Quantum Computation and Decision Trees

Edward Farhi*
Center for Theoretical Physics
Massachusetts Institute of Technology
Cambridge, MA 02139
farhi@mitlns.mit.edu

Sam Gutmann
Department of Mathematics
Northeastern University
Boston, MA 02115
sgutm@nuhub.neu.edu

Abstract

Many interesting computational problems can be reformulated in terms of decision trees. A natural classical algorithm is to then run a random walk on the tree, starting at the root, to see if the tree contains a node n levels from the root. We devise a quantum mechanical algorithm that evolves a state, initially localized at the root, through the tree. We prove that if the classical strategy succeeds in reaching level n in time polynomial in n, then so does the quantum algorithm. Moreover, we find examples of trees for which the classical algorithm requires time exponential in n, but for which the quantum algorithm succeeds in polynomial time. The examples we have so far, however, could also be solved in polynomial time by different classical algorithms.

MIT-CTP-2651, quant-ph/9706062 June 1997

1 Introduction

Many of the problems of interest to computation experts are, or are reducible to, decision problems. These are problems that for a given input require the determination of a yes or no answer to a specified question about the input. For example the traveling salesman problem is (polynomial time) equivalent to the decision problem that asks whether or not for a given set of intercity distances there is a route passing through all of the cities whose length is less than a given fixed length. Another example that we will later use for concreteness in this paper is the $0-1$ integer programming problem called "exact cover"[1]. Here we are given an m by n matrix, A, all of whose entries are either 0 or 1. The number of columns m is

*This work was supported in part by The Department of Energy under cooperative agreement DE-FC02-94ER40818.

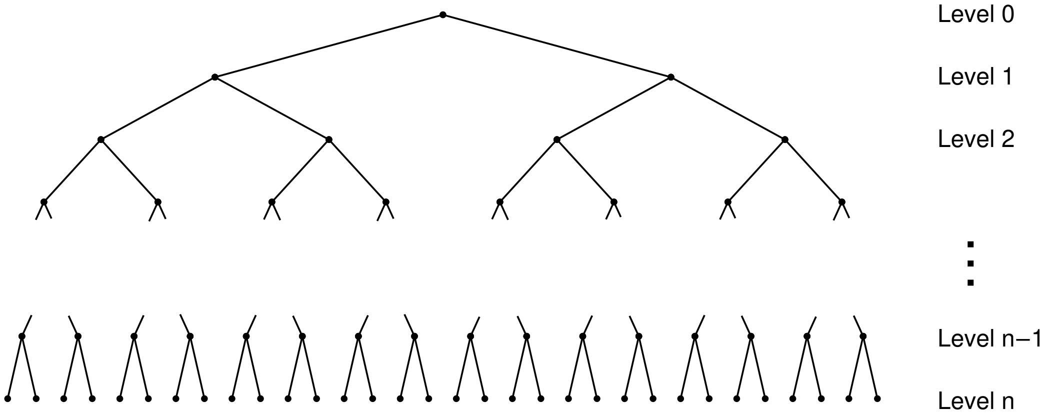

Figure 1: The underlying branching tree. At level m there are 2^m nodes.

$\leq n$. We are asked if there exists a solution to the m equations

$$\sum_{k=1}^{n} A_{jk} x_k = 1 \qquad \text{for } j = 1, m \tag{1.1}$$

with the x_k restricted to be 0 or 1. The brute force approach to this problem is to try the 2^n possible choices of $\vec{x} = (x_1, \ldots, x_n)$. For each choice of $\vec{x}$, checking to see if Eq. (1.1) is satisfied takes at most of order mn operations, which is polynomial in the input size. However, checking all 2^n possible choices for $\vec{x}$ is prohibitively time consuming even for moderately large values of n.

For the exact cover problem, with a given instance of the input matrix A, it is not actually necessary to check all 2^n values of $\vec{x}$ to see if Eq. (1.1) can be satisfied. Note that generically x_1 can take the values 0 or 1 and (x_1, x_2) can have the values $(0,0), (0,1), (1,0)$ or $(1,1)$. However, suppose that for some j the matrix A has $A_{j1} = A_{j2} = 1$. In this case the choice $x_1 = x_2 = 1$ is eliminated and no vectors $\vec{x}$ of the form $(1, 1, x_3, \ldots, x_n)$ need be tried. If we consider vectors $\vec{x}$ that begin with $x_1, x_2, \ldots, x_\ell$ then if for some j we have $\sum_{k=1}^{\ell} A_{jk} x_k \geq 2$, then all vectors $\vec{x}$ beginning with $x_1, x_2, \ldots, x_\ell$ are eliminated. We can picture this in terms of a decision tree as follows. Before imposing any constraints we construct an underlying branching tree. This tree starts at the top with one starting node that branches to two nodes corresponding to the two choices for x_1. This then branches to the four choices for (x_1, x_2) and so on until we have all 2^n choices for $(x_1 \cdots x_n)$ at the n^{th} level. However if we impose the constraints and see that a particular node is eliminated, then we can also eliminate all nodes connected to that node that lie below it in the tree. The decision tree is the underlying branching tree that has been trimmed as a result of the constraints. Note that the exact cover problem has a solution if and only if the decision tree has one or more nodes left at the bottom (n^{th}) level.

More generally we view decision problems as having an underlying bifurcating branching tree with n levels as in Fig. 1. The specific form (or instance) of the problem imposes constraints that eliminates nodes from the tree as in Fig. 2. When a node is excluded the

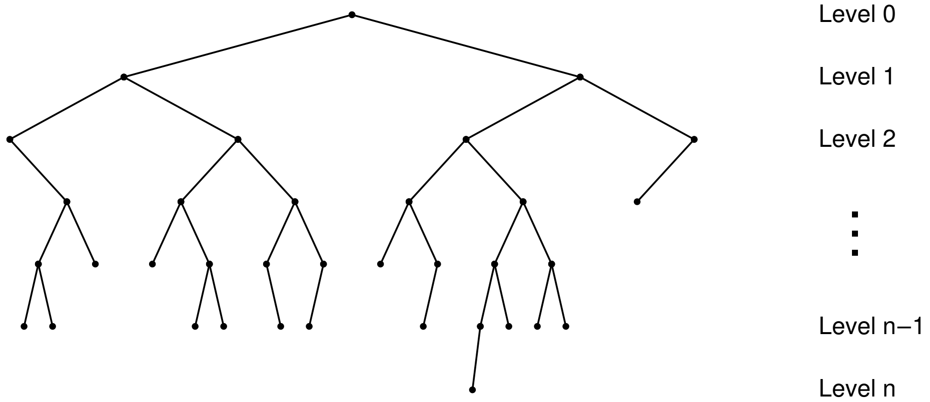

Figure 2: An example of a decision tree, T_n, with one node at level n. For aesthetic reasons we will no longer put breaks in trees – they are still to be thought of as being many levels deep.

whole branch with that node as its topmost point is also cut from the tree. The decision question we wish to answer is "are there any nodes left at the n^{th} level after all constraints have been imposed?"

Consider a family of decision problems indexed by a size n. Particular instances of the problem of size n give rise to particular decision trees that either have or do not have nodes at the n^{th} level. The computational concern is how much time, or how many algorithmic steps, are required to answer the decision question as n gets big. Roughly speaking if the time grows like n^A for fixed $A > 0$, the problem is considered easy whereas if the time grows like a^n with $a > 1$, the problem requires an "exponential amount of time" and is considered computationally hard.

One approach to solving a decision problem is to systematically check every path that starts at the top of the tree and moves downward through the tree. If a path reaches a dead end you try the next path (in some list of paths) until you find a path that has a node at the n^{th} level or else, after having checked all paths, you discover that the answer to the decision question is "no". An alternative to systematically exploring the whole tree is to move through the tree with a probabilistic rule. For example you could use the rule that if you are at a given node you move to the other nodes that are connected to it with equal probability. Thus if you are at a node that connects to two nodes below it you have a 1/3 chance of moving back up the tree; if the node connects to just one below you have a 1/2 chance of moving back up whereas if the node is a dead end you definitely move back. If you start at the top of the tree and move with this probabilistic rule you will eventually visit every node in the tree.

Consider a family of decision trees that are associated with underlying branching trees that are n levels deep. An individual instance of the decision tree either has or does not have nodes at the n^{th} level. If it does have nodes at the n^{th} level and we use a probabilistic rule for moving through the tree, then we say that the tree is penetrable if there is a good chance of reaching the n^{th} level in not too great a time. More precisely we define the family of trees as penetrable if:

> For those trees with a node (or nodes) at the n^{th} level there exist $A, B > 0$ such that for some $t < n^A$ the probability of being at the n^{th} level by t is greater than $(1/n)^B$. (P)

This means that in polynomial time the probability of reaching the n^{th} level is at worst polynomially small. If (P) is met, then by running the process order n^B times we achieve a probability of order 1 of reaching the n^{th} level in time n^{A+B}. If (P) is not met it means that it either takes more than polynomial time to reach the n^{th} level or the probability of reaching the n^{th} level is always smaller than $(\frac{1}{n})$ to any power. Therefore if condition (P) is not met instances of the trees with nodes at the n^{th} level cannot practically be distinguished from instances with no nodes at the n^{th} level. In this case the corresponding decision problem is not solvable in polynomial time by this algorithm. We will divide families of decision trees indexed by n into two classes, those that satisfy (P) and those that do not, which we call impenetrable.

We are interested in using quantum mechanics to move through decision trees. We imagine that nodes of the decision tree correspond to quantum states, which give a basis for the Hilbert space. We further imagine constructing a Hamiltonian $\hat{H}$ with nonzero off-diagonal matrix elements only between states that are connected in the corresponding decision tree. (We will be more specific about constructing the Hilbert space and $\hat{H}$ later.) We start the quantum system in the state corresponding to the topmost node and let it evolve with its time evolution determined by $\hat{H}$ so that the unitary time evolution operator is

$$\hat{U}(t) = \exp(-i\hat{H}t) \quad . \tag{1.2}$$

At any time t we have a pure state that can be expressed as a (complex) superposition of basis states corresponding to the nodes. Given $\hat{H}$ and the initial state, the probability (the amplitude squared) of finding the system at the n^{th} level at time t is determined. We then say that a family of trees indexed by size n is quantum penetrable if condition (P) is met and it is quantum impenetrable if condition (P) is not met.

