Constant-gap adiabatic quantum computation and the QMA promise gap

change this?

Steven T. Flammia* Daniel Gottesman* Spyridon Michalakis†

April 25, 2010

Abstract

Spiros did most of the brilliant work.

1 Introduction

Of the many proposals for constructing a quantum computer, perhaps none is more controversial than the adiabatic model. In its most basic instantiation, a computation in this model begins by preparing a quantum system consisting of n subsystems in a unique ground state of a simple non-interacting potential, such as the ground state $|0\rangle^{\otimes n}$ of the Hamiltonian $H_0 = -\sum_j \sigma_z^{(j)}$. Then one slowly turns on an auxiliary Hamiltonian H_1 which encodes in its ground state the solution to the problem of interest while simultaneously turning off the original Hamiltonian. That is, the time-dependent Hamiltonian for the system is

$$H(s) = (1-s)H_0 + sH_1\,, \tag{1}$$

where $s = t/\tau$ is a scaled time with $0 \leq s \leq 1$ and τ is the maximum amount of time for which the system evolves. The system is measured at $s = 1$ to read out the answer to the computation. Then as long as the conditions of the adiabatic theorem are met—in particular that the total amount of time τ is large enough compared to the inverse of the minimum energy gap Δ between the ground state and the first excited state—the probability of error should be small and the computation succeeds with high probability.

The evolution in Eq. (1), known as adiabatic quantum optimization, was proposed by Farhi *et al.* [1, 2] as an alternative to the circuit model that could take advantage of the natural robustness of quantum ground states to errors and decoherence. Independently, Mizel *et al.* had previously developed a model based on a similar principle [3, 4, 5], but where the solution was prepared by cooling to the ground state, rather than by adiabatic evolution. This model, called ground state quantum computing, had the advantage of being universal, but it was unclear how to cool to the ground state efficiently; however, the couplings in this model can be turned on sequentially to recover an adiabatic-type algorithm. Subsequent work [6, 7, 8] showed that adiabatic quantum

*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5 Canada

†T-4 and Center for Nonlinear Studies, Los Alamos National Labs, Los Alamos, New Mexico, 87544 USA

computation was equivalent to the circuit model if τ is allowed to scale as an inverse polynomial in Δ. Thus, in addition to the issues of error correction [9] and fault tolerance [10], answering the question of how exactly this gap Δ scales with the number of subsystems n is crucial to understanding the efficiency of adiabatic quantum computers.

The universal adiabatic algorithm proposed in Ref. [6] has a dependence of $\Delta \sim \text{poly}(1/n)$ in the worst case [6, 7, 8]. By contrast, the evolution in Eq. (1) applied to random instances of certain NP-complete problems encoded into H_1 seems to have an exponentially small gap on average [11]. This demonstrates that the choice of adiabatic path can greatly affect the scaling of the gap. It remained open whether the worst-case scaling $\Delta \sim \text{poly}(1/n)$ was *necessary* for a universal adiabatic quantum computer. In particular, the possibility of an adiabatic algorithm with a *constant* gap could not be ruled out despite certain theoretical obstacles [12, 13].

Recently, Mizel proposed a model claiming exactly this: a quantum adiabatic algorithm with a constant gap, independent of the system size [14]. His evidence for the gap scaling was partly analytic and partly numeric; the problem of rigorously proving (or disproving) the constant scaling of the gap was left open.

In this paper, we prove that indeed the gap scaling in the Mizel model is lower-bounded by a constant independent of n. ...explain what we do...

Promise gap for QMA. Violation of area law in 2D? Probably not, but worth looking at.

2 The Mizel model

We begin by discussing the one-dimensional version of the Mizel model which performs single-qubit unitary operations. Later we will add two-qubit gates to make the model universal. In addition to a review, we also provide a few additional insights which don't appear in Ref. [14].

The 1D Mizel model is defined on a spin chain consisting of an alternating sequence of 5- and 3-level quantum systems with a distinguished 2-level system on the right-hand boundary of the chain. Thus, the Hilbert space is given by

$$\mathcal{H} = \mathbb{C}^5 \otimes \mathbb{C}^3 \otimes \mathbb{C}^5 \otimes \mathbb{C}^3 \otimes \ldots \otimes \mathbb{C}^5 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2 \,. \tag{2}$$

One can easily write down related models that have a block structure in the bulk which repeats as $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$ instead (again with an additional $\mathbb{C}^2$ on the right-hand boundary), and this may have some advantage with respect to numerical calculations. These models differ slightly in that they have three-body terms.

The computation proceeds by simulating gate teleportation [15]. To accomplish this, we will divide the spins according to two distinct functions. The 5-level systems are essentially program registers which each store the coefficients of a single-qubit unitary operator at a given time step in the quantum circuit, as well as a clock that keeps track of whether or not the gate has been applied. The three-level systems will provide an entanglement resource similar to a Bell pair which will allow teleportation of the gates along the chain. The distinguished qubit on the right-hand boundary is a logical qubit which will store the answer to the computation, and it functions exactly like the 5-level systems, but it doesn't need to store any additional gates. The model simulates the evolution of a quantum circuit with time flowing from left to right along the spin chain and placing the distinguished qubit on the right-hand side merely sets the direction of the flow of time.

