)$ space. Algorithms for these tasks have been designed by Reif [29].[1]

Finally, we will need new results from the Hamiltonian simulation literature:

Theorem 12 ([7, 8, 9])**.** *Given as input is the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ Hermitian matrix H. Then treated as a Hamiltonian, the time evolution $\exp(-iHt)$ can be simulated using $\mathrm{poly}(n, k, \|H\|_{max}, t, \log(1/\epsilon))$ operations and $\mathcal{O}(k + \log(t/\epsilon))$ space.*

While the space complexity was not explicitly stated in [7, 8, 9], it can be seen from the analysis (see e.g. [8]). The crucial thing to notice in Theorem 12 is the polylogarithmic scaling in the error ϵ; this implies that we can obtain polynomial precision in $\exp(-iHt)$ using only polynomially many operations. Also note that the maximum eigenvalue of H, $\|H\|$, satisfies $\|H\| \leq \mathrm{poly}(n)\|H\|_{max}$.

---

[1] Reif's algorithms take only $\mathcal{O}(\log\log n \log\log\log n)$ space, but for simplicity we only use the $\mathcal{O}(\log n)$ bound.

3 The Well-Conditioned Matrix Inversion Problem

We begin with a formal statement of the problem:

Definition 13 ($k(n)$-*Well-conditioned Matrix Inversion*)**.** Given as input is the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ positive semidefinite matrix H with a known upper bound $\kappa = 2^{\mathcal{O}(k(n))}$ on the condition number, so that $\kappa^{-1}I \preceq H \preceq I$, and $s, t \in \{0,1\}^{k(n)}$. It is promised that either $|H^{-1}(s,t)| \geq b$ or $|H^{-1}(s,t)| \leq a$ for some constants $0 \leq a < b \leq 1$; determine which is the case.

Theorem 14. *For $\Omega(\log(n)) \leq k(n) \leq \mathrm{poly}(n)$, $\mathcal{O}(k(n))$-Well-conditioned Matrix Inversion is complete for $\mathsf{BQ_USPACE}[\mathcal{O}(k(n))]$ under classical reductions using $\mathrm{poly}(n)$ time and $\mathcal{O}(k(n))$ space.*

Proof. We begin by giving a new space efficient algorithm for this matrix inversion problem:

Theorem 15. *Fix functions $k(n)$, $\kappa(n)$, and $\epsilon(n)$. Suppose we are given the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ PSD matrix H such that $\kappa^{-1}I \preceq H \preceq I$. We are also given $\mathrm{poly}(n)$-time $\mathcal{O}(k + \log(\kappa/\epsilon))$-space uniform quantum circuits U_a and U_b acting on k qubits and using at most T gates. Let $U_a|0\rangle^{\otimes k(n)} = |a\rangle$ and $U_b|0\rangle^{\otimes k(n)} = |b\rangle$. The following tasks can be performed with $\mathrm{poly}(n)$-time $\mathcal{O}(k + \log(\kappa/\epsilon))$-space uniformly generated quantum circuits with $\mathrm{poly}(T, k, \kappa, 1/\epsilon)$ gates and $\mathcal{O}(k + \log(\kappa/\epsilon))$ qubits:*
1. *With at least constant probability, output an approximation of the quantum state $H^{-1}|b\rangle / \|H^{-1}|b\rangle\|$ up to error ϵ.*
2. *Approximate $\|H^{-1}|b\rangle\|$ to precision ϵ.*
3. *Approximate $|\langle a|H^{-1}|b\rangle|$ to precision ϵ.*

These circuits do not require intermediate measurements.

In fact our algorithm is much stronger: to solve $k(n)$-*Well-conditioned Matrix Inversion* we merely need to approximate $|\langle s|H^{-1}|t\rangle|$ to constant precision, while Theorem 15 actually gives an approximation to precision $2^{-\mathcal{O}(k)}$ in $\mathcal{O}(k(n))$ unitary quantum space. Moreover our algorithm does not require s and t to be computational basis states.

We note that we can modify our definition of unitary quantum space-bounded classes to include computing functions, for instance by adding a write-only one-way output tape of qubits to the Turing machine (see the discussion in Appendix A), that are all measured at the very end of the computation. The error reduction result (Corollary 33) later in our work allows the total error to be reasonably controlled. With such a modification we can compute the whole matrix inverse in unitary quantum logspace. We will not pursue this modified model further in this work.

Proof. We first briefly summarize the algorithm of Ta-Shma [32], which is based on the linear systems solver of Harrow, Hassidim and Lloyd [20]. Ta-Shma shows that an $n \times n$ matrix with condition number at most $\mathrm{poly}(n)$ can be inverted by a quantum logspace algorithm with intermediate measurements; in our language this corresponds to solving $\mathcal{O}(\log n)$-*Well-conditioned Matrix Inversion*.

Our algorithm and Ta-Shma's share the same initial procedure. In particular it is shown:

Lemma 16 (Implicit in [20, 32])**.** *There is a poly-time $\mathcal{O}(k + \log(\kappa/\epsilon'))$-space uniform quantum unitary transformation W_H over $k + \ell = \mathcal{O}(k + \log(\kappa/\epsilon'))$ qubits and using $\mathrm{poly}(k, \kappa/\epsilon')$ gates, such that for any k-qubit input state $|b\rangle$,*

$$W_H(|0\rangle^{\otimes \ell} \otimes |b\rangle) = \alpha|0\rangle_{out} \otimes |\psi_b\rangle + \sqrt{1 - \alpha^2}|1\rangle_{out} \otimes |\psi'_b\rangle, \tag{5}$$

where $|\psi_b\rangle$ and $|\psi'_b\rangle$ are normalized states such that $\||\psi_b\rangle - |0\rangle^{\otimes \ell - 1} \otimes \frac{H^{-1}|b\rangle}{\|H^{-1}|b\rangle\|}\| \leq \epsilon'$, α is a positive number satisfying $|\alpha - \frac{\|H^{-1}|b\rangle\|}{\kappa}| \leq \epsilon'$, and "out" is a 1-qubit register.

This lemma can be obtained by combining the Hamiltonian simulation algorithms of Berry et al. (Theorem 12) with the analysis of Harrow, Hassidim and Lloyd [20]; a version without the time bound is implicit in the proof of [32, Theorem 6.3]. For completeness, we sketch the proof below.

Proof sketch. Decompose $|b\rangle$ into the eigenbasis of H: $|b\rangle = \sum_\lambda a_\lambda |v_\lambda\rangle$, where λ are eigenvalues of H and $H|v_\lambda\rangle = \lambda|v_\lambda\rangle$. The following procedure satisfies Lemma 16 (all steps are approximate):

1. Perform phase estimation on the operator $\exp(iH)$ and state $|b\rangle$ to compute the eigenvalues of H into an ancilliary register, obtaining the state $\sum_\lambda a_\lambda |v_\lambda\rangle|\lambda\rangle$.

2. Implement the unitary transformation $|\lambda\rangle|0\rangle \to |\lambda\rangle[(\kappa\lambda)^{-1}|0\rangle + (\sqrt{1-(\kappa\lambda)^{-2}}|1\rangle]$, to obtain the state $\sum_\lambda a_\lambda |v_\lambda\rangle|\lambda\rangle[(\kappa\lambda)^{-1}|0\rangle + (\sqrt{1-(\kappa\lambda)^{-2}}|1\rangle]$.

3. Uncompute the eigenvalues λ by running phase estimation in reverse, obtaining the state $\sum_\lambda a_\lambda |v_\lambda\rangle|0\rangle^{\ell-1}[(\kappa\lambda)^{-1}|0\rangle + (\sqrt{1-(\kappa\lambda)^{-2}}|1\rangle]$. Note that $\sum_\lambda a_\lambda |v_\lambda\rangle|0\rangle^{\ell-1}(\kappa\lambda)^{-1}|0\rangle = \frac{1}{\kappa}H^{-1}|b\rangle$.

An appropriate error analysis of this procedure is the technical bulk of the proof; we refer the reader to [20]. For Step 1, Ta-Shma showed how to implement $\exp(iH)$ in $\mathcal{O}(k + \log(1/\epsilon))$ space [32, Theorem 4.1] (their proof works for general matrices with efficient encodings); recent sparse Hamiltonian simulation algorithms (Theorem 12) give a time efficient way to do this. $\square$

Intuitively, Lemma 16 gives a space-efficient quantum algorithm that produces a state proportional to $H^{-1}|b\rangle$ with probability at least $1/\kappa$. Our goal is to produce amplify the probability from $1/\kappa$ to a constant, to produce a state with constant overlap to the state $|0\rangle_{out} \otimes |\psi_b\rangle$ together with an estimate for $\alpha \approx \|H^{-1}|b\rangle\|$. From here our algorithm differs from Ta-Shma's and uses a combination of amplitude amplification and phase estimation. This sidesteps both the somewhat involved analysis and intermediate measurements of Ta-Shma's algorithm.

Specifically, consider the two projectors

$$\Pi_0 = |0\rangle\langle 0|^{\otimes \ell} \otimes |b\rangle\langle b|, \quad \Pi_1 = W_H^\dagger(|0\rangle\langle 0|_{out} \otimes I)W_H. \tag{6}$$

Π_0 projects onto the initial subspace, while Π_1 projects onto the initial states that would be accepted by the final measurement. The rotation $R = -(I - 2\Pi_1)(I - 2\Pi_0)$ has eigenvalues $e^{\pm i2\sin^{-1}\alpha}$ with eigenvectors $|\psi_\pm\rangle$, such that $|0\rangle^{\otimes \ell} \otimes |b\rangle = (|\psi_+\rangle + |\psi_-\rangle)/\sqrt{2}$ is a uniform superposition of the two eigenvectors. Therefore phase estimation on the operator R and input state $|0\rangle^{\otimes \ell} \otimes |b\rangle$ suffices to give an estimate of α. Furthermore both eigenvectors have constant overlap with $W_H^\dagger|\psi_b\rangle$, so applying W_H to the residual state of phase estimation allows us to complete the first task as well.