In the next section we will give a specific form for the quantum Hamiltonian $\hat{H}$ and then prove that any family of trees that is classically penetrable is associated with a closely related family of trees that is quantum penetrable. This will demonstrate that our model for quantum mechanically solving decision problems is at least as powerful as the classical probabilistic method. In Section Three we will go further and give an example of a family of decision trees that is classically impenetrable but which is quantum mechanically penetrable. This means that the quantum penetration is exponentially faster than the corresponding classical penetration for these trees. However, we have not yet identified general characteristics of a problem that guarantee that its associated decision trees are quantum penetrable. Furthermore, for the example considered, the problem associated with the classically impenetrable trees can be reformulated so that it is computationally simple to solve by an alternative, classical method.

In Section Four we discuss the construction of the Hilbert space and the Hamiltonian $\hat{H}$. The usual paradigm for quantum computation[2] envisages a string of, say, ℓ spin-1/2 particles that give rise to a 2^{ℓ} dimensional Hilbert space. Each elementary operation is a unitary transformation that acts on one or two spins at a time. We will show that the Hilbert space for our system can be constructed using ℓ spin 1/2 particles just as in a conventional quantum computer. Furthermore, for a large class of problems, the Hamiltonian that we construct is a sum of Hamiltonians that act on a fixed number of spins. In this sense[3] our quantum evolution through decision trees lies in the framework of conventional quantum computation.

2 Classical vs. Quantum Evolution through Trees

In the introduction we discussed a classical (that is, non-quantum) probabilistic rule for moving through decision trees. Here we are going to be more specific and state the rule in a way that gives rise to a continuous time Markov process. The rule is simply that if you are at a given node then you move to a connected node with a probability per unit time γ where γ is a fixed, time independent, constant. This means that in a time ϵ where $\gamma\epsilon \ll 1$, the probability of moving to a connected node is $\approx \gamma\epsilon$. Using a continuous time process, as opposed to saying that you move at discrete times, will make it easier when we compare with the continuous time evolution dictated by the unitary operator in (1.2).

We are now going to introduce some formalism that looks quantum mechanical but we are going to apply it to describe the classical Markov process. Suppose we are given a decision tree that has N nodes. (N may be as large as 2^{n+1} where n is the number of levels.) Index the nodes in some way by the integers $a = 1, \ldots, N$. Corresponding to the tree we construct an N-dimensional Hilbert space that has an orthonormal basis $\{|a\rangle\}$ with $a = 1, \ldots, N$ and accordingly $\langle a|b\rangle = \delta_{ab}$. Now we define a Hamiltonian $\hat{H}$ through its matrix elements in this basis:

$$\langle b|\hat{H}|a\rangle = \begin{cases} -\gamma & \text{for } a \neq b \text{ if node } a \text{ is connected to node } b \\ 0 & \text{for } a \neq b \text{ if node } a \text{ is not connected to node } b \end{cases}$$

$$\langle a|\hat{H}|a\rangle = \begin{cases} 3\gamma & \text{node } a \text{ is connected to three other nodes} \\ 2\gamma & \text{node } a \text{ is connected to two other nodes} \\ \gamma & \text{node } a \text{ is connected to one other node} \end{cases} \quad . \tag{2.1}$$

Return to the classical probabilistic rule for moving through a fixed tree and let

$$p_{ba}(t) = \text{Prob (go from } a \text{ to } b \text{ in time } t) \quad . \tag{2.2}$$

For a time ϵ where $\gamma\epsilon \ll 1$ we have

$$p_{ba}(\epsilon) = \begin{cases} -\epsilon\langle b|\hat{H}|a\rangle + \mathcal{O}(\epsilon^2) & \text{for } b \neq a \\ 1 - \epsilon\langle a|\hat{H}|a\rangle + \mathcal{O}(\epsilon^2) & \text{for } b = a \end{cases} \tag{2.3}$$

as a consequence of the definition of $\hat{H}$. For a classical Markov process, the probability of moving depends only on current position, not on history, so we have for any t_1 and t_2,

$$p_{ba}(t_1 + t_2) = \sum_c p_{bc}(t_2) p_{ca}(t_1) \quad . \tag{2.4}$$

Therefore

$$p_{ba}(t + \epsilon) = \sum_c p_{bc}(\epsilon) p_{ca}(t) \tag{2.5}$$

which for ϵ small gives

$$p_{ba}(t + \epsilon) = p_{ba}(t) - \epsilon \sum_c \langle b|\hat{H}|c\rangle p_{ca}(t) + \mathcal{O}(\epsilon^2) \tag{2.6}$$

where we have used (2.3). We see therefore that $p_{ba}(t)$ obeys the differential equation

$$\frac{d}{dt} p_{ba}(t) = -\sum_c \langle b|\hat{H}|c\rangle p_{ca}(t) \tag{2.7}$$

with the boundary condition $p_{ba}(0) = \delta_{ab}$. The solution to (2.7) is

$$p_{ba}(t) = \langle b|e^{-\hat{H}t}|a\rangle \quad . \tag{2.8}$$

Again, $p_{ba}(t)$ given by (2.8) is the **classical** probability of going from a to b in time t if you move through the tree with a probability per unit time γ of moving to a connecting node. As a check we should have that

$$\sum_b p_{ba}(t) = 1 \quad . \tag{2.9}$$

To see that this is the case note that $\hat{H}$ defined by (2.1) has a zero eigenvector,

$$|E = 0\rangle = \frac{1}{\sqrt{N}} \sum_{b=1}^{N} |b\rangle \quad . \tag{2.10}$$

Therefore

$$\begin{aligned}\sum_b p_{ba}(t) &= \sqrt{N}\, \langle E = 0|e^{-\hat{H}t}|a\rangle \\ &= \sqrt{N}\, \langle E = 0|a\rangle \\ &= 1 \quad .\end{aligned} \tag{2.11}$$

We have constructed the Hamiltonian $\hat{H}$ because of its utility in describing a classical Markov process. We now propose using the same Hamiltonian $\hat{H}$ to quantum mechanically evolve through the tree. Let

$$A_{ba}(t) = \langle b|e^{-i\hat{H}t}|a\rangle \tag{2.12}$$

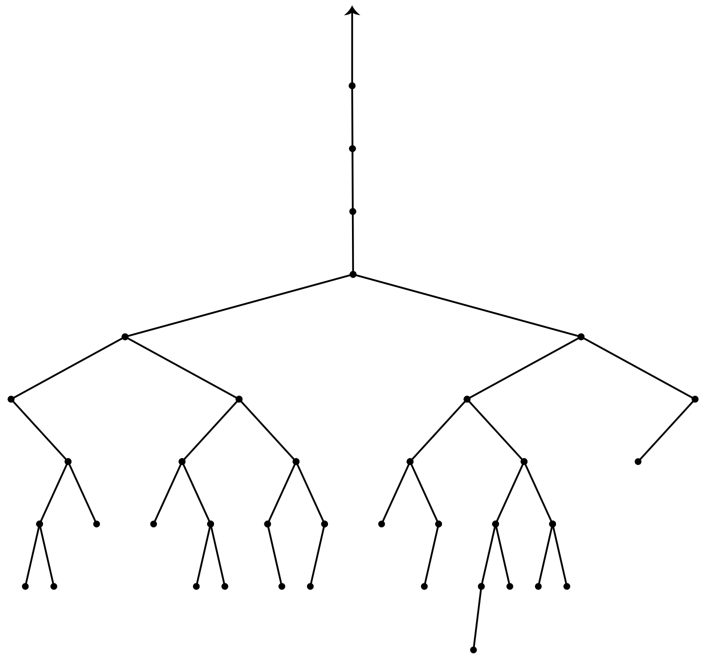

Figure 3: The tree T_n' obtained from the tree T_n of Fig. 2 by appending a semi-infinite line of nodes at the starting node of T_n.

be the quantum amplitude to be found at node b at time t given that you are at node a at time 0. In this case the probability is $|A_{ba}(t)|^2$ with

$$\sum_b |A_{ba}(t)|^2 = 1 \tag{2.13}$$

as a consequence of the fact that $\hat{H}$ is Hermitian. With this quantum Hamiltonian we will now show that if a family of trees is classically penetrable then there is a related family of trees that is also quantum mechanically penetrable.

Imagine we are given a family of decision trees $\{T_n\}$ where each T_n is n levels deep and does have nodes at the n-th level. For simplicity we will take the worst case possible and assume that there is only one node at level n. In order to establish our result we are going to consider another family of trees $\{T_n'\}$ where each T_n' is obtained from T_n by appending a semi-infinite line of nodes to the starting node of T_n. The rule for classically moving on the semi-infinite line is the same as the rule for moving on the rest of the tree: with a probability per unit time γ you move to an adjoining node.

We can see that if $\{T_n\}$ is classically penetrable so is $\{T_n'\}$. Roughly speaking, starting at 0 on T_n', the probability of reaching the n^{th} level is not appreciably reduced because of the time some paths spend on the semi-infinite line. (We now prove this statement, but the reader who is already convinced that it is true can skip to the next paragraph.) Suppose that for $\{T_n\}$ we have condition (P) so that

$$\text{Prob (go from 0 to } n \text{ in time } t) \geq \frac{1}{n^B} \tag{2.14}$$

for some $\gamma t \leq n^A$. At level 1 of the decision tree only one of the two nodes is on the branch that contains n, the unique node at level n. Denote this level 1 node by $\bar{1}$. Now for each path (on T_n) that reaches n from 0 in time t there is a time $t-s$ at which the path last jumps from 0 to $\bar{1}$. Thus

$$\begin{aligned} &\text{Prob (go from 0 to } n \text{ in time } t) = \\ \int_0^t ds \;\; &\text{Prob (go from 0 to 0 in time } t-s) \;\cdot\; \gamma \; ds \;\cdot \\ &\cdot \text{Prob (go from } \bar{1} \text{ to } n \text{ without hitting 0 in time } s) \;\; . \end{aligned} \tag{2.15}$$

Using (2.14) it follows that for some $\gamma s \leq \gamma t \leq n^A$,

$$\text{Prob (go from } \bar{1} \text{ to } n \text{ without hitting 0 in time } s) \;\geq \frac{1}{\gamma t n^B} \geq \frac{1}{n^{A+B}} \;\; . \tag{2.16}$$

However this last probability is the same for T_n' as for T_n. Turning to the trees T_n' we see that the node 0 is connected to three other nodes, the node at level -1 on the semi-infinite tree and the two nodes at level 1. In time $\frac{1}{\gamma}$ there is an n-independent lower bound on the probability of going from 0 to $\bar{1}$. Combining this fact with (2.16) we see that in a time $s + \frac{1}{\gamma}$ there is a probability of going from 0 to n on T_n' which is greater that $\frac{1}{n}$ to a power. Thus if $\{T_n\}$ is classically penetrable so is $\{T_n'\}$.