Let us examine some general features of the model. The total Hamiltonian will be a sum of projection operators with a unique zero-energy ground state. The gap to the first excited state

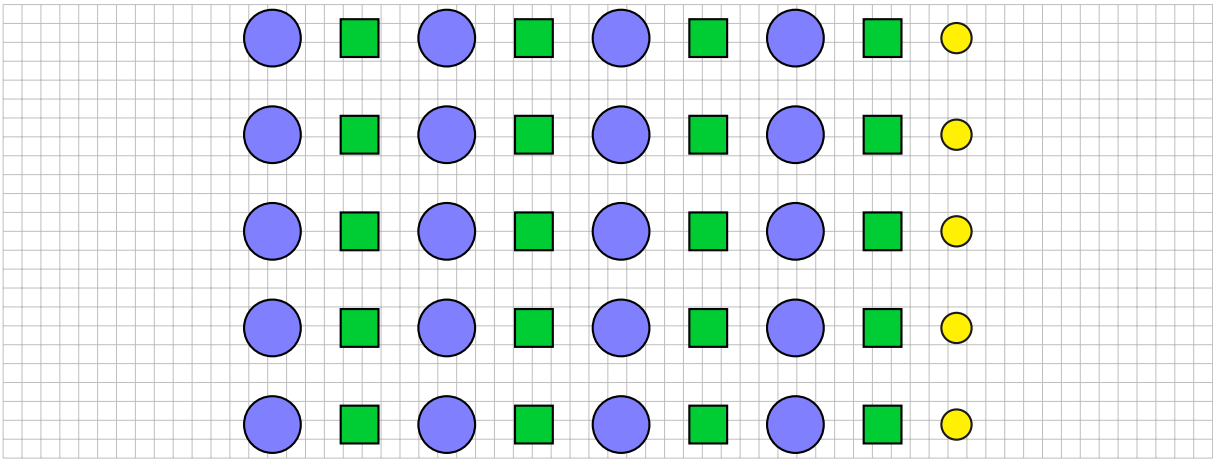

Figure 1: Notation for describing the systems along the spin chain in the 1-d Mizel model. This is obviously just a placeholder for now.

depends on a parameter Λ (to be defined shortly), but it is independent of the choice of single-qubit unitary operators. The parameter Λ determines the success probability for teleporting a gate. When $\Lambda = 0$, the ground state is essentially a product state: it is a tensor product of Bell pairs and unentangled states. For every finite non-zero value of Λ, the probability of successfully teleporting any particular gate is less than one, and the probability of teleporting a fiducial gate across the entire chain is exponentially small in the length of the chain. What makes the model interesting is that for sufficiently large values of Λ the teleportation failure probability is below the fault-tolerance threshold. Thus (when we go to two dimensions) we can employ error correction to achieve a universal fault-tolerant scheme.

Before introducing the Hamiltonian couplings which will enable us to encode the circuit into the ground state, we need a good notation to describe the various systems along the chain. The first 5-level system is block 0, and contains only the system q_0. Then counting the j^{th} pair of spins after that (grouped as $\mathbb{C}^3 \otimes \mathbb{C}^5$) defines the j^{th} block, with subsystems called b_j (the 3-level system) and q_j (the 5-level system). The final block is the n^{th} block, consisting of only b_n and q_n, the final two-level system. This notation is illustrated in Fig. 1.

We label the basis states for each system q_j as

$$|0_\circ\rangle \ , \ |1_\circ\rangle \ , \ |0_\bullet\rangle \ , \ |1_\bullet\rangle \ , \ |I\rangle \, . \tag{3}$$

with the exception that the qubit q_n retains only the states $|0_\circ\rangle$ and $|1_\circ\rangle$. These states label a single logical qubit together with a subscript "clock" label which tells whether or not a unitary gate has been applied. The final state $|I\rangle$ is a non-computational state (I stands for "idle") which heralds a successful gate teleportation.

We label the basis for each b_j system as

$$|0\rangle \ , |1\rangle \ , |I\rangle \, . \tag{4}$$

The first two states serve as a qubit for preparing a singlet state which can teleport a gate, while the last state is again a non-computational basis state heralding when the teleportation succeeds.

Now we can finally discuss the coupling terms in the Hamiltonian. Each q_j system is coupled with an on-site term $H_U = [H_U]_{q_j}$ which encodes the unitary operator U. The choice of U can depend on the index j, but we suppress this in our notation. (Here the notation $[X]_y$ means that operator X acts on system y and as the identity on every other system. We will suppress these labels whenever there is no risk of confusion.) Explicitly, this operator can be written in the basis listed in Eq. (3) as the matrix check the daggers are in the right place

$$H_U = \frac{1}{2}\begin{pmatrix} I & -U^\dagger \\ -U & I \end{pmatrix}, \tag{5}$$

where we have restricted the action to just the subspace spanned by the logical and clock degrees of freedom. The idle state $|I\rangle$ is annihilated by H_U.