We have addressed the first two tasks in Theorem 15. For the third task (approximating $|\langle a|H^{-1}|b\rangle|$), we can choose $\Pi_1' = W_H^\dagger(|0\rangle\langle 0|_{out} \otimes I)(I_{anc} \otimes |a\rangle\langle a|)(|0\rangle\langle 0|_{out} \otimes I)W_H$ instead, and phase estimation on $R = -(I - 2\Pi_1)(I - 2\Pi_0')$ will give an estimate for $|\langle a|H^{-1}|b\rangle|$. See Appendix B for the full proof. $\square$

We establish that $k(n)$-*Well-conditioned Matrix Inversion* is $\mathsf{BQ_U SPACE}[\mathcal{O}(k)]$-hard using a similar argument to Harrow, Hassidim, and Lloyd [20], in which given a quantum circuit acting on $k(n)$ qubits we construct a efficiently encoded well-conditioned $2^{\mathcal{O}(k)} \times 2^{\mathcal{O}(k)}$ matrix H, so that a single element of H^{-1} is proportional to the success probability of the circuit. See Appendix C. $\square$

4 The Minimum Eigenvalue Problem

Our second characterization of unitary quantum space is based on the following problem:

Definition 17 ($k(n)$-*Minimum Eigenvalue* problem)**.** Given as input is the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ PSD matrix H, such that $\|H\|_{max} = \max_{s,t} |H(s,t)|$ is at most a constant. Let λ_{min} be the minimum eigenvalue of H. It is promised that either $\lambda_{min} \leq a$ or $\lambda_{min} \geq b$, where $a(n)$ and $b(n)$ are numbers such that $b - a > 2^{-\mathcal{O}(k(n))}$. Output 1 if $\lambda_{min} \leq a$, and output 0 otherwise.

Theorem 18. *For* $\Omega(\log(n)) \leq k(n) \leq \mathrm{poly}(n)$, $\mathcal{O}(k(n))$-*Minimum Eigenvalue is complete for* $\mathsf{BQ_USPACE}[\mathcal{O}(k(n))]$ *under classical reductions using* $\mathrm{poly}(n)$ *time and* $\mathcal{O}(k(n))$ *space.*

In the process of proving this result, we will also show the following equivalence:

Theorem 19. $\mathsf{BQ_USPACE}[\mathcal{O}(k(n))]$ *is equivalent to the class of problems characterized by having quantum Merlin Arthur proof systems running in polynomial time,* $\mathcal{O}(k(n))$ *witness size and space, and* $2^{-\mathcal{O}(k(n))}$ *completeness-soundness gap. Or in other words,*

$$\mathsf{BQ_USPACE}[\mathcal{O}(k(n))] = \bigcup_{c-s \geq 2^{-\mathcal{O}(k(n))}} (\mathrm{poly}, \mathcal{O}(k(n)))\text{-}bounded\ \mathsf{QMA}_{\mathcal{O}(k(n))}(c,s)$$

Our proof will consist of three steps. Lemma 20 will show that $k(n)$-*Minimum Eigenvalue* is in the generalized $\mathsf{PreciseQMA}$ class defined in Theorem 19. Lemma 21 will show that this generalized $\mathsf{PreciseQMA}$ class is contained in $\mathsf{BQ_USPACE}[k(n)]$. Finally, Lemma 22 will show that $\mathsf{BQ_USPACE}[k(n)]$-hardness of $k(n)$-*Minimum Eigenvalue*.

Lemma 20. $k(n)$-*Minimum Eigenvalue is contained in* $(\mathrm{poly}, \mathcal{O}(k(n)))$-*bounded* $\mathsf{QMA}_{\mathcal{O}(k(n))}(c,s)$ *for some* c, s *such that* $c - s > 2^{-\mathcal{O}(k(n))}$.

Proof. We are given the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ PSD matrix H, and it is promised that the smallest eigenvalue λ_{min} of H is either at most a or at least b. Merlin would like to convince us that $\lambda_{min} \leq a$; he will send us a purported k-qubit eigenstate $|\psi\rangle$ of H with eigenvalue λ_{min}. Choose $t = \pi/(\mathrm{poly}(n)\|H\|_{max}) \leq \pi/\|H\|$; then all eigenvalues of Ht lie in the range $[0, \pi]$, and the output of phase estimation on $\exp(-iHt)$ will be unambiguous. We perform, on ψ, phase estimation of $\exp(-iHt)$ with one bit of precision:

$$\begin{array}{c} |0\rangle - \boxed{H} - \bullet - \boxed{H} - \quad \frac{1+e^{-i\lambda t}}{2}|0\rangle + \frac{1-e^{-i\lambda t}}{2}|1\rangle \\[2mm] |\psi\rangle - \boxed{e^{-iHt}} - \quad |\psi\rangle \end{array} \qquad (7)$$

Here the H gates on the first qubit are Hadamard gates (and have nothing to do with the matrix H). Theorem 12 gives an implementation of $\exp(-iHt)$ up to error $\epsilon = 2^{-\Theta(k(n))}$ using $\mathrm{poly}(n)$ operations and $\mathcal{O}(k(n))$ space.

In Circuit (7) we've assumed $|\psi\rangle$ is an eigenstate of H with eigenvalue λ. If we measure the control qubit at the end, the probability we obtain 0 is $(1+\cos(\lambda t))/2$. Therefore if ψ is a eigenstate with eigenvalue at most a, we can verify this with probability at least $c = (1+\cos(at))/2-\epsilon$, where ϵ is the error in the implementation of $\exp(-iHt)$. Otherwise if $\lambda_{min} \geq b$, no state ψ will be accepted with probability more than $s = (1 + \cos(bt))/2 + \epsilon$. The separation between c and s is at least

$$(\cos(at) - \cos(bt)) - 2\epsilon = 2\sin\left(\frac{(a+b)t}{2}\right)\sin\left(\frac{(b-a)t}{2}\right) - 2\epsilon \geq 2^{-\mathcal{O}(k)} \qquad (8)$$

since $\sin x = \Omega(x)$ for $x \in [0, 1]$, $(a+b)t \geq (b-a)t = 2^{-\mathcal{O}(k(n))}$, as long as we choose $\epsilon = 2^{-\Theta(k(n))}$ to be sufficiently small enough. This therefore gives a $(\mathrm{poly}, \Theta(k(n)))$-*bounded* $\mathsf{QMA}_{\Theta(k)}(c,s)$ protocol for $c - s = 2^{-\mathcal{O}(k(n))}$, as desired. $\qquad\square$

Lemma 21. $\bigcup_{c-s\geq 2^{-\mathcal{O}(k)}} (\mathrm{poly},\mathcal{O}(k))\text{-}bounded\ \mathsf{QMA}_{\mathcal{O}(k)}(c,s) \subseteq \mathsf{BQ_U SPACE}[k(n)]$.

Proof sketch. We only give a high level overview of the proof here; for the complete proof see Appendix D. The core of the proof is to develop and use new *space-efficient* QMA error reduction procedures. Our procedures are based on the "in-place" QMA amplification procedure of Marriott and Watrous [26], which allows the error in a QMA proof system to be reduced without requiring additional copies of the witness state. This was improved by Nagaj, Wocjan, and Zhang [27], whose phase-estimation based procedure reduces the error to 2^{-r} using only $\mathcal{O}\left(r\log\frac{1}{c-s}\right)$ additional qubits and $\mathcal{O}(r/(c-s))$ repetitions of the circuit and its inverse. We derive a procedure (Lemma 32) that gives the same error bounds while using only $\mathcal{O}\left(r+\log\frac{1}{c-s}\right)$ additional qubits, but still using only $\mathcal{O}(r/(c-s))$ repetitions of the circuit; the improved space bound will be required for our purposes[2].

Thus we can amplify the gap in our QMA protocols to still use $\mathcal{O}(k)$ space, but with completeness $1-2^{-\mathcal{O}(k)}$ and soundness $2^{-\mathcal{O}(k)}$. We can now replace the witness by the completely mixed state (or alternatively half of many EPR pairs), which gives us a computation with *no* witness such that the resulting completeness and soundness are both exponentially small, but are still separated by $2^{-\mathcal{O}(k)}$. Finally, we can once again apply our space-efficient amplification procedure to this witness-free protocol, obtaining a computation in $\mathsf{BQ_U SPACE}[\mathcal{O}(k)]$. $\qquad\square$

Lemma 22. $\mathcal{O}(k(n))\text{-}Minimum\ Eigenvalue$ is $\mathsf{BQ_U SPACE}[k(n)]\text{-}hard\ under\ classical\ poly\text{-}time$ $\mathcal{O}(k(n))\text{-}space\ reductions.$

Proof sketch. Again we only give an overview; see Appendix E for the full proof. Recall that our uniformity condition on $k(n)$-*Minimum Eigenvalue* implies that every language in $k(n)$-*Minimum Eigenvalue* can be decided by a quantum circuit of size at most $2^{\mathcal{O}(k(n))}$. We first use our space-efficient error reduction procedure to amplify the gap; then we apply a variant of Kitaev's clock construction [25] to construct a Hamiltonian from this amplified circuit. We use a *binary* clock instead of a unary one to save space; since the number of gates is at most $2^{\mathcal{O}(k(n))}$, the clock only needs to be of size $\mathcal{O}(k(n))$, and the total dimension of the system is $2^{\mathcal{O}(k(n))}$ as required. Therefore the Hamiltonian is not local, but it remains sparse (with only a constant number of nonzero terms in each row). Kitaev's analysis then shows that we can obtain a gap inverse polynomial in the circuit size, or inverse exponential in $k(n)$. $\qquad\square$

Proof of Theorems 18 and 19. Immediate from Lemmas 20, 21, and 22. $\qquad\square$

Note the polynomial space case in Theorem 19 is Corollary 3, that $\mathsf{PreciseQMA} = \mathsf{PSPACE}$.