We are now going to compare the classical and quantum evolution through the family of trees $\{T_n'\}$. From this point on we set $\gamma = 1$.We will return to finite trees later in this section but for now the device of appending a semi-infinite line to the trees of interest actually makes the analysis simpler. Again call the starting node (which is at level 0 of the tree T_n') 0 and call the unique node at the n^{th} level n. Then

$$p(t) = \langle n | e^{-\hat{H}t} | 0 \rangle \tag{2.17}$$

is the probability to go from 0 to n in time t if you evolve with the classical rule. Similarly

$$A(t) = \langle n | e^{-i\hat{H}t} | 0 \rangle \tag{2.18}$$

is the quantum amplitude to be at n at time t if at $t = 0$ you are at 0 and you evolve with the quantum Hamiltonian $\hat{H}$. (Of course $\hat{H}, p(t)$ and $A(t)$ are all sequences that depend on the sequence $\{T_n'\}$ but we will not bother to place an n label on these quantities.)

The Hamiltonian $\hat{H}$ is defined by (2.1) for each tree T_n' but now the number of nodes is infinite so the Hilbert space is infinite dimensional. We will see shortly that the spectrum of $\hat{H}$ is purely continuous, that is, there are no normalizable eigenvectors.[1] Call the energy eigenvectors $|E\rangle$ where

$$\hat{H}|E\rangle = E|E\rangle$$

and

$$\langle E|E'\rangle = \delta(E - E') \quad . \tag{2.19}$$

Now for any Hermitian operator $\hat{H}$, with matrix elements H_{ab}, any eigenvalue E of $\hat{H}$ must lie[4] in the union (over a) of the intervals

$$|E - H_{aa}| \leq \sum_{b \neq a} |H_{ba}| \tag{2.20}$$

which, given the form (2.1), implies that $\hat{H}$ has no negative eigenvalues.

Using the completeness of the $|E\rangle$'s we can write (2.17) as

$$p(t) = \int dE e^{-Et} \langle n|E\rangle\langle E|0\rangle \tag{2.21}$$

and (2.18) as

$$A(t) = \int dE e^{-iEt} \langle n|E\rangle\langle E|0\rangle \tag{2.22}$$

where the integral dE is over the eigenvalue range of $\hat{H}$ (which is ≥ 0). From (2.22) we have

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dt' \; e^{iwt'} \; A(t') = \int dE \; \delta(w - E) \langle n|E\rangle\langle E|0\rangle \quad . \tag{2.23}$$

Multiply both sides by e^{-wt} and integrate dw from 0 to ∞ to get, for $t > 0$,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dt' \frac{A(t')}{t - it'} = p(t) \tag{2.24}$$

which could have been obtained using the Cauchy integral formula. Now in the $|a\rangle$ basis $\hat{H}$ is real and symmetric and from (2.18) it then follows that $A(t) = A^*(-t)$. This allows us to write (2.24) as

$$p(t) = \frac{1}{\pi} \; \mathrm{Re} \int_0^{\infty} dt' \frac{A(t')}{t - it'} \quad . \tag{2.25}$$

We will now use (2.25) to show that if a family of trees $\{T_n'\}$ is classically penetrable it is also quantum penetrable. Pick some time T and let ϵ be the maximum of $|A(t')|$ for

[1] In truth, there may be normalizable states, completely localized on the ends of the branching structure T_n, but they have strictly zero overlap with the states of interest and can be safely ignored.

$0 \le t' \le T$. Now

$$
\begin{aligned}
p(t) &= \frac{1}{\pi}\ \mathrm{Re}\ \left\{ \int_0^T dt' \frac{A(t')}{t - it'} + \int_T^\infty dt' \frac{A(t')}{t - it'} \right\} \\
&\le \frac{\epsilon}{\pi} \int_0^T dt' \frac{1}{(t^2 + t'^2)^{1/2}} + \frac{1}{\pi} \left| \int_T^\infty dt' \frac{A(t')}{t - it'} \right| \\
&= \frac{\epsilon}{\pi}\ \ln\ \left[\frac{(T^2 + t^2)^{1/2} + T}{t} \right] + \frac{1}{\pi} \left| \int_T^\infty dt' \frac{A(t')}{t - it'} \right| \quad .
\end{aligned}
\tag{2.26}
$$

The magnitude of the last integral in (2.26) is actually less than $C/T^{1/4}$ for large T where C is an n-independent constant. We will show this shortly. With this result we then have that

$$
p(t) \le \frac{\epsilon}{\pi}\ \ln\ \left[\frac{(T^2 + t^2)^{1/2} + T}{t} \right] + \frac{C}{T^{1/4}} \quad . \tag{2.27}
$$

Now we are assuming that the family of trees is classically penetrable. This means that for some $t \le n^A$ we have $p(t) > 1/n^B$ for some A and B. For large n, this penetration time t is clearly ≥ 1. Since the ln term in (2.27) is a decreasing function of t, we have

$$
\frac{1}{n^B} \le \frac{\epsilon}{\pi} \ln \left[(T^2 + 1)^{1/2} + T \right] + \frac{C}{T^{1/4}} \quad . \tag{2.28}
$$

Now let $T = n^D$ for $D > 4B$. We then have, for large enough n,

$$
\frac{1}{n^B} \le \frac{\epsilon}{\pi} \ln(n^D) \tag{2.29}
$$

which means that the maximum of $|A(t)|$ for $t < n^D$ is bigger than a constant times $1/n^{B+1}$ Thus we have the result that if a family of trees $\{T'_n\}$ is classically penetrable, it is also quantum penetrable.

Before going further we want to verify that the last integral in (2.26) is actually bounded as claimed. To this end we need to establish some facts about the eigenfunctions of $\hat{H}$. Label the nodes on the semi-infinite line of T'_n by j with $j = 0, -1, -2, \ldots$ so that $j = 0$ is the starting node of T_n. On the semi-infinite line

$$
\hat{H}|j\rangle = 2|j\rangle - |j+1\rangle - |j-1\rangle \qquad \text{for } j \le -1 \quad . \tag{2.30}
$$

The state $|\theta\rangle$ with $\langle j|\theta\rangle$ proportional to $e^{ij\theta}$ is an eigenstate of (2.30) with energy

$$
E(\theta) = (2 - 2\cos\theta) = 4\sin^2\theta/2 \quad . \tag{2.31}
$$

Now $e^{ij\theta}$ and $e^{-ij\theta}$ correspond to the same energy but because of the finite branching part of the tree (T_n, which is connected at $j = 0$), only one linear combination is an eigenfunction of the full $\hat{H}$,

$$
\langle j|\theta\rangle = \frac{1}{(2\pi)^{1/2}} \left[e^{ij\theta} + R(\theta) e^{-ij\theta} \right] \tag{2.32}
$$

with $0 \le \theta \le \pi$, and $R(\theta)$ is determined by the structure of T_n. Because $\hat{H}$ in the node basis is real, (2.32) must be real up to an overall j independent phase. This implies that $R(\theta)$ is of the form $e^{-2i\delta(\theta)}$, that is, $|R(\theta)| = 1$. (The form (2.32) is an "in" state for scattering off of the tree T_n at the end of the semi-infinite line. The fact that $|R(\theta)| = 1$ is also a consequence of the unitarity of the S matrix.) We can rewrite (2.32) as

$$\langle j|\theta\rangle = e^{-i\delta(\theta)} \frac{2}{(2\pi)^{1/2}} \cos(j\theta + \delta(\theta)) \tag{2.33}$$

and then absorb the phase in the definition of $|\theta\rangle$ to get

$$\langle j|\theta\rangle = \left(\frac{2}{\pi}\right)^{1/2} \cos(j\theta + \delta(\theta)) \quad . \tag{2.34}$$

We can also now see that there are no normalizable energy eigenvectors. To be normalizable and an energy eigenvector, on the semi-infinite line the amplitude must be of the form $e^{j\alpha}$ with α real and positive. However, this state has energy $(2 - 2\cosh\alpha)$ that is less than zero and therefore excluded since $\hat{H}$ has no negative eigenvalues.