On the left-hand boundary of the chain, we need an additional term to initialize the computation so that H_U can prepare a proper history state. This term is simply

$$H_{\text{in}} = [H_{\text{in}}]_{q_0} = |1_\circ\rangle\langle 1_\circ| . \tag{6}$$

Next, there is a term which creates a singlet state, $H_\Phi = [H_\Phi]_{b_j,q_j}$. Specifically, in the local block containing b_j and q_j, let

$$|\phi^-\rangle_{b_j,q_j} = \frac{1}{\sqrt{2}}(|0\rangle_{b_j}|1_\circ\rangle_{q_j} - |1\rangle_{b_j}|0_\circ\rangle_{q_j}) . \tag{7}$$

Now consider the 4-dimensional subspace of the joint b_j-q_j space spanned by the states $\{|0\rangle|0_\circ\rangle$, $|1\rangle|0_\circ\rangle$, $|0\rangle|1_\circ\rangle$, $|1\rangle|1_\circ\rangle\}$, and let 1_Φ be the projector onto that space. Then

$$H_\Phi = 1_\Phi - |\phi^-\rangle\langle\phi^-| . \tag{8}$$

Thus, H_Φ penalizes the three local Bell states other than the singlet, and annihilates the singlet and all the states not in the span of 1_Φ.

The final term we need will teleport the gate U. It decomposes into a sum of two orthogonal projectors, and acts on the systems q_j and b_{j+1}. To define the first and most important piece, we define the vector

$$|\Lambda\rangle_{q_j,b_{j+1}} = \frac{1}{\sqrt{1+\Lambda^2}}\left(\Lambda\frac{|1_\bullet\rangle_{q_j}|0\rangle_{b_{j+1}} - |0_\bullet\rangle_{q_j}|1\rangle_{b_{j+1}}}{\sqrt{2}} + |I\rangle_{q_j}|I\rangle_{b_{j+1}}\right) . \tag{9}$$

The other piece is

$$\big[H_I\big]_{q_j,b_{j+1}} = \big[|I\rangle\langle I|\big]_{q_j} \otimes \big[1 - |I\rangle\langle I|\big]_{b_{j+1}} + \big[1 - |I\rangle\langle I|\big]_{q_j} \otimes \big[|I\rangle\langle I|\big]_{b_{j+1}} . \tag{10}$$

Then the total term is

$$H_\Lambda = |\Lambda\rangle\langle\Lambda| + H_I . \tag{11}$$

To see how this teleports a gate, let's restrict our attention to the smallest interesting 1-dimensional system, consisting of just three spins labeled q_0, b_1 and q_1. The total Hamiltonian acts on $\mathbb{C}^5 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2$ and is given by

$$\mathcal{H}(\Lambda) = H_{\text{in}} + H_U + H_\Lambda + H_\Phi , \tag{12}$$

where we have suppressed the dependence of $\mathcal{H}$ on the gate U.

continue.... The ground state of $H_{\text{in}} + H_U$ is given by,

$$|\Psi_U\rangle = \frac{1}{\sqrt{2}}(|0_\circ\rangle + U_{00}|0_\bullet\rangle + U_{10}|1_\bullet\rangle)\,. \tag{13}$$

(The other ground state is the state $|I\rangle$.)

When the terms H_Λ and H_Φ are added, the ground state of the block Hamiltonian becomes (unnormalized)

$$|g.s.\rangle = |\Psi_U\rangle|\Phi\rangle + \frac{\Lambda}{2\sqrt{2}}|I\rangle|I\rangle|U\rangle\,. \tag{14}$$

2.1 Some intuition about the form of the ground state.

The ground state per block is defined as the weighted sum of two terms: $|\Psi_U\rangle|\Phi\rangle$ and $|I\rangle|I\rangle|U\rangle$. When we act on the first term with the state $\langle\Lambda_1|_{Q_0,B_1} = \Lambda\langle\Phi^-|_{Q_0,B_1} + \langle S|_{Q_0} \otimes \langle F|_{B_1}$, we effectively teleport the state $u_{0,0}|0_1\rangle + u_{1,0}|1_1\rangle$, with probability $(1/2)*(1/4)$ from Q_0 to Q_1. This follows from $|\Psi^0_{U_0}\rangle_{Q_0}$ being an equal superposition of $|0_0\rangle$ and $u_{0,0}|0_1\rangle + u_{1,0}|1_1\rangle$ and the choice of the (one-out-of-four, antisymmetric) Bell state measurement $|\Phi^-\rangle\langle\Phi^-|_{Q_0,B_1}$ that acts only on the stage 1 states $|0_1\rangle, |1_1\rangle$. Finally, the teleportation is done with the aid of another singlet $|\Phi^-\rangle_{B_1,Q_1}$, leading to an overall phase of -1. This state exactly cancels the computational state $|U_0\rangle_{Q_1}$, which remains after we act with $\langle\Lambda_1|_{Q_0,B_1}$ on $|S\rangle_{Q_0} \otimes |F\rangle_{B_1} \otimes |U_0\rangle_{Q_1}$.

Finish introducing model. Mention the 2d case.

3 Proof that the gap is constant

In this section we prove that the gap between the ground state and the first excited state in the Mizel model is bounded from below by a constant that depends only on Λ, and not on the circuit or the system size. We emphasize that *whenever* we say "constant" we mean independent of the system size n and the circuit, but we allow constants to depend on Λ.

The proof proceeds in three steps. First we show that the model can be reduced to a different model that is slightly better behaved (it is a sum of commuting projectors) and prove that the gap of this reduced model bounds the gap of the original model. We specifically analyze this model in the trivial case where $\Lambda = 0$. Next, we show that for small values of Λ the gap at $\Lambda = 0$ persists in a neighborhood around the origin. Specifically, we show that for some sufficiently small constant ϵ, the gap is bounded from below by a constant for all $\Lambda \in [0, \epsilon)$. Finally, we derive a differential equation whose solution gives a lower bound on the decay of the gap as a function of Λ. We prove that this decay holds for all $\Lambda > 0$, but it contains an essential singularity at the origin. By enforcing the boundary condition of this differential equation in the interval $(0, \epsilon)$ instead of at the origin, we derive a constant lower bound for all Λ.