Finally, we end with two results particular to the polynomial space case. First of all, in the equality $\mathsf{PreciseQMA} = \mathsf{PSPACE}$, we can actually achieve perfect completeness ($c = 1$) for the QMA proof protocol, assuming the underlying gate set contains the Hadamard and Toffoli gates. Moreover for perfect completeness we do not require that $c - s > 2^{-\mathrm{poly}}$:

Proposition 23. *Let* $\mathsf{QMA}(c,s) = (\mathrm{poly},\mathrm{poly})\text{-}bounded\ \mathsf{QMA}_{\mathrm{poly}}(c,s)$. *Then*

$$\mathsf{PSPACE} = \mathsf{QMA}(1, 1-2^{-\mathrm{poly}}) = \bigcup_{s<1}\mathsf{QMA}(1,s),$$

where we assume that the gateset we use contains the Hadamard and Toffoli gates. In the last term, the union is taken over all functions $s(n)$ such that $s(n) < 1$ for all n.

---

[2]In recent work we improved this result to achieve such amplification using only $\log\frac{r}{c-s}$ additional space [17].

The containment $\mathsf{QMA}(1, s) \subseteq \mathsf{PSPACE}$ is known [21]. We prove this proposition in Appendix F.

Our second result concerns the QMA-complete Local Hamiltonian problem. We show that if we allow the promise gap to be exponentially small, then the problem becomes PSPACE-complete.

Definition 24 (*Precise k-Local Hamiltonian*)**.** Given as input is a k-local Hamiltonian $H = \sum_{j=1}^{r} H_j$ acting on n qubits, satisfying $r \in \mathrm{poly}(n)$ and $\|H_j\| \leq \mathrm{poly}(n)$, and numbers $a < b$ satisfying $b - a > 2^{-\mathrm{poly}(n)}$. It is promised that the smallest eigenvalue of H is either at most a or at least b. Output 1 if the smallest eigenvalue of H is at most a, and output 0 otherwise.

Theorem 25. *For any $3 \leq k \leq \mathcal{O}(\log(n))$, Precise k-Local Hamiltonian is* $\mathsf{PreciseQMA}$*-complete, and hence* PSPACE*-complete.*

See Appendix G for a proof. Combined with the perfect completeness results of Appendix F, this will also give a proof that determining whether a local Hamiltonian is frustration-free is a PSPACE-complete problem (Theorem 39 in Appendix G).

5 Complete problems for time- and space- bounded classes

As we noted in the introduction, variants of the problems we consider are already known to be complete for other time-bounded quantum complexity classes. For example, consider the problem of inverting an efficiently encoded $2^{\mathcal{O}(k(n))} \times 2^{\mathcal{O}(k(n))}$ matrix with condition number at most $\kappa(n)$. If $\kappa(n), k(n) = \mathrm{poly}(n)$, this problem is BQP-complete [20]. Theorem 14 says that this problem is instead $\mathsf{BQ_USPACE}[\mathcal{O}(k)]$-complete if $\kappa = 2^{\mathcal{O}(k)}$. Similarly, consider the problem of determining whether the minimum eigenvalue of an efficiently encoded $2^{\mathcal{O}(k(n))} \times 2^{\mathcal{O}(k(n))}$ matrix is at least b or at most a, with $b - a = g(n)$. If $g = 1/\mathrm{poly}$ and $k = \mathrm{poly}$ then this problem is QMA-complete [25, 2]. Theorem 18 says that this problem is instead $\mathsf{BQ_USPACE}[\mathcal{O}(k)]$-complete if $g = 2^{-\mathcal{O}(k)}$.

In both of the problems we consider, we have two parameters that we can vary: for matrix inversion, the condition number κ and the matrix size k; and for minimum eigenvalue, the promise gap size $g = b - a$ and the matrix size k. Varying these two parameters independently gives complete problems for quantum classes that are simultaneously bounded in time and space.

Theorem 26. *Consider the class of problems solvable by a unitary quantum algorithm using* $\mathrm{poly}(T(n))$ *gates and* $\mathcal{O}(k(n))$ *space, where* $\Omega(\log(n)) \leq k(n) \leq T(n) \leq 2^{\mathcal{O}(k)} \leq 2^{\mathrm{poly}(n)}$. *This class has the following complete problem under classical* $\mathrm{poly}(n)$*-time and* $\mathcal{O}(k(n))$*-space reductions:*

Given as input is the size-n efficient encoding of a $2^{\mathcal{O}(k)} \times 2^{\mathcal{O}(k)}$ *positive semidefinite matrix* H *with a known upper bound* $\kappa = \mathrm{poly}(T)$ *on the condition number, so that* $\kappa^{-1}I \preceq H \preceq I$, *and* $s, t \in \{0, 1\}^{k(n)}$. *It is promised that either* $|H^{-1}(s, t)| \geq b$ *or* $|H^{-1}(s, t)| \leq a$ *for some constants* $0 \leq a < b \leq 1$; *determine which is the case.*

Theorem 27. *For functions $k(n)$, $T(n)$ satisfying $\Omega(\log(n)) \leq k(n) \leq T(n) \leq 2^{\mathcal{O}(k)} \leq 2^{\mathrm{poly}(n)}$,*

$$\bigcup_{c-s \geq \frac{1}{\mathrm{poly}(T)}} (\mathrm{poly}(n), \mathcal{O}(k))\text{-}bounded\ \mathsf{QMA}_{\mathcal{O}(k)}(c, s) = (\mathrm{poly}(T), \mathcal{O}(k))\text{-}bounded\ \mathsf{QMA}_{\mathcal{O}(k)}(2/3, 1/3)$$

Furthermore, the following problem is complete for this class under classical $\mathrm{poly}(n)$*-time and* $\mathcal{O}(k(n))$*-space reductions:*

Given as input is the size-n efficient encoding of a $2^{\mathcal{O}(k)} \times 2^{\mathcal{O}(k)}$ *PSD matrix* H, *such that* $\|H\|_{max} = \max_{s,t} |H(s, t)|$ *is at most a constant. Let* λ_{min} *be the minimum eigenvalue of* H. *It is promised that either* $\lambda_{min} \leq a$ *or* $\lambda_{min} \geq b$, *where* $a(n)$ *and* $b(n)$ *are numbers such that* $b - a \geq 1/\mathrm{poly}(T)$. *Output 1 if* $\lambda_{min} \leq a$, *and output 0 otherwise.*

We omit the proofs; they are straightforward generalizations of the proofs in our paper. These results interpolate between the time-bounded and space-bounded case: when $T = \text{poly}(k)$ the time-bound dominates and we obtain a time-bounded class; while when $T = 2^{\mathcal{O}(k)}$ we obtain a space-bounded class. Note that when $T = 2^{\mathcal{O}(k)}$ then the complexity class in Theorem 27 is equal to $\mathsf{BQ_USPACE}[\mathcal{O}(k)]$, as shown in Theorem 19.

6 Open Problems

This work leaves open several questions that may lead to interesting follow-up work:
1. Can we use our $\mathsf{PreciseQMA} = \mathsf{PSPACE}$ result to prove upper or lower bounds for other complexity classes?
2. Here we have shown $\mathsf{PreciseQMA} = \mathsf{PSPACE}$. Ito, Kobayashi and Watrous have shown that QIP with doubly-exponential small completeness-soundness gap is equal to EXP [21]. What can be said about the power of QIP with exponentially small completeness-soundness gap?
3. In this paper we studied unitary quantum space complexity classes, and showed that $k(n)$-*Well-conditioned Matrix Inversion* and $k(n)$-*Minimum Eigenvalue* characterize unitary quantum space complexity. Can similar hardness results be shown for non-unitary quantum space complexity classes?

7 Acknowledgements

A More details on space-bounded computation

For this section, it would be helpful to keep in mind that we always assume the space bound $k(n)$ always satisfies $\Omega(\log(n)) \leq k(n) \leq \text{poly}(n)$.

We start with the definitions of classical bounded space computation. In discussion of space-bounded classes, we usually consider Turing machines with two tapes, a read-only input tape and a work tape; only the space used on the work tape is counted. For $k : \mathbb{N} \to \mathbb{N}$, a function $f : \{0,1\}^* \to \{0,1\}^*$ is said to be computable in $k(n)$ space if any bit of $f(x)$ can be computed by a deterministic Turing machine using space $\mathcal{O}(k(|x|))$ on the work tape. For example, L is the class of functions that can be computed in $\mathcal{O}(\log n)$ space.

We now discuss quantum space-bounded complexity classes; for a fuller discussion see [39]. A straightforward way to define quantum space-bounded classes is to consider a Turing machine with three tapes: a read-only classical input tape, a classical work tape, and a quantum work tape (with two heads) consisting of qubits. This is the model considered in [32] and [38], except that they allow intermediate measurements (and [38] allows even more general quantum operations). In this work we consider only computations with no intermediate measurements: we can therefore impose that there are no measurements on the quantum work tape until the register reaches a specified end state, following which a single measurement is performed on the quantum tape and the algorithm accepts or rejects according to the measurement. Therefore the operations performed

by the algorithm will not depend on the quantum tape, since there is no way to read information out of it until the end of the algorithm.