The states $|\theta\rangle$ form a complete set of delta function normalized eigenstates, i.e.,

$$\langle\theta|\theta'\rangle = \delta(\theta - \theta')$$

and

$$\langle j|k\rangle = \int_0^\pi d\theta \langle j|\theta\rangle\langle\theta|k\rangle = \delta_{jk} \qquad \text{for } j,k = 0,-1,-2,\ldots \quad . \tag{2.35}$$

We have introduced the states $|\theta\rangle$ because we could (fairly) easily normalize them, that is, pick the coefficient in (2.32) so that (2.35) holds. The energy eigenstates $|E\rangle$ given by (2.19) are proportional to the $|\theta\rangle$'s. To maintain both (2.19) and (2.35) we have

$$|E\rangle = \left(\frac{dE}{d\theta}\right)^{-1/2} |\theta\rangle = (4E - E^2)^{-1/4} |\theta\rangle \tag{2.36}$$

where again $E = 4\sin^2\theta/2$. In the node basis on the semi-infinite line we then have

$$\langle j|E\rangle = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{(4E - E^2)^{1/4}} \cos(j\theta + \delta(E)) \quad . \tag{2.37}$$

We now use these states to help us bound the integral of interest. First note that

$$A(t') = \langle n|e^{-i\hat{H}t'}|0\rangle = \int_0^4 dE \langle n|E\rangle\langle E|0\rangle e^{-iEt'} \tag{2.38}$$

and the integral in (2.26) we wish to bound is (after dividing by i)

$$\begin{aligned} \int_T^\infty dt' \frac{A(t')}{t' + it} &= \int_T^\infty dt' \int_0^4 dE \langle n|E\rangle\langle E|0\rangle e^{-iEt'} \frac{1}{t' + it} \\ &= \int_T^\infty dt' \int_0^4 dE\, \langle n|E\rangle\langle E|0\rangle e^{-iEt'} \int_0^\infty d\mu\, e^{-\mu(t'+it)} \qquad (2.39) \\ &= \int_0^4 dE \int_0^\infty d\mu\, \langle n|E\rangle\langle E|0\rangle e^{-i\mu t} e^{-iET} e^{-\mu T} \frac{1}{\mu + iE} \quad . \end{aligned}$$

Taking the absolute value we get

$$\Big|\int_T^\infty dt'\frac{A(t')}{t'+it}\Big| \le \int_0^4 dE \int_0^\infty d\mu \, \frac{|\langle n|E\rangle\langle E|0\rangle| e^{-\mu T}}{(\mu^2+E^2)^{1/2}} \quad . \tag{2.40}$$

The matrix element $\langle E|0\rangle$ is given by (2.37) so we have $|\langle E|0\rangle| \le C/(4E-E^2)^{1/4}$ where C is a generic constant. Thus we have

$$\Big|\int_T^\infty dt'\frac{A(t')}{t'+it}\Big| \le C\int_0^\infty d\mu\, e^{-\mu T}\int_0^4 dE\frac{|\langle n|E\rangle|}{(4E-E^2)^{1/4}(\mu^2+E^2)^{1/2}} \quad . \tag{2.41}$$

Now by the Cauchy-Schwarz inequality

$$\Big|\int_T^\infty dt'\frac{A(t')}{t'+it}\Big| \le C\int_0^\infty d\mu\, e^{-\mu T}\Big\{\int_0^4 dE'|\langle n|E'\rangle|^2\Big\}^{1/2}\Big\{\int_0^4 dE\frac{1}{(4E-E^2)^{1/2}(\mu^2+E^2)}\Big\}^{1/2} \tag{2.42}$$

Because $\langle n|n\rangle = 1$ and the $|E\rangle$ states are compete, the integral dE' gives 1. The integral dE is

$$\int_0^4 dE\frac{1}{(4E-E^2)^{1/2}}\,\frac{1}{\mu^2+E^2} = \int_0^\pi d\theta\frac{1}{\mu^2+(4\sin^2\theta/2)^2} < \int_0^\pi d\theta\frac{1}{\mu^2+(2/\pi)^4\theta^4} \le \frac{C}{\mu^{3/2}} \quad . \tag{2.43}$$

Now the inequality (2.42) becomes

$$\Big|\int_T^\infty dt\ \frac{A(t')}{t'+it}\Big| \le C\int_0^\infty d\mu\, e^{-\mu T}\frac{1}{\mu^{3/4}} \le \frac{C}{T^{1/4}} \quad . \tag{2.44}$$

which is the desired result. This was the last step we needed in showing that if $\{T_n'\}$ is classically penetrable then it is quantum penetrable.

Of course we are not ultimately interested in quantum evolving on the family of infinite trees $\{T_n'\}$ because we only imagine building a quantum computer with a finite number of building blocks. However we now argue that if the family $\{T_n'\}$ is quantum penetrable there is a closely related family of finite trees $\{T_n^f\}$ that is also quantum penetrable. In fact T_n^f is obtained from T_n' by chopping off the semi-infinite line at some node that is far, but not exponentially far as a function of n, from the node 0. Alternatively we can view T_n^f as arising from T_n by appending to T_n at 0 a finite number of linearly connected nodes.

To understand when infinite and very long give rise to the same quantum evolution consider an infinite line of nodes by itself with the Hamiltonian given by (2.30). In this case it is possible to explicitly evaluate the amplitude to go from j to k in time t:

$$\langle k|e^{-i\hat{H}t}|j\rangle = e^{-2it}i^{(k-j)}J_{k-j}(2t) \tag{2.45}$$

where J_{k-j} is a Bessel function of integer order. For fixed t this amplitude dies rapidly if $|k-j|$ is bigger than $2t$. Imagine starting at $j=0$ at $t=0$. The quantum amplitude spreads out with speed 2 (recall that we have set γ=1). Chopping off the infinite system at the nodes $\pm L$ will not affect the evolution from $j = 0$ as long as $L \gg 2t$.

Return to the family of quantum penetrable trees $\{T'_n\}$. These trees have the property that starting at 0, which is at the end of the semi-infinite line, there is a substantial quantum amplitude for being at the node on the n^{th} level of the branching tree at a time $t \leq n^{\bar{A}}$ for a fixed $\bar{A}$. Lopping off the infinite tree at a node of order $(n^{\bar{A}})^2$ down from 0 will not affect this result. Thus the family of finite trees $\{T^f_n\}$, which are obtained from the family of classically penetrable trees $\{T_n\}$ by adding a finite number of linearly connected nodes, is quantum penetrable.

It is reasonable to ask why we bother with the family of infinite trees $\{T'_n\}$ when we are only actually interested in finite trees. Why didn't we prove directly that the family of classically penetrable trees $\{T_n\}$ is also quantum penetrable? Of course the answer is we would have if we could have. The difficulty lies in the fact that for an arbitrary finite tree with an exponential number of nodes there are an exponential number of energy eigenvalues falling in a fixed interval and we were unable to establish the requisite facts about the density of states needed for a proof.

Let us summarize the results of this section. We started with a given family of trees $\{T_n\}$ that was assumed to be classically penetrable. We then constructed the closely related family of trees $\{T'_n\}$ that has a semi-infinite line of nodes attached to the starting node of each T_n. The trees $\{T'_n\}$ are also classically penetrable. Then, using the analytic relationship between the classical probabilities and quantum amplitudes of $\{T'_n\}$ we were able to prove that $\{T'_n\}$ is quantum penetrable. We also argued that cutting the semi-infinite line at some node that is a polynomial in n distance from 0 cannot affect the quantum penetrability as long as this distance is much greater than the quantum penetration time. Therefore the family $\{T^f_n\}$ of trees that is made from $\{T_n\}$ by appending a long (but finite) string of nodes to the starting node of each T_n is quantum penetrable if the original $\{T_n\}$ is classically penetrable. Clearly $\{T_n\}$ and $\{T^f_n\}$ are addressing precisely the same decision question. Therefore any problem that can be solved by classically random walking through a decision tree can be solved by quantum evolving through a very closely related tree.

3 A Family of Trees that is Quantum, but Not Classically, Penetrable

If we know enough about the structure of a family of trees we can decide if it is classically penetrable and if it is quantum penetrable. Here we will show examples of families of trees that are quantum but not classically penetrable. We begin by discussing the calculations in the quantum case. As in the last section we consider a family of trees $\{T_n\}$ whose members have only one node at the n^{th} level, called n. This time we construct the family $\{T''_n\}$ where each tree T''_n has two semi-infinite lines of nodes, one connected to the starting node of T_n, and the other semi-infinite line of nodes attached to the node n of T_n. For calculational

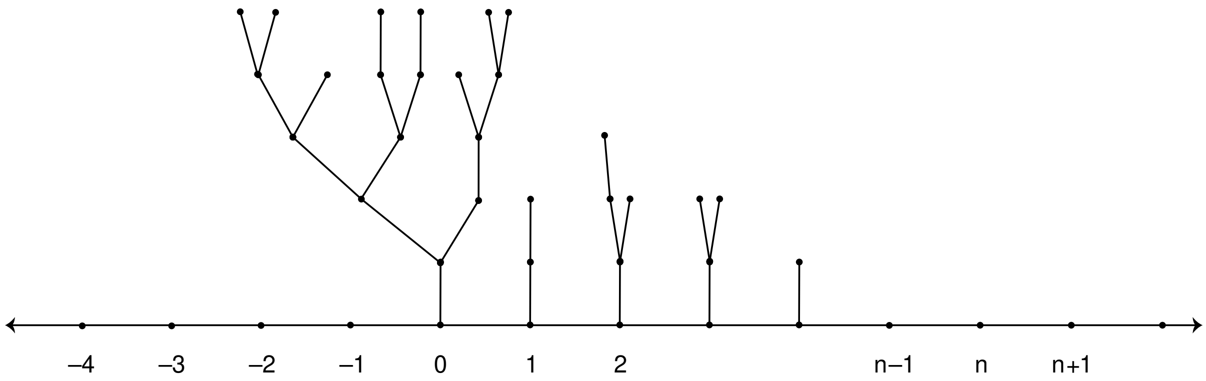

Figure 4: The tree T_n'' obtained from the tree T_n of Fig. 2 by appending two semi-infinite lines, one connected at the starting node and one connected to the node n. The tree is drawn with the direct line of nodes from 0 to n along the base.

purposes we make these two extra lines of nodes semi-infinite but ultimately we envisage making them of length n to a power.

For convenience we redraw our trees so that the direct line of nodes from 0 to n lies along the base. In this way the tree depicted in Fig. 2 with two semi-infinite lines appended becomes that of Fig. 4. We use "bush" to denote a group of nodes coming out of a node on the base. Here we label the nodes on the base by j. The nodes $j = -1, -2, -3, \ldots$ are on the semi-infinite starting line. The nodes $j = n+1, n+2, \ldots$ are on the appended ending line. The nodes $j = 0, \ldots, n$ are all on the original tree T_n and the nodes $0, \ldots, n-2$ may have bushes coming out them although the nodes $n-1$ and n do not. (If node $n-1$ had a bush then n would not be the unique level n node.) What we imagine doing is building a quantum state localized near 0 on the starting line and then calculating the quantum amplitude for penetrating the tree and being on the ending line. To this end we now set up the formalism for calculating the energy dependent transmission coefficient $T(E)$ and then evaluate it in certain specific cases of families of trees.