3.1 The reduced model and $\Lambda = 0$

The first ingredient we need is a lemma that will help us write the Mizel model as a sum of commuting projectors when $\Lambda = 0$. Many variants of this lemma have been proven in the literature,

and in many different contexts; the essential insights go back at least to Halmos [16]. We first define the *angle* between two distinct projectors.

Definition 1 (Angle)**.** *Let A and B be distinct projection operators acting on $\mathbb{C}^d$. Define the* **angle** *$\gamma = \gamma(A, B)$ to be the largest eigenvalue of $(ABA)^{1/2}$ which is strictly less than 1.*

We call this quantity an angle, but this is a slight abuse of terminology. It is actually the cosine of an angle in the interval $[0, \pi/2]$, since it always lies in the interval $[0, 1]$. This quantity will play a key role in the analysis of the Mizel model for small values of Λ. Note that the eigenvalues of ABA are the same as the eigenvalues of BAB, so this definition is actually symmetric, despite appearances to the contrary.

Lemma 2. *Let A and B be distinct projection operators acting on $\mathbb{C}^d$ with angle γ, and let $P_{A,B}$ be the projector onto the support of $A + B$. Then*

$$A + B \geq \big(1 - \gamma\big) P_{A,B}\,, \tag{15}$$

where the inequality holds in the positive semidefinite ordering.

Proof. In the subspace where A and B commute, the bound is trivially true. In the subspace where they don't commute, we use the following well known simultaneous canonical form for two projectors:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} C^2 & CS \\ CS & S^2 \end{pmatrix}, \tag{16}$$

where the matrices in each block are diagonal and $\ker C = \ker S = 0$ and $C^2 + S^2 = 1$. Thus, the problem of computing a bound on $A + B$ reduces to a direct sum of 2×2 blocks. Each block can be easily solved exactly. The eigenvalues or $A + B$ are given by $1 \pm C$ within that block. Since the angle $\gamma = \|C\|$ this indeed gives us the desired bound. □

We will apply this lemma to the pairs of terms in the Mizel Hamiltonian that don't commute even when $\Lambda = 0$. These pairs of terms are (H_{in}, H_U), (H_Φ, H_U) and (H_Φ, H_{CP}) whenever these terms overlap at a sight, of course. There is a subtlety for the pair (H_Φ, H_{CP}), since actually an H_{CP} term intersects *two* different H_Φ terms, once on the top wire (t) and once on the bottom wire (b). But this is easily fixed by the chain of inequalities

$$H_\Phi^{\text{t}} + H_\Phi^{\text{b}} + H_{\text{CP}} \geq P_\Phi^{\text{t,b}} + H_{\text{CP}} \geq (1 - \gamma) P_{\Phi,\text{CP}}\,, \tag{17}$$

where

$$P_\Phi^{\text{t,b}} = H_\Phi^{\text{t}} + H_\Phi^{\text{b}} - H_\Phi^{\text{t}} H_\Phi^{\text{b}} \tag{18}$$

is the projector onto the support of $H_\Phi^{\text{t}} + H_\Phi^{\text{b}}$, and the last step follows from Lemma 2. For each of these pairs, we can compute the angle, and we find the following result.

Lemma 3. *For each of the pairs of projectors (H_{in}, H_U), (H_Φ, H_U) and $(P_\Phi^{t,b}, H_{CP})$, the angle γ is the same for all pairs and is given by*

$$\gamma = 1/\sqrt{2}\,. \tag{19}$$

Proof. The proof is simply a series of straightforward and tedious calculations. □

Now that we have computed the angles, we can use Lemma 2 to write a new Hamiltonian with nicer properties. Let $\mathcal{K}(\Lambda)$ be the Hamiltonian defined as follows. For every pair of non-commuting projectors (A, B) from Lemma 3 that appears in the original Mizel model, replace them with $P_{A,B}$, the projector onto the support of $A + B$. We call $\mathcal{K}$ the "reduced" Hamiltonian. Then we have the main result of this section.

Theorem 4. *The reduced Hamiltonian $\mathcal{K}(\Lambda)$ defined above satisfies*

$$\mathcal{H}(\Lambda) \geq \left(1 - \frac{1}{\sqrt{2}}\right) \mathcal{K}(\Lambda) \tag{20}$$

for all values of $\Lambda \geq 0$. In particular, let $\Delta(\Lambda)$ be the gap of $\mathcal{K}(\Lambda)$. Then $(1 - 1/\sqrt{2})\Delta(\Lambda)$ gives a lower bound on the gap of $\mathcal{H}(\Lambda)$ for all Λ which is tight at $\Lambda = 0$. Furthermore, $\mathcal{K}(\Lambda)$ has the same unique frustration-free ground state for all values of Λ that $\mathcal{H}(\Lambda)$ has. Finally, when $\Lambda = 0$, $\mathcal{K}(0)$ is a sum of commuting projectors with a factorized ground state.

Proof. The claim that $\mathcal{H}(\Lambda) \geq (1 - 1/\sqrt{2})\mathcal{K}(\Lambda)$ follows immediately from the previous lemmas. One simply observes that the inequality holds term-by-term. The gap $\Delta(\Lambda)$ of $\mathcal{K}(\Lambda)$ provides a tight lower bound on the gap of $\mathcal{H}(\Lambda)$ since the inequality is tight at $\Lambda = 0$.