Instead of working with Turing machines, in quantum computation it is much more customary (and convenient) to work with quantum circuits. For the setup above, since the operations on the quantum tape are completely classically controlled, we can equivalently consider a quantum circuit generated by a classical space-bounded Turing machine that computes the quantum gates one-by-one and applies them in sequence. If the classical Turing machine is $\mathcal{O}(k(n))$-space bounded, it has at most $2^{\mathcal{O}(k)}$ configurations, and therefore there are at most $2^{\mathcal{O}(k)}$ quantum gates in the circuit.

Moreover, the $\mathcal{O}(k)$-space bounded classical Turing machine can be replaced by a classical circuit on $\mathcal{O}(k)$ bits, such that there is a poly(n)-time $\mathcal{O}(k)$-space Turing machine that on input i generates the i-th gate of the circuit (see e.g. [4, Section 6.8]). The classical circuit can then be bundled into the quantum circuit, and we obtain a quantum circuit with at most $2^{\mathcal{O}(k)}$ gates, such that each individual gate can be generated in classical poly(n)-time and $\mathcal{O}(k)$-space. This justifies the definition of the complexity class $\mathsf{Q_USPACE}[k(n)](c,s)$:

Definition 4. Let $k(n)$ be a function satisfying $\Omega(\log(n)) \leq k(n) \leq \text{poly}(n)$. A promise problem $L = (L_{yes}, L_{no})$ is in $\mathsf{Q_USPACE}[k(n)](c,s)$ if there exists a poly$(|x|)$-time $\mathcal{O}(k)$-space uniformly generated family of quantum circuits $\{Q_x\}_{x \in \{0,1\}^*}$, where each circuit $Q_x = U_{x,T}U_{x,T-1}\cdots U_{x,1}$ has $T = 2^{\mathcal{O}(k)}$ gates, and acts on $\mathcal{O}(k(|x|))$ qubits, such that:
If $x \in L_{yes}$:
$$\left\langle 0^k \middle| Q_x^\dagger |1\rangle\langle 1|_{out} Q_x \middle| 0^k \right\rangle \geq c. \tag{9}$$

Whereas if $x \in L_{no}$:
$$\left\langle 0^k \middle| Q_x^\dagger |1\rangle\langle 1|_{out} Q_x \middle| 0^k \right\rangle \leq s. \tag{10}$$

Here *out* denotes a single qubit we measure at the end of the computation; no intermediate measurements are allowed. Furthermore, we require c and s to be computable in classical $\mathcal{O}(k(n))$-space.

B   Proof that $\mathcal{O}(k(n))$-*Well-conditioned Matrix Inversion* $\in \mathsf{BQ_USPACE}[k(n)]$

Theorem 15. *Fix functions $k(n)$, $\kappa(n)$, and $\epsilon(n)$. Suppose we are given the size-n efficient encoding of a $2^{k(n)} \times 2^{k(n)}$ PSD matrix H such that $\kappa^{-1}I \preceq H \preceq I$. We are also given $\text{poly}(n)$-time $\mathcal{O}(k + \log(\kappa/\epsilon))$-space uniform quantum circuits U_a and U_b acting on k qubits and using at most T gates. Let $U_a|0\rangle^{\otimes k(n)} = |a\rangle$ and $U_b|0\rangle^{\otimes k(n)} = |b\rangle$. The following tasks can be performed with $\text{poly}(n)$-time $O(k + \log(\kappa/\epsilon))$-space uniformly generated quantum circuits with $\text{poly}(T, k, \kappa, 1/\epsilon)$ gates and $\mathcal{O}(k + \log(\kappa/\epsilon))$ qubits:*
1. *With at least constant probability, output an approximation of the quantum state $H^{-1}|b\rangle/\|H^{-1}|b\rangle\|$ up to error ϵ.*
2. *Estimate $\|H^{-1}|b\rangle\|$ up to precision ϵ.*
3. *Estimate $|\langle a|H^{-1}|b\rangle|$ up to precision ϵ.*

These circuits do not require intermediate measurements.

For convenience we also restate the following lemma:

Lemma 16 (Implicit in [20, 32]). *There is a poly-time $\mathcal{O}(k + \log(\kappa/\epsilon'))$-space uniform quantum unitary transformation W_H over $k + \ell = \mathcal{O}(k + \log(\kappa/\epsilon'))$ qubits and using $\text{poly}(k, \kappa/\epsilon')$ gates, such that for any k-qubit input state $|b\rangle$,*
$$W_H(|0\rangle^{\otimes \ell} \otimes |b\rangle) = \alpha|0\rangle_{out} \otimes |\psi_b\rangle + \sqrt{1-\alpha^2}|1\rangle_{out} \otimes |\psi_b'\rangle, \tag{11}$$

where $|\psi_b\rangle$ and $|\psi_b'\rangle$ are normalized states such that $\||\psi_b\rangle - |0\rangle^{\otimes \ell-1} \otimes \frac{H^{-1}|b\rangle}{\|H^{-1}|b\rangle\|}\| \le \epsilon'$, α is a positive number satisfying $|\alpha - \frac{\|H^{-1}|b\rangle\|}{\kappa}| \le \epsilon'$, and "out" is a 1-qubit register.

Before we start the proof, we note that for ease of exposition we will actually use a limited number of intermediate measurements of up to $\mathcal{O}(k + \log(\kappa/\epsilon))$ qubits. These intermediate measurements are not necessary, since they can be deferred to the end of the computation using only $\mathcal{O}(k + \log(\kappa/\epsilon))$ extra space, which fits within our space bound. (It is only when the number of intermediate measurements used is superlinear in the amount of space available that we cannot defer measurements.)

Proof of Theorem 16. We first show the first item, i.e. generating the state $H^{-1}|b\rangle/\|H^{-1}|b\rangle\|$. Choose $\epsilon' = \mathcal{O}(\epsilon/\kappa)$ in the statement of Lemma 16, keeping in mind for the rest of the proof that $\log(\kappa/\epsilon') = \mathcal{O}(\log(\kappa/\epsilon))$. Note that $|\psi_b\rangle$ can be obtained by computing $W_H U_b |0\rangle_{all} = W_H(|0\rangle_{anc} \otimes |b\rangle)$, and then postselecting on the output qubit being in state $|0\rangle$. To obtain $|\psi_b\rangle$ with high probability we can repeat this procedure many times until success. We can then get a good approximation to $\frac{H^{-1}|b\rangle}{\|H^{-1}|b\rangle\|}$ by tracing out the other ancilla qubits. For our setting we would like to get by with a low space requirement and without using intermediate measurements, so instead of repeating until success, we will apply amplitude amplification to the above unitary W_H. Define the projectors Π_0 and Π_1 by

$$\Pi_0 = |0\rangle\langle 0|_{anc} \otimes |b\rangle\langle b| \tag{12}$$

$$\Pi_1 = W_H^\dagger(|0\rangle\langle 0|_{out} \otimes I)W_H \tag{13}$$

Define $|v\rangle = |0\rangle_{anc} \otimes |b\rangle$, and write

$$|v\rangle = \sin\theta|w\rangle + \cos\theta|w^\perp\rangle, \tag{14}$$

$$|w\rangle = \sin\theta|v\rangle + \cos\theta|v^\perp\rangle \tag{15}$$

where $|v^\perp\rangle$, $|w\rangle$, and $|w^\perp\rangle$ are normalized states such that

$$\Pi_0|v\rangle = |v\rangle, \quad \Pi_0|v^\perp\rangle = 0 \tag{16}$$

$$\Pi_1|w\rangle = |w\rangle, \quad \Pi_1|w^\perp\rangle = 0. \tag{17}$$

Note that

$$(\langle 0|_{anc} \otimes I)W_H(|0\rangle\langle 0|_{anc} \otimes |b\rangle\langle b|)(|0\rangle_{anc} \otimes |b\rangle) = W_H\Pi_1\Pi_0(|0\rangle_{anc} \otimes |b\rangle) \propto W_H|w\rangle, \tag{18}$$

and therefore $W_H|w\rangle$ is the postmeasurement state we desire. The success probability of the postselection step is simply $\langle v|\Pi_0\Pi_1\Pi_0|v\rangle = \alpha^2 = \sin^2\theta$.

Consider now the operator $R = -(I - 2\Pi_1)(I - 2\Pi_0)$; by analogy from Grover's algorithm, it is easy to see that R is a rotation operator with angle 2θ. It has eigenvalues $e^{\pm i2\theta}$, with the following eigenvectors:

$$R|\psi_+\rangle = e^{i2\theta}|\psi_+\rangle, \quad |\psi_+\rangle \equiv \frac{1}{\sqrt{2}}\left(|v\rangle + i|v^\perp\rangle\right) = \frac{ie^{-i\theta}}{\sqrt{2}}\left(|w\rangle - i|w^\perp\rangle\right) \tag{19}$$

$$R|\psi_-\rangle = e^{-i2\theta}|\psi_-\rangle, \quad |\psi_-\rangle \equiv \frac{1}{\sqrt{2}}\left(|v\rangle - i|v^\perp\rangle\right) = \frac{-ie^{i\theta}}{\sqrt{2}}\left(|w\rangle + i|w^\perp\rangle\right) \tag{20}$$

Note our initial state $|v\rangle$ is a uniform superposition of the eigenstates of R: $|v\rangle = 2^{-1/2}(|\psi_+\rangle + |\psi_-\rangle)$. We perform phase estimation of the rotation operator R on the state $|v\rangle$ with precision $\mathcal{O}(\epsilon')$ and failure probability $\delta = 2^{-\mathcal{O}(k)}\mathrm{poly}(\kappa/\epsilon)$; this requires $\mathrm{poly}(T, k, \kappa, 1/\epsilon)$ gates and $\mathcal{O}(k + \log(\kappa/\epsilon))$ extra ancilla qubits to perform, and with probability $1 - \delta$ outputs an estimate for either θ or $-\theta$ with error $\mathcal{O}(\epsilon')$. Taking the absolute value of the sine of the output, we obtain an estimate of $\alpha = \sin\theta$ with error $\mathcal{O}(\epsilon') = \mathcal{O}(\epsilon/\kappa)$. Since $|\alpha - \frac{\|H^{-1}|b\rangle\|}{\kappa}| \le \epsilon'$, this allows us to calculate an estimate for $\|H^{-1}|b\rangle\|$ up to precision $\mathcal{O}(\kappa\epsilon') = \mathcal{O}(\epsilon)$, with probability $1 - 2^{-\mathcal{O}(k)}\mathrm{poly}(\kappa/\epsilon)$. This completes the second task.