For the tree depicted in Fig. 4 with an infinite base, for each energy E with $0 \leq E \leq 4$, there are two energy eigenstates. (Here again we have set γ equal to 1). On the semi-infinite lines they are, in the node basis, of the form $e^{ij\theta}$ and $e^{-ij\theta}$ where again $E = 4\sin^2\theta/2$ and $0 \leq \theta \leq \pi$. Superpositions of the $e^{ij\theta}$ are used to make right moving packets whereas superpositions of $e^{-ij\theta}$ make left movers. Consider the state $|E, +\text{in}\rangle$ that on the starting and ending lines is of the form

$$\begin{aligned} \langle j|E, +\text{in}\rangle &= N(E)[e^{ij\theta} + R(E)e^{-ij\theta}] \qquad && j = -1, -2, \ldots \\ \langle j|E, +\text{in}\rangle &= N(E)T(E)e^{ij\theta} && j = n-1, n, n+1, \ldots \end{aligned} \tag{3.1}$$

with

$$N(E) = \frac{1}{(2\pi)^{1/2}} \frac{1}{(4E - E^2)^{1/4}} \quad .$$

At this point we say nothing about $\langle a|E,+\text{in}\rangle$ if a is a node on T_n. Superpositions of $|E,+\text{in}\rangle$ make states that at early times represent right moving packets on the starting line headed towards the tree structure T_n. At late times the packet splits into a reflected piece, proportional to R, left moving on the starting line, and a transmitted piece, proportional to T, which is right moving on the ending line. Similarly we can define $|E,-\text{in}\rangle$, which represents a state left moving on the ending line at early times that at late times is split into a right mover on the ending line and a transmitted part left moving on the starting line. For $|E,-\text{in}\rangle$ we have

$$\begin{aligned} \langle j|E,-\text{in}\rangle &= N(E)[e^{-ij\theta} + \bar{R}(E)e^{ij\theta}] \qquad && j = n-1, n, n+1, \dots \\ \langle j|E,-\text{in}\rangle &= N(E)\bar{T}(E)e^{-ij\theta} && j = -1, -2, \dots \quad . \end{aligned} \tag{3.2}$$

The states $|E,+\text{in}\rangle$ and $|E,-\text{in}\rangle$ are a complete set of scattering states useful for discussing tree penetration. Equivalently there is the set $|E,+\text{out}\rangle$ and $|E,-\text{out}\rangle$ that at late times represents respectively a right mover on the ending line and a left mover on the starting line. From (3.1) and (3.2) we get

$$\begin{aligned} |E,+\text{in}\rangle &= R(E)|E,-\text{out}\rangle + T(E)|E,+\text{out}\rangle \\ |E,-\text{in}\rangle &= \bar{R}(E)|E,+\text{out}\rangle + \bar{T}(E)|E,-\text{out}\rangle \quad . \end{aligned} \tag{3.3}$$

This transformation from the out states to the in states is called the S-matrix,

$$S = \begin{pmatrix} R & T \\ \bar{T} & \bar{R} \end{pmatrix} \tag{3.4}$$

which is necessarily unitary so we have

$$\begin{aligned} |R(E)|^2 + |T(E)|^2 &= 1 \\ |\bar{R}(E)|^2 + |\bar{T}(E)|^2 &= 1 \\ R^*(E)T(E) + \bar{T}^*(E)\bar{R}(E) &= 0 \quad . \end{aligned} \tag{3.5}$$

The standard interpretation of $T(E)$ is as follows. Suppose we build a state $|\psi\rangle$ completely on the starting line, that is, $\langle a|\psi\rangle$ is nonzero only for nodes a on the starting line. Furthermore suppose that $|\psi\rangle$ expanded as a superposition of energy eigenstates is made only of states whose energy is close to some E_0. If we quantum mechanically evolve $|\psi\rangle$ with the unitary operator $e^{-i\hat{H}t}$, then at late times the probability of being on the ending line is $|T(E_0)|^2$. Thus $|T(E)|^2$ has a direct interpretation as the E dependent transmission probability through the tree.

Of course any state $|\psi\rangle$ that is highly localized in energy is necessarily highly delocalized in the node basis. (This can be viewed as a consequence of the uncertainty principle.) We don't want our constructions to rely on building states that are very spread out on the starting line since we eventually do wish to chop it off not too far from the node 0. Suppose we start at a specific node, j on the starting line, and we want the amplitude for being at node k on the ending line at time t. This is given by

$$\begin{aligned}
A_{kj}(t) &= \langle k|e^{-i\hat{H}t}|j\rangle \\
&= \int_0^4 dE\,\{\langle k|E,+\text{in}\rangle\langle E,+\text{in}|j\rangle + \langle k|E,-\text{in}\rangle\langle E,-\text{in}|j\rangle\}\,e^{-iEt} \qquad (3.6)\\
&= \int_0^4 dEN^2(E)\left\{T(E)e^{ik\theta}(e^{-ij\theta}+R^*(E)e^{ij\theta}) + (e^{-ik\theta}+\bar{R}(E)e^{ik\theta})\bar{T}^*(E)e^{ij\theta}\right\}e^{-iEt}
\end{aligned}$$

where we have used the explicit forms for $|E,\pm\text{in}\rangle$ on the starting and ending lines. Now using the last equation in (3.5) we get

$$A_{kj}(t) = \int_0^4 dEN^2(E)\left\{T(E)e^{i(k-j)\theta} + \bar{T}^*(E)e^{-i(k-j)\theta}\right\}e^{-iEt} \quad . \tag{3.7}$$

As a consequence of (3.5) we have that $|T(E)| = |\bar{T}(E)|$. Furthermore $\hat{H}$ is real in the node basis, so we can show that $T(E) = \bar{T}(E)$. Thus even for a state initially localized at a single node, the function $T(E)$ determines the amplitude for being at any node on the ending line at time t.

We now turn to calculating $T(E)$, which clearly depends on the structure of the tree to which we have added the semi-infinite starting and ending lines of nodes. For each of the nodes $m = 0, 1, \ldots, n-2$ along the base of the tree – see Fig. 4 – that has a bush sprouting up from it, let us define

$$y_m(E) = \frac{\langle \text{node above } m|E,+\text{in}\rangle}{\langle m|E,+\text{in}\rangle} \tag{3.8}$$

where $|\text{node above } m\rangle$ is the state corresponding to the node one level up from the base above the node m. Now for fixed E, $y_m(E)$ is determined solely by the bush coming out of the node m; it does not depend on the other bushes. To see this suppose that the bush coming out of node m has N nodes above the base node m. Label these nodes by $a = 1, \ldots, N$. Now $\hat{H}|a\rangle$ gives a superposition of $|a\rangle$ and the states connected to a. Thus

$$\langle a|\hat{H}|E,+\text{in}\rangle = E\langle a|E,+\text{in}\rangle \tag{3.9}$$

is N equations for the $(N+1)$ quantities $\langle a|E,+\text{in}\rangle$ and $\langle m|E,+\text{in}\rangle$. Divide through by $\langle m|E,+\text{in}\rangle$ and we get N equations for the N ratios $\langle a|E,+\text{in}\rangle/\langle m|E,+\text{in}\rangle$ so we see that (3.8) is determined by the bush alone. Furthermore the equations that were used to determine $y_m(E)$ are all real so $y_m(E)$ is also real. For any given bush $y_m(E)$ can be calculated recursively by looking at sub-bushes and it is not actually necessary to solve the N equations (3.9).

Let m be a node on the base with a bush coming off. Now, from (2.1),

$$\begin{aligned}
\langle m|\hat{H}|E,+\text{in}\rangle &= 3\langle m|E,+\text{in}\rangle \\
&\qquad -\langle m+1|E,+\text{in}\rangle - \langle m-1|E,+\text{in}\rangle - \langle \text{node above } m|E,+\text{in}\rangle \\
&= E\langle m|E,+\text{in}\rangle \qquad (3.10)
\end{aligned}$$

which implies that

$$\langle m+1|E,+\mathrm{in}\rangle = (3 - E - y_m(E))\langle m|E,+\mathrm{in}\rangle - \langle m-1|E,+\mathrm{in}\rangle \tag{3.11}$$

where we have used (3.8). If m has no bush coming out of it, a parallel argument gives

$$\langle m+1|E,+\mathrm{in}\rangle = (2 - E)\langle m|E,+\mathrm{in}\rangle - \langle m-1|E,+\mathrm{in}\rangle \quad . \tag{3.12}$$

We can use (3.11) for nodes with bushes as well as without if we define $y_m(E) = 1$ for nodes on the base with no bushes above. Equation (3.11) can be written as a matrix equation

$$\begin{bmatrix} \langle m+1|E,+\mathrm{in}\rangle \\ \langle m|E,+\mathrm{in}\rangle \end{bmatrix} = \begin{bmatrix} (3 - E - y_m(E)) & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \langle m|E,+\mathrm{in}\rangle \\ \langle m-1|E,+\mathrm{in}\rangle \end{bmatrix} \quad . \tag{3.13}$$

We then have

$$\begin{bmatrix} \langle n|E,+\mathrm{in}\rangle \\ \langle n-1|E,+\mathrm{in}\rangle \end{bmatrix} = M \begin{bmatrix} \langle 0|E,+\mathrm{in}\rangle \\ \langle -1|E,+\mathrm{in}\rangle \end{bmatrix} \quad . \tag{3.14}$$

where

$$M = M_{n-1}\, M_{n-2} \cdots M_0 \tag{3.15}$$

and

$$M_m = \begin{bmatrix} (3 - E - y_m(E)) & -1 \\ 1 & 0 \end{bmatrix} \quad . \tag{3.16}$$

Substituting the explicit form for $|E,+\mathrm{in}\rangle$ from (3.1) we get

$$\begin{bmatrix} T(E)e^{in\theta} \\ T(E)e^{i(n-1)\theta} \end{bmatrix} = M \begin{bmatrix} 1 + R(E) \\ e^{-i\theta} + R(E)e^{i\theta} \end{bmatrix} \quad . \tag{3.17}$$

If we know the matrix M, $T(E)$ is determined by these last two equations for $T(E)$ and $R(E)$. From (3.16) we see that M is the product of matrices of determinant 1 so $\det(M) = 1$. We can write

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{3.18}$$

with $ad - bc = 1$ and a, b, c, d all real. Solving for $T(E)$ we get

$$T(E) = e^{-in\theta} \frac{2i\sin\theta}{c - b + (d-a)\cos\theta + i(d+a)\sin\theta} \quad . \tag{3.19}$$

It is interesting to note that if for some E we have $y_m(E) = 1$ for all m, then $T(E) = 1$. To see this we construct $M = M(E)$ in this special case. From (3.15) and (3.16) we have

$$\begin{aligned} M(E) &= \begin{bmatrix} 2 - E & -1 \\ 1 & 0 \end{bmatrix}^n \\ &= \frac{1}{\sin(\theta)} \begin{bmatrix} \sin\big((n+1)\theta\big) & -\sin(n\theta) \\ \sin(n\theta) & -\sin\big((n-1)\theta\big) \end{bmatrix} \quad . \end{aligned} \tag{3.20}$$

Plugging into (3.18) and (3.19) we get $T(E)=1$. To understand why this comes about recall that a node with no bush is the same as a node with a bush for which $y_m(E)=1$ as far as the calculation of $T(E)$ is concerned. Therefore if all bushes have $y_m(E)=1$ at some E we have unimpeded transmission at that E.