The fact that $\mathcal{K}(\Lambda)$ has for all Λ the same unique frustration-free ground state as $\mathcal{H}(\Lambda)$ follows since the null space of each is the same term-by-term. Only the excited states change.

It is straightforward to check that $\mathcal{K}(0)$ is commuting. The factorized ground state consists of a tensor product over Bell pairs that have gates applied to them, either a local U gate or a CPHASE gate across two Bell pairs. □

3.2 Small values of Λ

Now that we have the reduced model $\mathcal{K}(\Lambda)$, we can derive a lower bound on the gap $\Delta(\Lambda)$ which holds for small Λ. Our proof strategy follows Fannes, Nachtergaele and Werner [17], from which we can infer the following more general theorem:

Theorem 5 (Generalized FNW)**.** *Let $H = \sum_j P_j$ be a Hamiltonian which is a sum of projectors P_j with a unique frustration-free ground state. For each pair of non-commuting projectors (P_j, P_k), compute their angle $\gamma_{j,k}$ and let $g = \max_{j,k} \gamma_{j,k}$. Then the gap Δ of H satisfies*

$$\Delta \geq 1 - 2g\,, \tag{21}$$

which gives a non-trivial bound whenever $g < 1/2$.

Proof. The proof in Ref. [17] is unnecessarily complicated for our purposes since it uses the language of C*-algebras to manage certain infinite limits, and it is specialized to the translation invariant case. We give a simple proof here following Ref. [18].

We begin with the insight that the gap Δ is the largest constant such that $H^2 \geq \Delta H$. Computing the square of H and ignoring some positive terms, we find

$$H^2 = \sum_j P_j^2 + \sum_{j \neq k} P_j P_k \geq H + \sum_{j \not\sim k} \{P_j, P_k\}\,, \tag{22}$$

where the notation $j \not\sim k$ means that we restrict the sum to only those terms where P_j and P_k don't commute.

Now we need to bound how negative the anticommutator terms can be. Using again the canonical form from Lemma 2, we find that

$$\{P_j, P_k\} \geq -\gamma_{j,k}(P_j + P_k)\,. \tag{23}$$

Putting this inequality back into Eq. (22) and performing the sum, we find

$$H^2 \geq H - \sum_{j \not\sim k} \gamma_{j,k}(P_j + P_k) \geq H - 2gH\,. \tag{24}$$

Thus, we have that $\Delta \geq 1 - 2g$, as claimed. □

Looking back, we now see why the projections to the reduced model were necessary. The value of the worst-case angle g in the original Mizel model is too large to get a non-trivial bound by this argument. However, with the reduced model, we have the following theorem.

Theorem 6 (Gap for small Λ)**.** *The reduced Hamiltonian $\mathcal{K}(\Lambda)$ satisfies the conditions of Theorem 5 with a constant*

$$g(\Lambda) = \frac{\Lambda}{\sqrt{1+\Lambda^2}}\,, \tag{25}$$

which means that $\Delta(\Lambda) > 0$ for all $0 \leq \Lambda < 1/\sqrt{3}$.

Proof. The proof is once again the tedious exercise of computing all of the angles γ for each of the types of non-commuting terms in $\mathcal{K}(\Lambda)$. It is actually not trivial to do this, since some of the terms are quite large matrices. We used *Mathematica* to symbolically compute each of the angles γ and the largest one was the $g(\Lambda)$ listed above. □

3.3 Large values of Λ

To obtain a lower bound on the gap for large values of Λ, we follow Mizel's idea of using the Hellman-Feynman theorem. This theorem states the following. If $E_k(\Lambda)$ is the kth eigenvalue of a Hamiltonian $H(\Lambda)$ with corresponding normalized eigenvector $|\Psi_k(\Lambda)\rangle$ (all differentiable with respect to Λ), then

$$\partial_\Lambda E_k(\Lambda) = \langle \Psi_k(\Lambda) | \partial_\Lambda H(\Lambda) | \Psi_k(\Lambda) \rangle\,. \tag{26}$$

I would like to review the differentiability and degeneracy conditions on this theorem. We would like to find a lower bound of the form $\partial_\Lambda E_k(\Lambda) \geq -f(\Lambda)E_k(\Lambda)$ for some $f(\Lambda) \geq 0$, since this will imply decay with a constant rate independent of the size of the system n. This is exactly the content of the next theorem.

Theorem 7 (Decay rate)**.** *For any eigenstate $|\Psi_k(\Lambda)\rangle$ of $\mathcal{H}(\Lambda)$ or $\mathcal{K}(\Lambda)$ with corresponding eigenvalue $E_k(\Lambda)$ and for all $\Lambda > 0$, the following inequality holds:*

$$\partial_\Lambda E_k(\Lambda) \geq -\frac{f(\Lambda)}{1+\Lambda^2} E_k(\Lambda)\,, \tag{27}$$

with

$$f(\Lambda) = (1+\sqrt{2})\Lambda + \left[(1+\sqrt{2})\Lambda\right]^{-1}. \tag{28}$$