To complete the first task, we note that the residual state after the phase estimation procedure is still a linear combination of $|\psi_+\rangle$ and $|\psi_-\rangle$. If the phase estimation procedure above had output an estimate $\tilde{\theta}_+ \approx_{\mathcal{O}(\epsilon')} \theta$ (this happens with probability at least $(1-\delta)/2$), then because the we set the failure probability to be δ, the residual state must be δ-close to $|\psi_+\rangle$. From (19) we see that this residual state has constant overlap with $|w\rangle$. Similarly, if the phase estimation procedure had instead output an estimate close to $-\theta$ (this happens with probability at least $(1-\delta)/2$), then we obtain a residual state δ-close to $|\psi_-\rangle$, and hence with constant overlap with $|w\rangle$. By applying W_H to our state and verifying the ancilla qubits are all zero, we obtain the desired state $W_H|w\rangle \approx H^{-1}|b\rangle$ with constant probability.

Finally, if an estimate for $|\langle a|H^{-1}|b\rangle|$ is desired, we can consider instead the following modification to Π_0 and Π_1:

$$\Pi_0 = |0\rangle\langle0|_{anc} \otimes |b\rangle\langle b| \tag{21}$$

$$\Pi_1' = W_H^\dagger(|0\rangle\langle0|_{out} \otimes I)(I_{anc} \otimes |a\rangle\langle a|)(|0\rangle\langle0|_{out} \otimes I)W_H \tag{22}$$

Since $(|0\rangle\langle0|_{out} \otimes I)W_H(|0\rangle^{\otimes\ell} \otimes |b\rangle) = \alpha|0\rangle_{out} \otimes |\psi_b\rangle$, we see that

$$\langle v|\Pi_0\Pi_1'\Pi_0|v\rangle = \alpha^2 \left|\left(\langle0|^{\otimes\ell-1} \otimes \langle a|\right)|\psi_b\rangle\right|^2. \tag{23}$$

$\langle v|\Pi_0\Pi_1'\Pi_0|v\rangle$ can be estimated in the same way that α has been estimated. Recalling that $\||\psi_b\rangle - |0\rangle^{\otimes\ell-1} \otimes \frac{H^{-1}|b\rangle}{\|H^{-1}|b\rangle\|}\| \le \epsilon'$ and $|\alpha - \frac{\|H^{-1}|b\rangle\|}{\kappa}| \le \epsilon'$, this allows us to estimate $|\langle a|H^{-1}|b\rangle|$ to $\mathcal{O}(\epsilon)$ precision. $\qquad\square$

C $\quad k(n)$-*Well-conditioned Matrix Inversion* is hard for $\mathsf{BQSPACE}[k(n)]$

We begin our hardness proof by considering the following simple hard problem for $\mathsf{BQ_USPACE}[k(n)]$:

Definition 28 ($k(n)$-*Quantum Circuit Acceptance*)**.** Given as input is the size-n classical algorithm (Turing machine) that generates a quantum circuit Q acting on $k(n)$ qubits with $T = 2^{\mathcal{O}(k(n))}$ 1- or 2-qubit gates: on input i, the algorithm outputs the gate i-th gate of the circuit in polynomial time and $\mathcal{O}(k(n))$ space. It is promised that either the matrix entry $|\langle0^{\otimes k(n)}|Q|0^{\otimes k(n)}\rangle| \ge 2/3$ or $|\langle0^{\otimes k(n)}|Q|0^{\otimes k(n)}\rangle| \le 1/3$; determine which is the case.

Lemma 29 (Implicit in [5, 14])**.** $\mathcal{O}(k(n))$-*Quantum Circuit Acceptance is* $\mathsf{BQ_USPACE}[k(n)]$-*hard under classical* $\mathrm{poly}(n)$-*time,* $\mathcal{O}(k(n))$-*space reductions.*

Proof. This lemma is implicit in e.g., [5, 14]. We include the proof here for completeness. Suppose we are given an $x \in \{0,1\}^n$ and would like to determine if $x \in L_{yes}$ for some $L = \{L_{yes}, L_{no}\} \in \mathsf{BQ_USPACE}[k]$. There is a quantum circuit on $k(n)$ qubits, $Q_x = U_T U_{T-1} \cdots U_1$ of size $T = 2^{\mathcal{O}(k(n))}$ that decides x. That is,

$$Q_x|0^{\otimes k(n)}\rangle = \sqrt{p_x}|1\rangle_{out}|\psi_x^1\rangle + \sqrt{1-p_x}|0\rangle_{out}|\psi_x^0\rangle. \tag{24}$$

14

where *out* indicates the designated output qubit, and $|\psi_x^1\rangle$, $|\psi_x^0\rangle$ are $(k-1)$-qubit states; p_x is the probability that the computation accepts, so $p_x \geq 2/3$ if $x \in L_{yes}$ and $p_x \leq 1/3$ if $x \in L_{no}$. Note that the uniformity condition on $\mathsf{BQ_USPACE}[k]$ guarantees the existence of a classical algorithm that generates Q_x.

We now describe a reduction which creates a related circuit $\tilde{Q}_x$ with a single matrix entry that is proportional to the acceptance probability of Q_x. This new circuit $\tilde{Q}_x$ takes the same number of input qubits as Q_x as well as an additional ancillary qubit. $\tilde{Q}_x$ runs Q_x, then using a single CNOT gate copies the state of the output qubit to the ancillary qubit, flips the ancillary qubit, and finally applies the inverse, $Q_x^\dagger$, to the input qubits. It is straightforward to check that

$$\langle 0|\langle 0^{\otimes k(n)}|\tilde{Q}_x|0^{\otimes k(n)}\rangle|0\rangle = p_x. \tag{25}$$

Therefore knowing the single matrix entry $\langle 0|\langle 0^{\otimes k(n)}|\tilde{Q}_x|0^{\otimes k(n)}\rangle|0\rangle$ is sufficient to decide if $x \in L_{yes}$. Moreover, $\tilde{Q}_x$ can be computed from Q_x using polynomial time and $\mathcal{O}(k)$ space, and this completes the proof. $\square$

Theorem 30. $\mathcal{O}(k(n))$*-Well-conditioned Matrix Inversion is* $\mathsf{BQ_USPACE}[k(n)]$*-hard under classical reductions computable in polynomial time and* $\mathcal{O}(k(n))$ *space.*

Proof. We will show that $k(n)$*-Well-conditioned Matrix Inversion* is as hard as $k(n)$*-Quantum Circuit Acceptance*. Given an instance of the latter, i.e., a circuit on $k(n)$ qubits, $Q = U_T U_{T-1} \cdots U_1$ with $T = 2^{\mathcal{O}(k(n))}$, define the following unitary of dimension $3T2^k$:

$$U = \sum_{t=1}^T |t+1\rangle\langle t| \otimes U_t + |t+T+1\rangle\langle t+T| \otimes I + |t+2T+1 \bmod 3T\rangle\langle t+2T| \otimes U_{T+1-t}^\dagger$$

Crucially, note that for any t in the range $[T, 2T]$ and any state $|\psi\rangle$ on $k(n)$ qubits:

$$U^t|1\rangle|\psi\rangle = |t+1\rangle \otimes Q|\psi\rangle \tag{26}$$

Furthermore $U^{3T} = I$, a fact we will soon exploit. We now construct the Hermitian matrix:

$$H = \begin{bmatrix} 0 & I - Ue^{-1/T} \\ I - U^\dagger e^{-1/T} & 0 \end{bmatrix} \tag{27}$$

H has dimension $N = 6T2^k$ and condition number $\kappa \leq 2T = 2^{\mathcal{O}(k)}$. Notice that given as input a description of Q we can compute each entry of H to within $2^{-\mathcal{O}(k(n))}$ error in $\mathcal{O}(k(n))$ space, via space efficient algorithms for exponentiation [29]. Furthermore, H^{-1} is the following matrix:

$$\begin{bmatrix} 0 & \left(I - U^\dagger e^{-1/T}\right)^{-1} \\ \left(I - Ue^{-1/T}\right)^{-1} & 0 \end{bmatrix} \tag{28}$$

$\left(I - Ue^{-1/T}\right)^{-1}$ is just the power series $\sum_{j=0}^\infty U^j e^{-j/T} = \frac{1}{1-e^{-3}}\sum_{j=0}^{3T-1} U^j e^{-j/T}$, where we've used $U^{3T} = I$. Therefore for any fixed $t \in [0, 3T-1]$,

$$\langle 0^{\otimes k(n)}|\langle t+1|\left(I - Ue^{-1/T}\right)^{-1}|1\rangle|0^{\otimes k(n)}\rangle = \frac{1}{1-e^{-3}}\langle 0^{\otimes k(n)}|\langle t+1|\left(\sum_{j'=0}^{3T-1} U^{j'} e^{-j'/T}\right)|1\rangle|0^{\otimes k(n)}\rangle \tag{29}$$

15

which is a particular entry in H^{-1}. In the second line we've used $j = 3Tx + j'$ for some integers x and $j' \in [0, 3T - 1]$, and that $U^{3T} = I$. For any $t \in [T, 2T]$, as a consequence of Equation 26 the above quantity equals

$$\frac{e^{-t/T}}{1 - e^{-3}} \langle 0^{\otimes k(n)} | Q | 0^{\otimes k(n)} \rangle. \tag{30}$$

In particular, an estimation of this entry of H^{-1} will solve $k(n)$-*Quantum Circuit Acceptance*. $\square$

D Proof of Lemma 21

D.1 In-place gap amplification of QMA protocols with phase estimation

We start out by proving the following lemma, which proves "in-place" gap amplification of QMA using phase estimation (see also the similar result of Nagaj et. al, Lemma 40 in Appendix H).