To recap, given a decision tree T_n with one node at level n, construct a new tree with semi-infinite lines attached to the starting node 0 and to the node at level n. Redraw the tree as in Fig. 4 with the direct line from 0 to n along the base. Suppose we can calculate the $n-1$ functions $y_0(E), y_1(E), \ldots, y_{n-2}(E)$. Substitute into (3.16) and (3.15) to get the matrix M as a function of E. The transmission coefficient $T(E)$ is then given by (3.19) where $E = 4\sin^2\theta/2$.

In order for a family of trees to be quantum penetrable, the function $|T(E)|$ must be not too small over a not too small range of E as can be seen from (3.7). Furthermore even if $|T(E)|$ is not small, $T(E)$ must not oscillate rapidly about zero or else the integral (3.7) may be small due to cancellations. It is interesting to note that for any tree $T(E) \to 1$ as $E \to 0$. To see this note that the zero-energy eigenvector of $\hat{H}$, $|E=0,+\mathrm{in}\rangle$, is constant in the node basis, that is, $\langle a|E=0,+\mathrm{in}\rangle$ is independent of a. Thus $y_m(0)$ defined by (3.8) is 1 for all nodes on the base and by the argument of the paragraph before last we have $T(0)=1$. For trees that are not quantum penetrable we will see that although $T(0)=1$, $T(E)$ falls to near zero for an exponentially small value of E.

Consider a decision tree that is perfectly bifurcating until level $n-1$ and then only one of the 2^{n-1} nodes at level $n-1$ continues on to level n. The associated tree T_n is shown in Fig. 5. This decision tree could arise from the following question. You are given a list of $N=2^{n-1}$ items with the knowledge that a single unspecified item may or may not be marked. The question is, "Is there a marked item?" (This is essentially the problem for which Grover [5] found a quantum algorithm requiring order $\sqrt{N}$ steps.) Any classical algorithm for solving this problem requires of order N steps. In particular the Markov process for moving through the decision tree gives a probability of being at the unique node at level n that is at most of order $1/N$, so this family of trees is classically impenetrable.

We now turn to quantum evolution through the same set of trees. Draw the tree in Fig. 5 with the direct line from 0 to n along the base and add semi-infinite starting and ending lines; see Fig. 6. We see that each bush coming out of the base at node m is a perfectly bifurcating bush of length $n-1-m$ for $m=0$ to $n-1$. The ratio $y_m(E)$ can be calculated for each of these bushes. Consider one such bush of length $k=n-1-m$ as depicted in Fig. 7. At height ℓ, with $1 \le \ell \le k$, there are $2^{\ell-1}$ nodes. At each height we define the normalized state

$$|\ell; pb\rangle = \frac{1}{(2^{\ell-1})^{1/2}} \sum_{a \text{ at height } \ell} |a\rangle \tag{3.21}$$

with $|0;pb\rangle$ being the state at the node on the bottom of the bush, that is, $|0;pb\rangle = |m\rangle$.

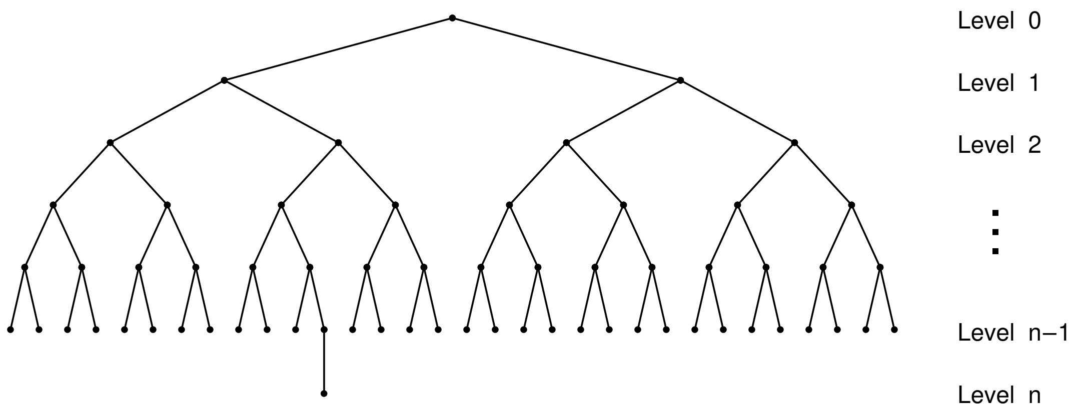

Figure 5: The tree, T_n, which is perfectly bifurcating for the first $n-1$ levels and then has only one node at level n.

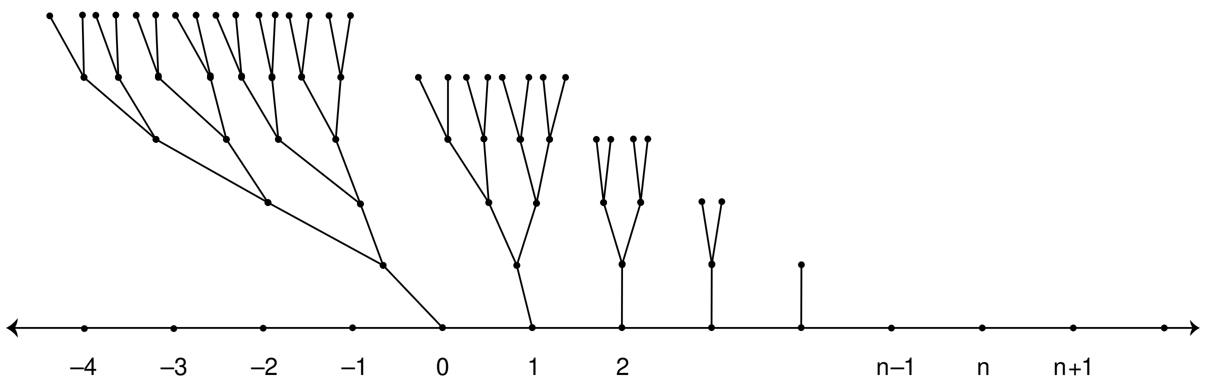

Figure 6: The tree T''_n constructed from T_n of Fig. 5 by appending two semi-infinite lines of nodes and drawing the direct line of nodes from 0 to n along the base.

With these labels, for these bushes, $y_m(E)$ defined by (3.8) is

$$y_m(E) = \frac{\langle 1; pb|E, + \text{ in}\rangle}{\langle 0; pb|E, + \text{ in}\rangle} \quad . \tag{3.22}$$

Note that $\hat{H}$ to any power acting on $|0;pb\rangle$ gives a linear superposition of states that only contains the states $|\ell;pb\rangle$ on the bush. Further note that

$$\langle \ell; pb|\hat{H}|\ell'; pb\rangle = 3\delta_{\ell\ell'} - \sqrt{2}[\delta_{\ell,\ell'+1} + \delta_{\ell,\ell'-1}] \qquad \text{for } 1 \le \ell,\ \ell' \le k-1 \tag{3.23}$$

so the bush in Fig. 7 can be replaced by the effective linear bush given in Fig. 8 where the number next to the node on the right gives the diagonal element of the Hamiltonian and the number by the connecting edge on the left gives the off-diagonal element. Up to an overall constant that drops out of 3.22, for $\ell = 1$ to k we have

$$\langle \ell; pb|E, + \text{ in}\rangle = \cos(\ell\theta' + \alpha)$$

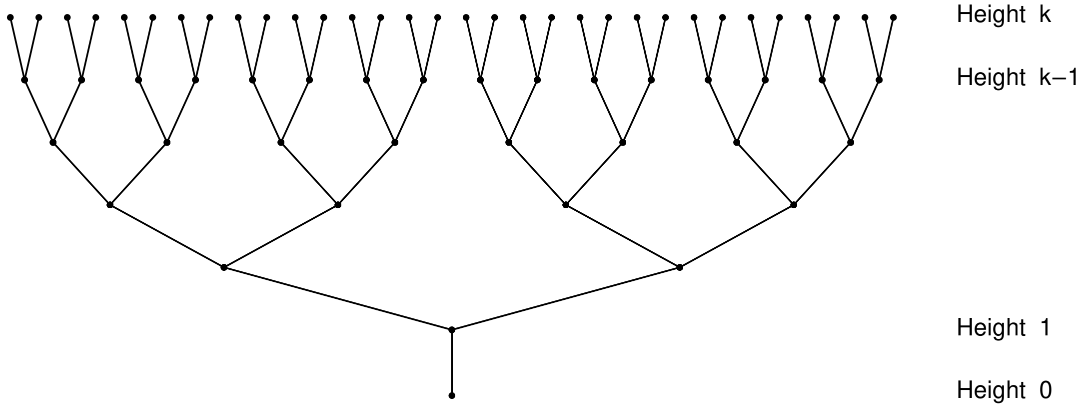

Figure 7: A perfectly bifurcating bush of height k coming out of the base of the tree in Fig. 6 at node $m = n - 1 - k$.

and

$$\langle 0; pb | E, +\text{in} \rangle = \sqrt{2} \cos \alpha \tag{3.24}$$

with

$$E = 3 - 2\sqrt{2} \cos \theta' \quad .$$

By applying $\hat{H}$ to the $\ell = k$ node we can determine α,

$$\tan(k\theta' + \alpha) = \frac{\cos(\theta') - \sqrt{2}}{\sin \theta'} \quad . \tag{3.25}$$

Going back to (3.22) we then have

$$y_m(E) = \frac{1}{\sqrt{2}} \left\{ \frac{\sqrt{2} \sin((k-1)\theta') - \sin(k\theta')}{\sqrt{2} \sin(k\theta') - \sin((k+1)\theta')} \right\} \tag{3.26}$$

where again $k = n - 1 - m$. Of course the calculation of $y_m(E)$ in this example was greatly facilitated by the regularity of the bush.