Proof. Consider a single term $|\Lambda\rangle\langle\Lambda|_j$ as defined in Eq. (9) acting on an arbitrary pair of spins labeled by j. We will first show that it is sufficient to prove the inequality

$$\langle \partial_\Lambda |\Lambda\rangle\langle\Lambda|_j\rangle \geq -\frac{f(\Lambda)}{1+\Lambda^2}\,\langle |\Lambda\rangle\langle\Lambda|_j\rangle\ , \tag{29}$$

where the expectation is taken with respect to $|\Psi_k(\Lambda)\rangle$. If this holds for an arbitrary such term, then we can sum both sides of this inequality over all such terms,

$$\sum_j \langle \partial_\Lambda |\Lambda\rangle\langle\Lambda|_j\rangle \geq -\frac{f(\Lambda)}{1+\Lambda^2}\sum_j \langle |\Lambda\rangle\langle\Lambda|_j\rangle\ . \tag{30}$$

By the Hellman-Feynman theorem, the left-hand side is exactly $\partial_\Lambda E_k(\Lambda)$. For the right-hand side, we can always subtract the rest of the terms in the Hamiltonian to make this side more negative, since every term is positive semidefinite. Therefore Eq. (29) implies Eq. (27), as claimed.

Let's look at the local states which comprise the projector $|\Lambda\rangle\langle\Lambda|$. Using a shorthand $|\Phi\rangle$ and $|I\rangle$ for the two orthogonal terms in Eq. (9), we can rewrite

$$|\Lambda\rangle = \frac{1}{\sqrt{1+\Lambda^2}}\big(\Lambda|\Phi\rangle + |I\rangle\big)\,. \tag{31}$$

Thus the inequality in Eq. (29) only depends on the *matrix minor* in the $\{|\Phi\rangle, |I\rangle\}$ basis of the density operator $\rho_k(\Lambda)$.

To take advantage of this insight, we can explicitly compute the derivative in Eq. (29) to get

$$\partial_\Lambda(|\Lambda\rangle\langle\Lambda|) = \frac{1}{(1+\Lambda^2)^2}\big[2\Lambda(|\Phi\rangle\langle\Phi| - |I\rangle\langle I|) + (1-\Lambda^2)(|\Phi\rangle\langle I| + |I\rangle\langle\Phi|)\big] \tag{32}$$

We can collect the terms on the other side and define the matrix

$$M(\Lambda) = \begin{pmatrix} f(\Lambda) - 2\Lambda & \Lambda f(\Lambda) + 1 - \Lambda^2 \\ \Lambda f(\Lambda) + 1 - \Lambda^2 & \Lambda^2 f(\Lambda) + 2\Lambda \end{pmatrix}\ , \tag{33}$$

written in the $\{|\Phi\rangle, |I\rangle\}$ basis such that Eq. (29) becomes simply

$$\langle M(\Lambda)\rangle \geq 0\,. \tag{34}$$

Our goal is to find a function $f(\Lambda) \geq 0$ such that this inequality is true. Now we use the fact that the matrix minor of any density operator is again a positive operator, but with a potentially sub-normalized trace [19]. Therefore if we choose the worst case operator consistent with this information we find

$$\langle M(\Lambda)\rangle \geq \min_{\substack{\rho\geq 0 \\ \mathrm{Tr}(\rho)\leq 1}} \mathrm{Tr}\big[M(\Lambda)\rho\big]\,. \tag{35}$$

The set of positive operators with trace ≤ 1 is just the Bloch sphere, so the optimization is straightforward.

Since $M(\Lambda)$ is real and symmetric, we only need to optimize over the real part. We can parameterize this cone as $\rho = \tau 1 + r\cos(\theta)X + r\sin(\theta)Z$ with trace and positivity constraints becoming $r \leq \tau \leq 1/2$. The minimization then becomes

$$\min_{0\leq r\leq\tau\leq 1/2} r\sin(\theta)(-f(\Lambda)\Lambda^2 + f(\Lambda) - 4\Lambda) + 2r\cos(\theta)(\Lambda(f(\Lambda)-\Lambda)+1) + \tau(\Lambda^2+1) \tag{36}$$

This proof is wrong. □

4 Conclusion

I think this section works better in the conclusion. Needs a revision to put it in this context, though. We also give an alternative proof of the second step of our argument. We show that the modified Mizel model at $\Lambda = 0$ satisfies the two conditions required for the proof of stability given recently for topologically ordered phases [20, 21]. Then all of the Λ-dependent terms in the Hamiltonian can be treated as perturbations in that framework, and the proof of the constant gap follows immediately for a neighborhood of constant size about the origin. This doesn't quite allow us to give a full independent proof however, since the best lower bound we can achieve with these methods decreases linearly with Λ. Thus, the lower bound becomes negative at some finite value of Λ, which could be less than the value required for fault tolerance. An additional drawback is that the bounds used in Refs. [20, 21] don't have explicit constants and hence they don't allow us to check the bound without additional work.

It is worth briefly mentioning why an adiabatic quantum algorithm with a constant *gap* does not imply a quantum algorithm with constant *time*. maybe put this in the intro?

Acknowledgments We thank Mick Bremner, Yi-Kai Liu and Ari Mizel for discussions. STF, DG and research at Perimeter Institute are supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation.

References

[1] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum Computation by Adiabatic Evolution," quant-ph/0001106 (2000).

[2] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, "A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem," Science **292**, 472 (2001).

[3] A. Mizel, M. W. Mitchell, and M. L. Cohen, "Energy barrier to decoherence," Phys. Rev. A **63**, 40302 (2001).