Lemma 31. *For any functions $t, k, r > 0$,*

$$(t, k)\text{-}bounded\ \mathsf{QMA}_m(c, s) \subseteq \left(\mathcal{O}\left(\frac{t2^r}{c - s}\right), \mathcal{O}\left(k + r + \log\left(\frac{1}{c - s}\right)\right) \right)\text{-}bounded\ \mathsf{QMA}_m(1 - 2^{-r}, 2^{-r}).$$

Proof. Let $L = (L_{yes}, L_{no})$ be a promise problem in $\mathsf{QMA}(c, s)$ and $\{V_x\}_{x \in \{0,1\}^*}$ the corresponding uniform family of verification circuits. Define the projectors:

$$\Pi_0 = I_m \otimes \left|0^k\right\rangle\left\langle 0^k\right|, \quad \Pi_1 = V_x^\dagger \left(|1\rangle\langle 1|_{out} \otimes I_{m+k-1}\right) V_x \tag{31}$$

and the corresponding reflections $R_0 = 2\Pi_0 - I, R_1 = 2\Pi_1 - I$. Define $\phi_c = \arccos\sqrt{c}/\pi$ and $\phi_s = \arccos\sqrt{s}/\pi$ (recalling that these functions can be computed to precision $\mathcal{O}(c - s)$ in space $\mathcal{O}(\log[1/(c - s)])$). Now consider the following procedure:

1. Perform phase estimation of the operator $R_1 R_0$ on the state $|\psi\rangle \otimes \left|0^k\right\rangle$, with precision $\mathcal{O}(c - s)$ and failure probability 2^{-r}.
2. Output YES if the phase is at most $(\phi_c + \phi_s)/2$; otherwise output NO.

Phase estimation of an operator U up to precision a and failure probability ϵ requires $\alpha := \lceil \log_2(1/a) \rceil + \log_2[2 + 1/(2\epsilon)]$ additional ancilla qubits and $2^\alpha = \mathcal{O}(1/(a\epsilon))$ applications of the control-U operation (see e.g. [28]). Thus, the above procedure can be implemented by a circuit of size $\mathcal{O}(2^r t/(c - s))$ using $\mathcal{O}(r + \log[1/(c - s)])$ extra ancilla qubits. Using the standard analysis of in-place QMA error reduction [26, 27], it can be shown that this procedure has completeness probability at least $1 - 2^{-r}$ and soundness at most 2^{-r}. $\square$

In Appendix H we will prove the following stronger error reduction lemma that gives the same space bound but uses less time. This better time bound will be required for proving Lemma 22.

Lemma 32. *For any functions $t, k, r > 0$,*

$$(t, k)\text{-}bounded\ \mathsf{QMA}_m(c, s) \subseteq \left(\mathcal{O}\left(\frac{rt}{c - s}\right), \mathcal{O}\left(k + r + \log\left(\frac{1}{c - s}\right)\right) \right)\text{-}bounded\ \mathsf{QMA}_m(1 - 2^{-r}, 2^{-r}).$$

Thus, we get the following corollaries:

Corollary 33. *For any $r = \mathcal{O}(k)$, $\mathsf{Q_U SPACE}[k](c, c - 2^{-\mathcal{O}(k)}) \subseteq \mathsf{Q_U SPACE}[\Theta(k)](1 - 2^{-r}, 2^{-r})$.*

This corollary shows that error reduction is possible for unitary quantum $\mathcal{O}(k)$-space bounded classes, as long as the completeness-soundness gap is at least $2^{-\mathcal{O}(k)}$.

Proof. This follows from Lemma [32] by taking $m = 0$, $s = c - 2^{-\Theta(k)}$, and $r = \Theta(k)$. $\qquad\square$

Corollary 34.

$$(t, k)\text{-}bounded\ \mathsf{QMA}_m(c, c - 2^{-\Theta(k)}) \subseteq \left(\mathcal{O}\left(t2^{\Theta(k)}\right), \mathcal{O}\left(k\right)\right)\text{-}bounded\ \mathsf{QMA}_m(1 - 2^{-(m+2)}, 2^{-(m+2)}).$$

Proof. This follows from Lemma [32] by taking $s = c - 2^{-\Theta(k)}$ and $r = m + 2$. $\qquad\square$

D.2 Removing the witness of an amplified QMA protocol

Theorem 35. *For any function $t = 2^{\mathcal{O}(k+m)}$,*

$$(t, k)\text{-}bounded\ \mathsf{QMA}_m(1 - 2^{-(m+2)}, 2^{-(m+2)}) \subseteq \mathsf{Q_U SPACE}[k + m](3/4 \cdot 2^{-m}, 1/4 \cdot 2^{-m}).$$

Proof. The proof is very similar to that of [26, Theorem 3.6]. For any functions m, k, consider a problem $L \in (t, k)$-*bounded* $\mathsf{QMA}_m(1 - 2^{-(m+2)}, 2^{-(m+2)})$, and let $\{V'_x\}_{x \in \{0,1\}^*}$ be a uniform family of verification circuits for L with completeness $1 - 2^{-(m+2)}$ and soundness $2^{-(m+2)}$. For convenience, define the $2^m \times 2^m$ matrix:

$$Q_x := \left(I_{2^m} \otimes \langle 0^p|\right) V'^{\dagger}_x |1\rangle\langle 1|_{out} V'_x \left(I_{2^m} \otimes |0^p\rangle\right). \tag{32}$$

Q_x is positive semidefinite, and $\langle \psi | Q_x | \psi \rangle$ is the acceptance probability of V'_x on witness ψ. Thus

$$x \in L_{yes} \Rightarrow \mathrm{tr}[Q_x] \geq 1 - 2^{-(m+2)} \geq 3/4 \tag{33}$$

since the trace is at least the largest eigenvalue, and $m \geq 0$; likewise,

$$x \in L_{no} \Rightarrow \mathrm{tr}[Q_x] \leq 2^m \cdot 2^{-(m+2)} = 1/4 \tag{34}$$

since the trace is the sum of the 2^m eigenvalues, each of which is at most $2^{-(m+2)}$.

Therefore our problem reduces to determining whether the trace of Q_x is at least $3/4$ or at most $1/4$. Now we show that using the totally mixed state $2^{-m} I_m$ (alternatively, preparing m EPR pairs and taking a qubit from each pair) as the witness of the verification procedure encoded by Q_x, succeeds with the desired completeness and soundness bounds. The acceptance probability is given by $\mathrm{tr}(Q_x 2^{-m} I_m) = 2^{-m} \mathrm{tr}(Q_x)$, which is at least $2^{-m} \cdot 3/4$ if $x \in L_{yes}$, and at most $2^{-m} \cdot 1/4$ if $x \in L_{no}$. Thus we have reduced the problem L to determining if a quantum computation with *no* witness, acting on $k + m$ qubits, accepts with probability at least $3/4 \cdot 2^{-m}$ or at most $s' = 1/4 \cdot 2^{-m}$. $\qquad\square$

We can finally finish the proof of Lemma [21].

Proof of Lemma [21]. This follows from Corollary [34], Theorem [35], and Corollary [33]. $\qquad\square$

E $\mathcal{O}(k(n))$-*Minimum Eigenvalue* is hard for $\mathsf{BQ_U SPACE}[k(n)]$

In this section we prove Lemma [22], stated here for convenience.

Lemma 22. $\mathcal{O}(k(n))$-*Minimum Eigenvalue is* $\mathsf{BQ_U SPACE}[k(n)]$-*hard under classical poly-time $\mathcal{O}(k(n))$-space reductions.*

Proof. Let $L = (L_{yes}, L_{no})$ be a problem in $\mathsf{BQ_U SPACE}[k(n)]$, and suppose it has a verifier that uses $t = 2^{\mathcal{O}(k)}$ gates with completeness $2/3$ and soundness $1/3$. By Lemma 32, we can amplify the gap to get a new verifier circuit that uses $T = \mathcal{O}(rt)$ gates, and has completeness $c = 1 - 2^{-r}$ and soundness $s = 2^{-r}$. Choose $r = \mathcal{O}(k)$ large enough so that $c \geq 1 - 1/T^3$ and $s \leq 1/T^3$, and suppose $V_x = V_{x,T}V_{x,T-1}\cdots V_{x,1}$ is the new (gap-amplified) verifier circuit for L acting on k qubits. Consider the Kitaev clock Hamiltonian:

$$H = H_{in} + H_{prop} + H_{out} \tag{35}$$

defined on the Hilbert space $\mathbb{C}^{2^k} \otimes \mathbb{C}^{T+1}$, where

$$H_{in} = \left|0^k\right\rangle\!\left\langle 0^k\right| \otimes |0\rangle\langle 0|, \quad H_{out} = (|1\rangle\langle 1|_{out} \otimes I_{k-1}) \otimes |T\rangle\langle T| \tag{36}$$

$$H_{prop} = \sum_{j=1}^{T} \frac{1}{2}\left[-V_{x,j} \otimes |j\rangle\langle j-1| - V_{x,j}^\dagger \otimes |j-1\rangle\langle j| + I \otimes (|j\rangle\langle j| + |j-1\rangle\langle j-1|)\right]. \tag{37}$$