With $y_m(E)$ determined for each bush we can evaluate $T(E)$ by substituting into (3.16), (3.15) and then (3.19). In Fig. 9 we show $|T(E)|$ for $n = 26$. At the $n-1$ level there are $2^{25} = 10^{7.5}$ nodes. Although $T(0) = 1$, $T(E)$ has fallen substantially by $E = 10^{-10}$. Most of the area under the curve comes from E of order 1. We can evaluate $T(E)$ explicitly at $E = 3$. Note from (3.24) that $\theta' = \pi/2$ at $E = 3$. In this case $y_m(3)$ is 1 if $k = n - 1 - m$ is even and $y_m(3)$ is $-1/2$ if k is odd. Thus $M(3)$ can be written as (for n even)

$$\begin{aligned} M(3) &= \left\{ \begin{bmatrix} 1/2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \right\}^{n/2} \\ &= (-1)^{n/2} \begin{bmatrix} 3/2 & 1/2 \\ 1 & 1 \end{bmatrix}^{n/2} \\ &= (-1)^{n/2} \begin{bmatrix} 1 & -1/3 \\ 1 & 2/3 \end{bmatrix} \begin{bmatrix} 2^{n/2} & 0 \\ 0 & 2^{-n/2} \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ -1 & 1 \end{bmatrix} \end{aligned} \tag{3.27}$$

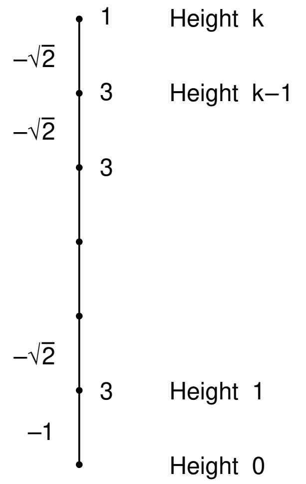

Figure 8: The effective bush of height k associated with the bush of Fig. 7. The number to the left of each edge gives the matrix element of $\hat{H}$ between the two states connected by the edge. The number next to the node gives the diagonal element of $\hat{H}$ for that state.

from which we conclude that $T(3) \sim 2^{-n/2}$. The transmission amplitude is of order $2^{-n/2}$ so the transmission probability goes like 2^{-n}. Here the quantum algorithm is doing no better than the classical algorithm.

Now we make a seemingly small modification of the tree. We take all of the odd-height bushes coming out of the base line of Fig. 6 and trim back one layer so all bushes are of even height. The magnitude of the transmission coefficient is shown in Fig. 10 where we see that for a substantial range of E near 3, $|T(E)|$ is very close to 1. In fact for all of these teeth, $y_m(3) = 1$, which by the argument given above implies that $T(3) = 1$. We can also see that $T(E)$ does not oscillate rapidly in this region by plotting the real part of $T(E)$, which is shown in Fig. 11, confirming a more tedious analytic evaluation. Therefore the family of trees is quantum penetrable.

It is easy to see that these trees with even-height bushes are not classically penetrable. Before trimming back the odd-height bushes we had the n-level tree shown in Fig. 5, T_n, which is associated with the tree T''_n shown in Fig. 6. These trees are not classically penetrable. Now, if we trim the odd-height bushes back one layer, the trimmed tree still contains all of the tree T''_{n-1}, which has even- and odd-height bushes. Since T''_{n-1} is not classically penetrable, the even-height bush family is also not classically penetrable, since, classically, any time you add nodes to bushes you necessarily decrease the chances of getting to the node n.

We have given a single example of a family of trees that is not classically penetrable but *is* quantum penetrable. Clearly there are many variants of this example using even-length, perfectly bifurcating bushes in all sorts of combinations; we will not pursue these other examples here. However, we are faced with the question of what problem this family of trees corresponds to.

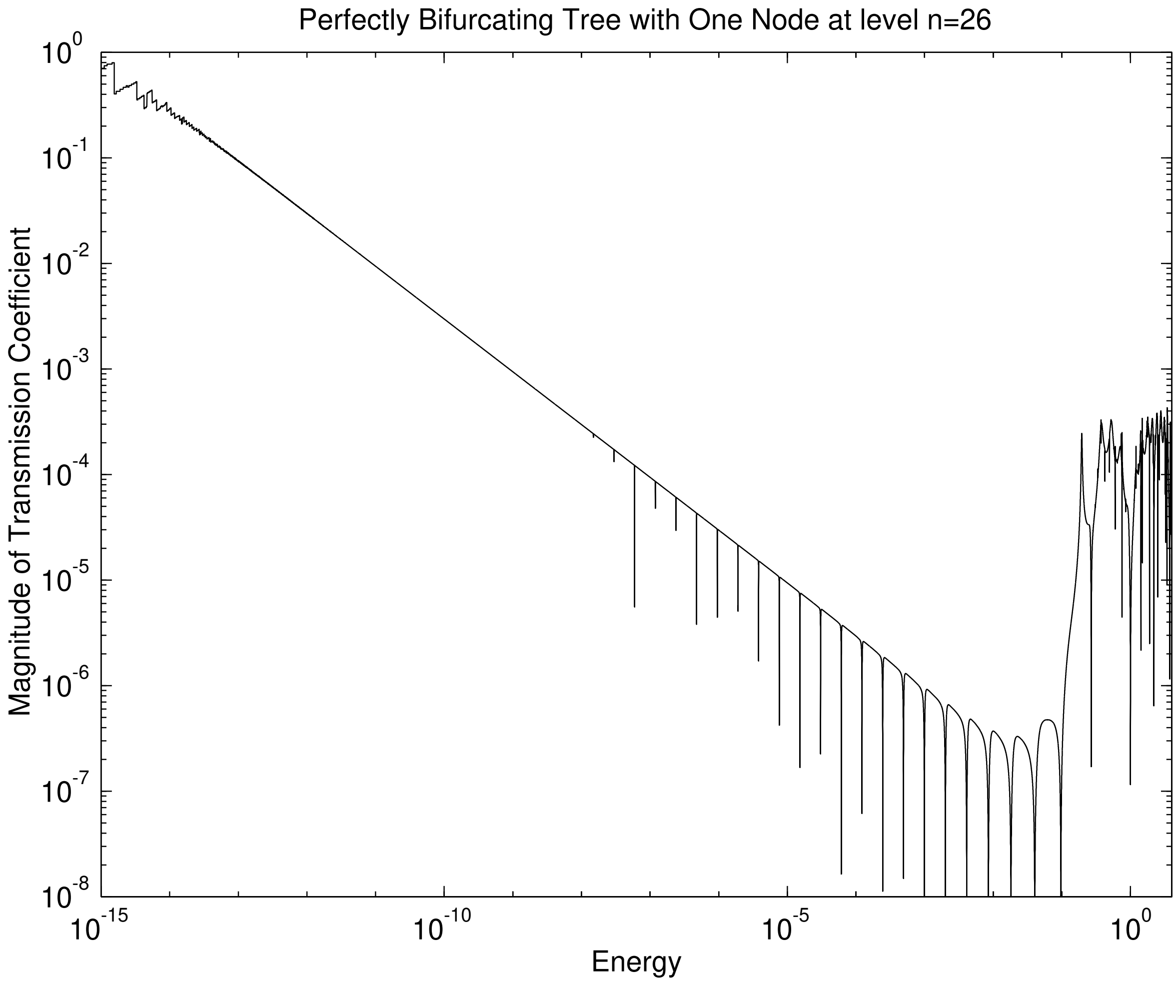

Figure 9: The magnitude of T versus E for E between 0 and 4 for the perfectly bifurcating tree with one node at the n^{th} level.

We can think of decision trees as associated with functions that impose constraints. At each level i there is a function f_i that depends on the first i bits. If $f_i(x_1 \cdots x_i) = 1$ then the i^{th}-level node $x_1 \cdots x_i$ is connected to the $(i-1)^{\text{th}}$-level node $x_1 \cdots x_{i-1}$. (The 0^{th}-level node needs no bits to describe it.) If $f_i(x_1 \cdots x_i) = 0$ then $x_1 \cdots x_i$ is absent from the tree. In terms of the functions f_i, the decision question is, "Is there an $x_1 \cdots x_n$ such that $f_i(x_1 \cdots x_i) = 1$ for all $i = 1$ to n?"

For the tree depicted in Fig. 5, the functions $f_1, \ldots, f_{n-1}$ are all identically 1. This gives the perfectly bifurcating structure. Then there is a function $f_n(x_1 \cdots x_n)$ that is guaranteed to be 0 for all but one of the 2^n values of $x_1 \cdots x_n$. At one special, but unknown, value f_n is either 0 or 1. (We draw the decision tree assuming there is a value for which f_n equals 1. Otherwise the transmission coefficient is 0 and there is nothing to calculate.) Without further information about f_n, any classical algorithm will need to search 2^n values of $x_1 \cdots x_n$ to see