[4] A. Mizel, M. W. Mitchell, and M. L. Cohen, "Scaling considerations in ground state quantum computation," Phys. Rev. A **65**, 022315 (2002).

[5] A. Mizel, "Mimicking time evolution within a quantum ground state: Ground-state quantum computation, cloning, and teleportation," Phys. Rev. A **70**, 012304 (2004).

[6] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, "Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation," in *45th Annual Symposium on the Foundations of Computer Science*, pp. 42–51 (IEEE Computer Society, Los Alamitos, CA, USA, 2004).

[7] A. Mizel, D. A. Lidar, and M. Mitchell, "Simple proof of equivalence between adiabatic quantum computation and the circuit model," Phys. Rev. Lett. **99**, 070502 (2007).

[8] P. Deift, M. B. Ruskai, and W. Spitzer, "Improved Gap Estimates for Simulating Quantum Circuits by Adiabatic Evolution," Quant. Info. Process. **6**, 121 (2007).

[9] S. P. Jordan, E. Farhi, and P. W. Shor, "Error-correcting codes for adiabatic quantum computation," Phys. Rev. A **74**, 052322 (2006).

[10] D. A. Lidar, "Towards Fault Tolerant Adiabatic Quantum Computation," Phys. Rev. Lett. **100**, 160506 (2008).

[11] B. Altshuler, H. Krovi, and J. Roland, "Adiabatic quantum optimization fails for random instances of NP-complete problems," arXiv:0908.2782 (2009).

[12] T. J. Osborne, "Simulating adiabatic evolution of gapped spin systems," Phys. Rev. A **75**, 032321 (2007).

[13] M. B. Hastings, "Quantum Adiabatic Computation with a Constant Gap Is Not Useful in One Dimension," Phys. Rev. Lett. **103**, 050502 (2009).

[14] A. Mizel, "Fixed-gap ground-state quantum computation and quantum concurrent processing," arXiv:1002.0846 (2010).

[15] D. Gottesman and I. L. Chuang, "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations," Nature **402**, 390 (1999).

[16] P. Halmos, "Two subspaces," Trans. Amer. Math. Soc. **144**, 381 (1969).

[17] M. Fannes, B. Nachtergaele, and R. F. Werner, "Finitely correlated states on quantum spin chains," Comm. Math. Phys. **144**, 443 (1992).

[18] S. T. Flammia, D. Gross, S. D. Bartlett, and R. Somma, "Heralded Polynomial-Time Quantum State Tomography," arXiv:1002.3839 (2010).

[19] R. Bhatia, *Positive definite matrices* (Princeton University Press, 2006).

[20] S. Bravyi, M. Hastings, and S. Michalakis, "Topological quantum order: stability under local perturbations," arXiv:1001.0344 (2010).

[21] S. Bravyi and M. B. Hastings, "A short proof of stability of topological order under local perturbations," arXiv:1001.4363 (2010).

Old Notes

Our goal is now to build a local gapped Hamiltonian whose (unnormalized) ground state is given by

$$|\Lambda_U\rangle = |\Phi\rangle|\Psi_U\rangle + \frac{\Lambda}{2\sqrt{2}}|U\rangle|B\rangle|H\rangle\,. \tag{37}$$

That is, we want the state to be a superposition of the initial state and a state where the second half of the Bell pair and the original history state are both in their idle configurations and the gate U has been teleported to the first logical qubit associated with the first half of the Bell pair.

4.1 Tedious calculation

Here we explicitly compute how H_Λ acts on $|\Lambda_U\rangle$. Let's start with the assumption that it is indeed a zero-energy ground state. Then we can ignore constant factors in front of either the state or the Hamiltonian.

Define

$$\langle\chi| = \frac{\Lambda}{\sqrt{2}}\langle 0|\langle 1_1| - \frac{\Lambda}{\sqrt{2}}\langle 1|\langle 0_1| - \langle B|\langle H|\,. \tag{38}$$

Then we have that

$$H_\Lambda \propto I_\mathrm{A} \otimes |\chi\rangle\langle\chi|\,. \tag{39}$$

Expanding the definition of $|\Lambda_U\rangle$ (up to a constant factor) we have

$$|\Lambda_U\rangle = |0\rangle|1\rangle|\Psi_U\rangle - |1\rangle|0\rangle|\Psi_U\rangle + \frac{\Lambda}{2}|U\rangle|B\rangle|H\rangle\,. \tag{40}$$

Now we want to compute the quantity $\langle\chi|H_\Lambda|\Lambda_U\rangle$, which is an vector living in A (hopefully the zero vector.) Writing it out explicitly, we find

$$\langle\chi|H_\Lambda|\Lambda_U\rangle = \langle\chi|1,\Psi_U\rangle|0\rangle - \langle\chi|0,\Psi_U\rangle|1\rangle + \frac{\Lambda}{2}\langle\chi|B,H\rangle|U\rangle\,, \tag{41}$$

where recall that $|U\rangle = U_{00}|0\rangle + U_{10}|1\rangle$.