H is a sparse matrix - in fact, there are only a constant number of nonzero terms in each row. Since each gate $V_{x,j}$ can be computed in classical polynomial time and $k(n)$ space, it follows that so can the nonzero entries of H. Moreover, let λ_{min} be the minimum eigenvalue of H; then it was shown by Kitaev [25] that if $x \in L_{yes}$ then $\lambda_{min} \leq a = (1-c)/(T+1)$, while if $x \in L_{no}$ then $\lambda_{min} \geq b = (1-s)/T^3$. Since $c \geq 1 - 1/T^3$ and $s \leq 1/T^3$ we see that $2^{-\mathcal{O}(k)} < b - a$. $\qquad\square$

F   Achieving Perfect Completeness for PreciseQMA

We now consider the problem of achieving perfect completeness for PreciseQMA. Specifically, we will show the following:

Proposition 23. *Let* $\mathsf{QMA}(c, s) = (\mathrm{poly}, \mathrm{poly})\text{-bounded } \mathsf{QMA}_{\mathrm{poly}}(c, s)$. *Then*

$$\mathsf{PSPACE} = \mathsf{QMA}(1, 1 - 2^{-\mathrm{poly}}) = \bigcup_{s<1} \mathsf{QMA}(1, s),$$

where we assume that the gateset we use contains the Hadamard and Toffoli gates. In the last term, the union is taken over all functions $s(n)$ *such that* $s(n) < 1$ *for all* n.

Since $\mathsf{PSPACE} = \mathsf{PreciseQMA}$, this proposition shows that any PreciseQMA protocol can be reduced to a different PreciseQMA protocol with perfect completeness, i.e. in the YES case Arthur accepts Merlin's witness with probability 1. The reduction is rather roundabout, however, and it would be interesting to see if a more direct reduction can be found.

The second equality follows from the first equality and the result by [21] that $\mathsf{QMA}(1, s) \subseteq \mathsf{PSPACE}$. We will therefore only prove the first equality.

Looking back at Circuit 7, we see that we *almost* have perfect completeness in our protocol already - if the Hamiltonian simulation of e^{-iHt} could be done without error, then indeed the protocol has perfect completeness. Our strategy will be perform a different unitary that can be performed exactly, but, like e^{-iHt}, also allows us to use phase estimation to distinguish the eigenvalues of H.

Given a sparse Hamiltonian H (with at most d nonzero entries per row) and a number $X \geq \max_{j,\ell} |H_{j\ell}|$ that upper bounds the absolute value of entries of H, Andrew Childs defined an efficiently implementable quantum walk [7, 12]. Each step of the quantum walk is a unitary U with eigenvalues $e^{i\tilde{\lambda}}$, where

$$\tilde{\lambda} = \arcsin \frac{\lambda}{Xd} \tag{38}$$

with λ representing eigenvalues of H. Note that the YES case $\lambda = 0$ corresponds to $\tilde{\lambda} = 0$, and the NO case $\lambda \geq 2^{-g(n)}$ corresponds to $\tilde{\lambda} \geq 2^{-g(n)}/(Xd)$ since $\arcsin x \geq x$ for $|x| \leq 1$. In the latter case the $\tilde{\lambda}$ can be at most exponentially small, and therefore the stripped down version of phase estimation still suffices to tell the two cases apart with exponentially small probability.

We now note that the Hamiltonian H we obtain from the hardness reduction from PSPACE (Lemma 22) is of a very special form. Specifically, since $\mathsf{BQ_USPACE}[\text{poly}] = \mathsf{PSPACE}$, we can assume the verifier circuit V_x is deterministic, so it has completeness 1 and soundness 0. Moreover, all of its gates are classical, and hence all entries of the Kitaev clock Hamiltonian H are $0, \pm 1/2$, or 1.

For the matrix H satisfying the above, U can be implemented exactly with a standard gateset; perfect completeness of the protocol will then follow. If H is a $N \times N$ matrix (where $N = 2^n$), U is (see presentation in [9, Section 3.1 and Lemma 10]) a unitary defined on the enlarged Hilbert space $\mathbb{C}^{2N} \otimes \mathbb{C}^{2N} = (\mathbb{C}^N \otimes \mathbb{C}^2) \otimes (\mathbb{C}^N \otimes \mathbb{C}^2)$, as follows:

$$U = ST(I_{2N} \otimes (I_{2N} - 2|0\rangle\langle 0|_{2N}))T^\dagger \tag{39}$$

where the $2N$ subscript indicates a register of dimension $2N$, the unitary S swaps the two registers, and the unitary T is defined by

$$T = \sum_{j=0}^{N-1} \sum_{b \in \{0,1\}} (|j\rangle\langle j| \otimes |b\rangle\langle b|) \otimes |\varphi_{jb}\rangle\langle 0|_{2N} \tag{40}$$

with $|\varphi_{j1}\rangle = |0\rangle_N |1\rangle$ and

$$|\varphi_{jb}\rangle = \frac{1}{\sqrt{d}} \sum_{\ell \in F_j} |\ell\rangle \left(\sqrt{\frac{H_{j\ell}^*}{X}} |0\rangle + \sqrt{1 - \frac{|H_{j\ell}^*|}{X}} |1\rangle \right), \tag{41}$$

where F_j index the nonzero entries in the j-th row. Recall that for any j, ℓ, $H_{j\ell} = 0, \pm 1/2$, or 1, and hence we can take $X = 1$. If we furthermore assume d is a power of 2 (which we can always do by adding indices of zero entries to F_j), it is straightforward to see that both S and T can be implemented using just Hadamard gates and classical gates (Pauli-X, controlled-X, and Toffoli gates) - the latter of which can be implemented using just Toffoli gates and access to a qubit in the $|1\rangle$ state (which can be provided by the prover). Therefore U can be exactly implemented in any gateset that allows Hadamard gates and Toffoli gates to be implemented exactly.

G   Precise Local Hamiltonian Problem

Recall the following definition:

Definition 24 (*Precise k-Local Hamiltonian*)**.** Given as input is a k-local Hamiltonian $H = \sum_{j=1}^r H_j$ acting on n qubits, satisfying $r \in \text{poly}(n)$ and $\|H_j\| \leq \text{poly}(n)$, and numbers $a < b$ satisfying $b - a > 2^{-\text{poly}(n)}$. It is promised that the smallest eigenvalue of H is either at most a or at least b. Output 1 if the smallest eigenvalue of H is at most a, and output 0 otherwise.

In this section we will prove the following:

Theorem 25. *For any $3 \leq k \leq \mathcal{O}(\log(n))$, Precise k-Local Hamiltonian is* $\mathsf{PreciseQMA}$-*complete, and hence* PSPACE-*complete.*

Proof. This proof follows straightforwardly by adapting the proof of [25] and [24]. The proof of containment in PreciseQMA is identical to the containment of the usual Local Hamiltonian problem in QMA; see [25] for details.

To show PreciseQMA-hardness, we note that for a QMA-verification procedure with T gates, completeness c and soundness s, [24] reduces this to a 3-local Hamiltonian with lowest eigenvalue no more than $(1-c)/(T+1)$ in the YES case, or no less than $(1-s)/T^3$ in the NO case. For this to specify a valid *Precise Local Hamiltonian* problem we need that

$$\frac{1-s}{T^3} - \frac{1-c}{T+1} > 2^{-\text{poly}(n)}. \tag{42}$$

Recalling that we showed that perfect completeness can be assumed for PreciseQMA-hard problems, we can take $c = 1$, $s = 1 - 2^{-\text{poly}(n)}$ and the above inequality trivially holds. Hence any problem in PSPACE can be reduced to a *Precise 3-Local Hamiltonian* problem. □

In fact, even just testing whether a k-Local Hamiltonian is frustration-free is PSPACE-complete:

Definition 38 (*Frustration-Free k-Local Hamiltonian*)**.** Given as input is a k-local Hamiltonian $H = \sum_{j=1}^{r} H_j$ acting on n qubits, satisfying $r \in \text{poly}(n)$, each term H_j is positive semidefinite, and $\|H_j\| \leq \text{poly}(n)$. Output 1 if the smallest eigenvalue of H is zero, and output 0 otherwise.

Theorem 39. *Frustration-Free k-Local Hamiltonian is* PSPACE-*complete.*

Proof. The containment in PSPACE follows from the proof of the containment of the usual Local Hamiltonian problem in QMA [25], along with Proposition 23. PSPACE-hardness follows from the proof of Theorem 25, by taking $c = 1$ in the proof. □

H In-place gap amplification

In this appendix we will prove Lemma 32. To do so we first start out with the following weaker result:

Lemma 40 (Implicit in Nagaj, Wocjan, and Zhang [27])**.** *For any functions* $t, k, r > 0$,

$$(t,k)\text{-bounded } \mathsf{QMA}_m(c,s) \subseteq \left(\mathcal{O}\left(\frac{rt}{c-s} \right), \mathcal{O}\left(k + r\log\left(\frac{1}{c-s} \right) \right) \right)\text{-bounded } \mathsf{QMA}_m(1 - 2^{-r}, 2^{-r}).$$

Proof. Let $L = (L_{yes}, L_{no})$ be a promise problem in $\mathsf{QMA}(c,s)$ and $\{V_x\}_{x \in \{0,1\}^*}$ the corresponding uniform family of verification circuits. Define the projectors:

$$\Pi_0 = I_m \otimes \left| 0^k \right\rangle\left\langle 0^k \right| \tag{43}$$

$$\Pi_1 = V_x^\dagger \left(|1\rangle\langle 1|_{out} \otimes I_{m+k-1} \right) V_x \tag{44}$$

and the corresponding reflections:

$$R_0 = 2\Pi_0 - I, \quad R_1 = 2\Pi_1 - I. \tag{45}$$

Define $\phi_c = \arccos\sqrt{c}/\pi$ and $\phi_s = \arccos\sqrt{s}/\pi$ (recalling that these functions can be computed to precision $\mathcal{O}(c-s)$ in space $\mathcal{O}(\log[1/(c-s)])$). Now consider the following procedure:

1. Perform r trials of phase estimation of the operator $R_1 R_0$ on the state $|\psi\rangle \otimes \left| 0^k \right\rangle$, with precision $\mathcal{O}(c-s)$ and $1/16$ failure probability.