We have to compute the matrix elements $\langle\chi|1,\Psi_U\rangle$, $\langle\chi|0,\Psi_U\rangle$ and $\langle\chi|B,H\rangle$. First computation:

$$\langle\chi|1,\Psi_U\rangle = -\frac{\Lambda}{\sqrt{2}}\langle 0_1|\Psi_U\rangle = -\frac{\Lambda U_{00}}{2}\,. \tag{42}$$

Second computation:

$$\langle\chi|0,\Psi_U\rangle = \frac{\Lambda}{\sqrt{2}}\langle 1_1|\Psi_U\rangle = \frac{\Lambda U_{10}}{2}\,. \tag{43}$$

Final computation:

$$\langle\chi|B,H\rangle = 1\,. \tag{44}$$

Plugging these back into the expression from Eq. (41) and simplifying gives

$$\langle \chi | H_\Lambda | \Lambda_U \rangle = \left(-\frac{\Lambda U_{00}}{2} \right) |0\rangle - \left(\frac{\Lambda U_{10}}{2} \right) |1\rangle + \frac{\Lambda}{2}(-1)|U\rangle \tag{45}$$

$$= \frac{\Lambda}{2}\big(-U_{00}|0\rangle - U_{10}|1\rangle - |U\rangle\big) \tag{46}$$

$$= \frac{\Lambda}{2}\big(-|U\rangle + |U\rangle\big) = 0\,. \tag{47}$$

and everything cancels.

5 Multiple gates on a single qubit

We've shown that the state $|\Lambda_U\rangle$ is indeed a zero-energy ground state of the combined Hamiltonian (with the small caveat there there is a sign error in Mizel's paper.)

What happens when we want to do more than one gate? Let's stick to a single logical qubit for now.

6 Proof That the Gap is Constant

The structure of our proof is as follows:

1. Prove the ground state is unique when $\Lambda = 0$ and the system is gapped with a constant gap.
2. Prove that there is a frustration-free ground state $|\Psi(\Lambda)\rangle$, which is a zero-energy eigenstate of the full Hamiltonian for all values of $\Lambda \geq 0$.
3. Use the Feynman-Hellman theorem to prove that the energy of the first excited state has a rate of change bounded from below.
4. Verify that the computation succeeds with high probability by checking that the ground state really looks like a sum over "faulty" histories, and that measurement will reveal gate errors.

6.1 Uniqueness

When $\Lambda = 0$, the ground state factorizes into a product state. The Hamiltonian is a sum of commuting projectors, except for two terms that don't commute. We just compute the gap there, and we're done. In particular, it can be shown easily that the ground state given below (for $\Lambda = 0$) uniquely satisfies each projector in the Hamiltonian.

6.2 Frustration-Free

This requires that we have a good description of the ground state. The ground state $|\Psi_0^n(\Lambda)\rangle$ of the 1-dimensional n-gate chain is given recursively by:

$$|\Psi_0^n(\Lambda)\rangle = \left(1 - \Lambda|S_{Q_{n-1}} \otimes F_{B_n}\rangle\langle\Phi^-_{Q_{n-1},B_n}|\right) |\Psi_0^{n-1}(\Lambda)\rangle \otimes |\Phi^-_{U_n}\rangle, \tag{48}$$

$$|\Psi_0^0(\Lambda)\rangle = |\Psi^0_{U_0}\rangle, \quad |\Phi^-_{U_n}\rangle = \frac{1}{\sqrt{2}}\left(|0_{B_{n-1}}\rangle \otimes |\Psi^1_{U_n}\rangle - |1_{B_{n-1}}\rangle \otimes |\Psi^0_{U_n}\rangle\right), \tag{49}$$

$$|\Psi^0_{U_n}\rangle = \frac{1}{\sqrt{2}}\left(|0_0\rangle + U_n(0,0)|0_1\rangle + U_n(1,0)|1_1\rangle\right), \tag{50}$$

$$|\Psi^1_{U_n}\rangle = \frac{1}{\sqrt{2}}\left(|1_0\rangle + U_n(0,1)|0_1\rangle + U_n(1,1)|1_1\rangle\right). \tag{51}$$

6.3 Invariance of spectrum

Using (**??**-**??**), it is easy to show that all terms of the Hamiltonian in each block commute with each of the V_j, since the terms affected contain either stage 1 states, or $|S\rangle$ states, all of which act either on the 2^{nd} or 3^{rd} column of V_j), which are trivially commuting as seen from (**??**). The only other term one needs to check is the $[H_{\text{in}}]_{Q_0} = |1_0\rangle\langle 1_0|$ term, which also commutes trivially through the top-left 0-block of V_j. Hence, the global unitary

$$W_n = (V_0)_{Q_0} \otimes \mathbb{1}_{B_1} \otimes (V_1)_{Q_1} \otimes \mathbb{1}_{B_2} \cdots \otimes (V_{n-1})_{Q_{n-1}} \otimes \mathbb{1}_{B_n} \otimes \mathbb{1}_{Q_n}$$

satisfies:

$$H_n(\Lambda; U_0, U_1, \ldots, U_{n-1}) = W_n H_n(\Lambda; \mathbb{1}) W_n^\dagger. \tag{52}$$

6.4 Bound the gap

Here we need to just differentiate the Hamiltonian and put a lower bound on the rate of change of the energy of the excited states. Not so straightforward.

6.5 Fault tolerance

Here we need to check that when we measure the state at the end, the result really collapses to the history of a faulty quantum circuit where each gate fails with some probability p.

7 Promise gap for QMA

Daniel...

8 Violation of area law in 2D?

Could be true or not. Actually, having a PEPS description for this state doesn't mean we can contract it efficiently... so there is no contradiction if the area law still holds. But this state gives as good a shot as any for a violation.