2. If the median of the r results is at most $(\phi_c + \phi_s)/2$, output YES; otherwise output NO.
Phase estimation of an operator U up to precision a and failure probability ϵ requires $\alpha :=$ $\lceil \log_2(1/a) \rceil + \log_2[2 + 1/(2\epsilon)]$ additional ancilla qubits and $2^\alpha = \mathcal{O}(1/(a\epsilon))$ applications of the control-U operation (see e.g. [28]). Thus, the above procedure, which uses r applications of phase estimation to precision $\mathcal{O}(c - s)$, can be implemented by a circuit of size $\mathcal{O}(rt/(c - s))$ using $\mathcal{O}(r \log[1/(c - s)])$ extra ancilla qubits. Using the standard analysis of in-place QMA error reduction [26, 27], it can be seen that this procedure has completeness probability at least $1 - 2^{-r}$ and soundness at most 2^{-r}. $\qquad\square$

We can now prove Lemma 32, which we restate below:

Lemma 32. *For any functions $t, k, r > 0$,*

$$(t, k)\text{-bounded } \mathsf{QMA}_m(c, s) \subseteq \left(\mathcal{O}\left(\frac{rt}{c - s} \right), \mathcal{O}\left(k + r + \log\left(\frac{1}{c - s} \right) \right) \right)\text{-bounded } \mathsf{QMA}_m(1 - 2^{-r}, 2^{-r}).$$

Proof.

$$
\begin{aligned}
(t, k)\text{-bounded } \mathsf{QMA}_m(c, s) &\subseteq \left(\mathcal{O}\left(\frac{t}{c - s} \right), \mathcal{O}\left(k + \log\left(\frac{1}{c - s} \right) \right) \right)\text{-bounded } \mathsf{QMA}_m(3/4, 1/4) \\
&\subseteq \left(\mathcal{O}\left(\frac{rt}{c - s} \right), \mathcal{O}\left(k + r + \log\left(\frac{1}{c - s} \right) \right) \right)\text{-bounded } \mathsf{QMA}_m(1 - 2^{-r}, 2^{-r})
\end{aligned}
$$

where the first line follows by taking $r = 2$ in Lemma 31, and the second line follows from Lemma 40. $\qquad\square$

I Proof sketch of $\mathsf{PQP}^{O_{PEPS}}_{\|,\mathbf{classical}} = \mathsf{PP}$

Since $\mathsf{PP} \subseteq \mathsf{BQP}^{O_{PEPS}}_{\|,\text{classical}} \subseteq \mathsf{PQP}^{O_{PEPS}}_{\|,\text{classical}}$ [30], we only need to show that $\mathsf{PQP}^{O_{PEPS}}_{\|,\text{classical}} \subseteq \mathsf{PP}$. In [30] it was noted that all PEPS can be seen as the output of a quantum circuit followed by a postselected measurement. Therefore $\mathsf{PQP}^{O_{PEPS}}_{\|,\text{classical}}$ corresponds to the problems that can be decided by a quantum circuit, followed by a postselected measurement (since the queries to O_{PEPS} are classical and nonadaptive, we can compose them into one single postselection), followed by a measurement. In the YES case the measurement outputs 1 with probability at least c, whereas in the NO case the measurement outputs 1 with probability at most s, with $c > s$. The standard counting argument placing BQP inside PP then applies to this case as well; see for instance [1, Propositions 2 and 3].

References

[1] Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. *Proceedings of the Royal Society A*, 461(2063):3473–3482, 2005.

[2] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the 35th Annual ACM Symposium on the Theory of Computing (STOC)*, pages 20–29, 2003.

[3] Eric W. Allender and Klaus W. Wagner. Counting hierarchies: polynomial time and constant depth circuits. In G. Rozenberg and A. Salomaa, editors, *Current trends in Theoretical Computer Science*, pages 469–483. World Scientific, 1993.

[4] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 2009.

[5] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh V. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997.

[6] Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18(3):147–150, 1984.

[7] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Exponential improvement in precision for simulating sparse Hamiltonians. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 283–292, New York, NY, USA, 2014. ACM.

[8] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Physical Review Letters*, 114:090502, 2015.

[9] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Proceedings of the 56th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 792–809, 2015.

[10] A Borodin, S Cook, and N Pippenger. Parallel computation for well-endowed rings and space-bounded probabilistic machines. *Information and Control*, 58(1-3):113–136, July 1984.

[11] Sergey Bravyi. Efficient algorithm for a quantum analogue of 2-sat. arXiv preprint quant-ph/0602108, 2006.

[12] Andrew Childs. On the relationship between continuous- and discrete-time quantum walk. *Communications in Mathematical Physics*, 294:581–603, 2010.

[13] Stephen A. Cook. A taxonomy of problems with fast parallel algorithms. *Information and Control*, 64(1-3):2 – 22, 1985. International Conference on Foundations of Computation Theory.

[14] Christopher M. Dawson, Andrew P. Hines, Duncan Mortimer, Henry L. Haselgrove, Michael A. Nielsen, and Tobias Osborne. Quantum computing and polynomial equations over the finite field Z2. *Quantum Information & Computation*, 5(2):102–112, 2005.

[15] Dean Doron, Amir Sarid, and Amnon Ta-Shma. On approximating the eigenvalues of stochastic matrices in probabilistic logspace. Electronic Colloquium on Computational Complexity (ECCC) preprint TR16-120, 2016.

[16] Dean Doron and Amnon Ta-Shma. On the problem of approximating the eigenvalues of undirected graphs in probabilistic logspace. In *Proceedings of the 42nd International Colloquium on Automata, Languages and Programming (ICALP)*, pages 419–431, 2015.

[17] Bill Fefferman, Hirotada Kobayashi, Cedric Yen-Yu Lin, Tomoyuki Morimae, and Harumichi Nishimura. Space-efficient error reduction for unitary quantum computations. In *Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP)*, pages 14:1–14:14, 2016.

[18] François Le Gall. Solving Laplacian systems in logarithmic space. arXiv preprint 1608.01426, 2016.

[19] David Gosset and Daniel Nagaj. Quantum 3-SAT is QMA1-complete. In *Proceedings of the 54th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 756–765, 2013.

[20] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical Review letters*, 103(15):150502, 2009.

[21] Tsuyoshi Ito, Hirotada Kobayashi, and John Watrous. Quantum interactive proofs with weak error bounds. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)*, pages 266–275, 2012.

[22] Richard Jozsa, Barbara Kraus, Akimasa Miyake, and John Watrous. Matchgate and space-bounded quantum computations are equivalent. *Proceedings of the Royal Society A*, 466(2115):809–830, 2010.

[23] Richard Jozsa and Akimasa Miyake. Matchgates and classical simulation of quantum circuits. *Proceedings of the Royal Society A*, 464(2100):3089–3106, 2008.

[24] Julia Kempe and Oded Regev. 3-local Hamiltonian is QMA-complete. *Quantum Information & Computation*, 3(3):258–264, 2003.

[25] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, Boston, MA, USA, 2002.

[26] Chris Marriott and John Watrous. Quantum Arthur-Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

[27] Daniel Nagaj, Pawel Wocjan, and Yong Zhang. Fast amplification of QMA. *Quantum Information & Computation*, 9(11):1053–1068, 2011.

[28] M. A. Nielsen and I. L. Chuang. *Quantum Information and Computation*. Cambridge University Press, Cambridge, UK, 2000.

[29] John H. Reif. Logarithmic depth circuits for algebraic functions. *SIAM Journal on Computing*, 15(1):231–242, 1986.

[30] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac. Computational complexity of projected entangled pair states. *Physical Review Letters*, 98:140506, 2007.

[31] Yaoyun Shi. Both Toffoli and controlled-NOT need little help to do universal quantum computing. *Quantum Information & Computation*, 3(1):84–92, January 2003.

[32] Amnon Ta-Shma. Inverting well conditioned matrices in quantum logspace. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 881–890. ACM, 2013.

[33] Barbara M. Terhal and David P. DiVincenzo. Classical simulation of noninteracting-fermion quantum circuits. *Physical Review A*, 65(3):032325, 2002.

[34] Leslie G. Valiant. Quantum circuits that can be simulated classically in polynomial time. *SIAM Journal on Computing*, 31(4):1229–1254, 2002.

[35] Dieter van Melkebeek and Thomas Watson. Time-space efficient simulations of quantum computations. *Theory of Computing*, 8:1–51, 2012.

[36] F. Verstraete and J. I. Cirac. Renormalization algorithms for quantum-many body systems in two and higher dimensions. arXiv preprint cond-mat/0407066, 2004.

[37] John Watrous. Space-bounded quantum complexity. *Journal of Computer and System Sciences*, 59(2):281–326, 1999.

[38] John Watrous. On the complexity of simulating space-bounded quantum computations. *Computational Complexity*, 12(1):48–84, 2003.

[39] John Watrous. Quantum computational complexity. In Robert A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 7174–7201. Springer, 2009.