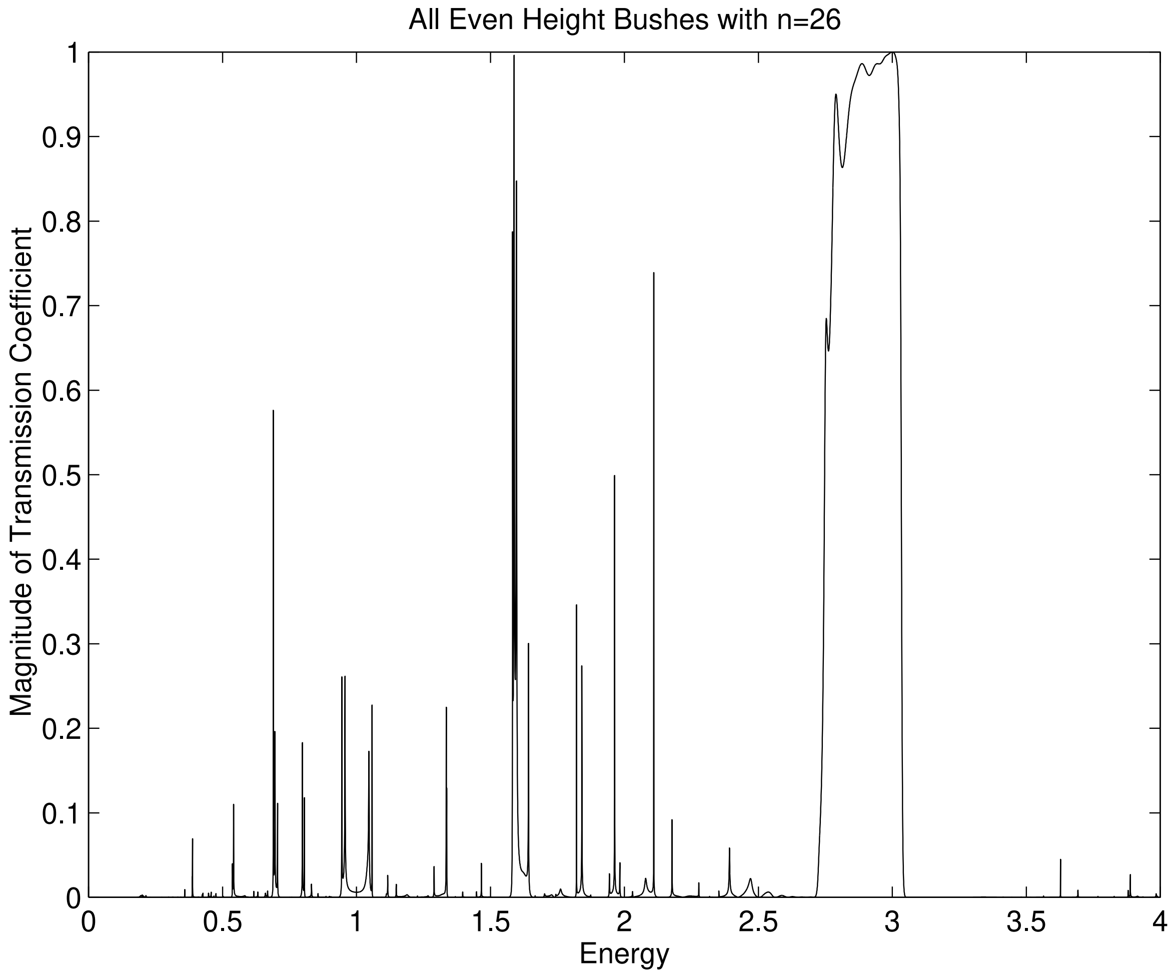

Figure 10: The magnitude of T versus E for the same tree used in Fig. 9 after removing one layer of nodes from each odd-length bush.

if there is a value at which f_n equals 1.

Let us turn to the functions that determine the quantum penetrable tree just discussed. At the n^{th} level there is the function $f_n(x_1 \cdots x_n)$ which may take the value 1 on one input, say $w_1 \cdots w_n$. To arrange for the bushes to all have even height, the tree must be trimmed at level $n-1$. For n even, the function $f_{n-1}(x_1 \cdots x_{n-1})$ is 0 if $x_1 \neq w_1$ or if $x_1 = w_1$, $x_2 = w_2$, and $x_3 \neq w_3$ or if $x_1 = w_1$, $x_2 = w_2$, $x_3 = w_3$, $x_4 = w_4$, and $x_5 \neq w_5$, etc. If we are allowed to call the function $f_{n-1}(x_1 \cdots x_{n-1})$, which we know has this much structure, we can determine (thanks to M. Sipser) $w_1 \cdots w_{n-1}$ with far fewer than order 2^n function calls. First try various inputs until you find an example $x_1 \cdots x_{n-1}$ such that f_{n-1} is 1 on this input. Then you know that $w_1 = x_1$. Trying inputs of the form $w_1 x_2 \cdots x_{n-1}$ will allow you to find w_2, etc. Once $w_1 \cdots w_{n-1}$ is determined, two function evaluations of $f_n(w_1 \cdots w_{n-1} x_n)$ with $x_n = 0, 1$ will answer the decision question. Of course what is occurring here is that the

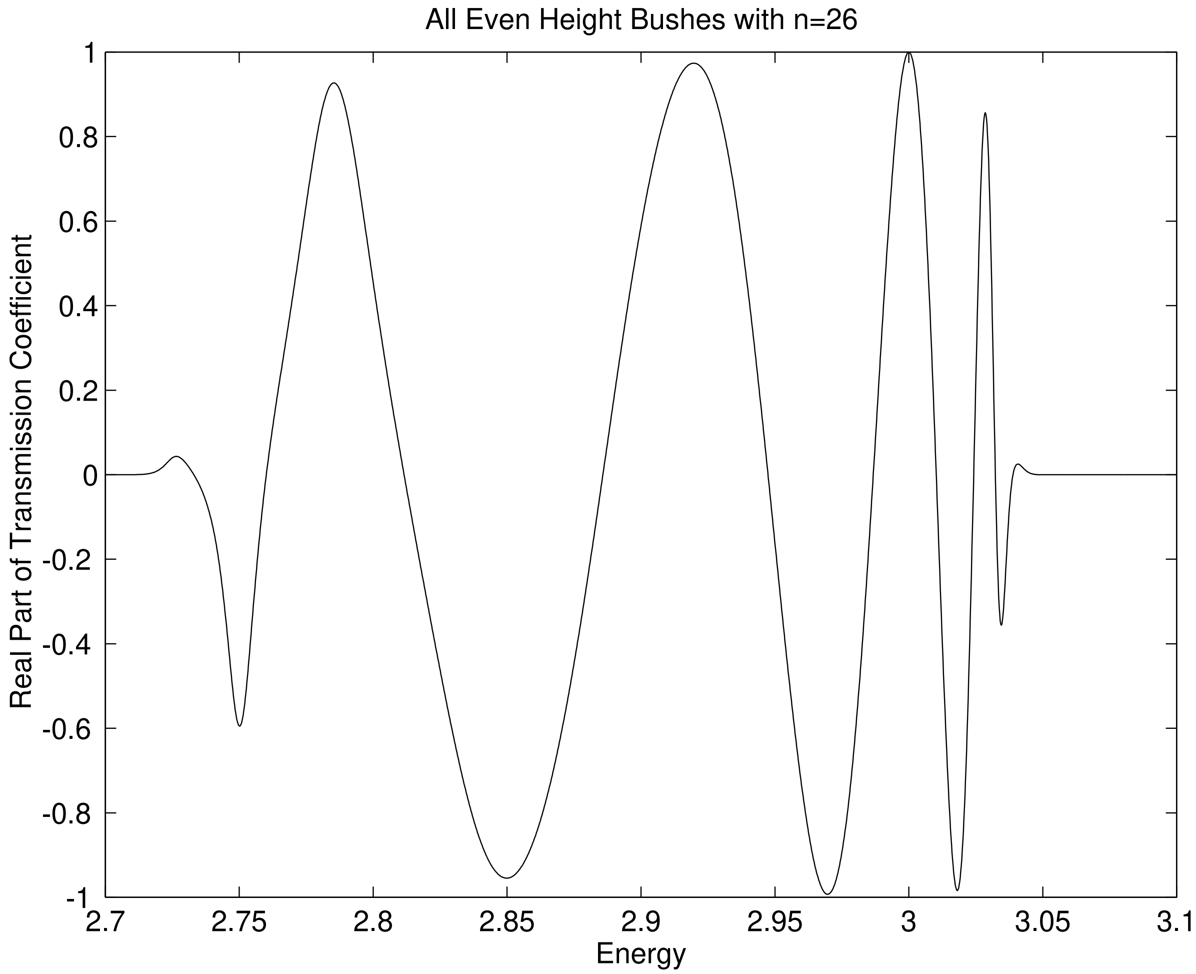

Figure 11: The real part of T versus E showing that T does not oscillate rapidly about zero close to where T is 1, for the same tree as Fig. 10.

extreme regularity of the tree, which guarantees its quantum penetrability, is also structuring the decision problem so that it can be answered much more efficiently than by a classical random walk, which is incapable of seeing larger structures.

4 Implementing the Quantum System

In this section, we show how to implement on a conventional quantum computer the quantum systems previously described. A conventional quantum computer consists of ℓ spin 1/2 particles that give rise to a 2^ℓ dimensional complex Hilbert space with basis elements $|z_1 z_2 \cdots z_\ell\rangle$ where we take each z_i to be 0 or 1. The computer program can be thought of as a sequence of unitary operators $\hat{U}_\alpha$ each of which acts on (at most) B bits. That is, for each $\hat{U}_\alpha$ in the sequence, there is a set $S_\alpha = \{i_1, i_2, \ldots, i_B\}$ that tells us which B bits are being acted on and a 2^B by 2^B unitary matrix whose elements we write as $U_\alpha(w'_1 \cdots w'_B; w_1 \cdots w_B)$. We then have for each $\hat{U}_\alpha$,

$$\langle z'_1 z'_2 \cdots z'_\ell | \hat{U}_\alpha | z_1 z_2 \cdots z_\ell \rangle = \prod_{j \notin S_\alpha} I(z_j = z'_j) U_\alpha(z'_{i_1} \cdots z'_{i_B}; z_{i_1} \cdots z_{i_B}) \quad . \tag{4.1}$$

Here $I(s)$ is the indicator function that is 1 if s is true and 0 if s is false. This formula is just a way of writing that $\hat{U}_\alpha$ acts on B bits.

In previous sections we described evolution through decision trees using the quantum Hamiltonian $\hat{H}$ that gives rise to the unitary time evolution operator $e^{-it\hat{H}}$. To find a sequence of unitary operators each of which acts on only several bits and whose product gives (approximately) the same evolution as $e^{-it\hat{H}}$, we follow the procedure given in [3]. Suppose

$$\hat{H} = \sum_{k=1}^{p} \hat{H}_k \tag{4.2}$$

where for each k, $\hat{H}_k$ and hence $e^{-it\hat{H}_k}$ acts only on (at most) B bits. The Trotter formula says,

$$e^{-it\hat{H}} \approx \left[e^{-it\hat{H}_1/m} \ e^{-it\hat{H}_2/m} \ \cdots \ e^{-it\hat{H}_p/m} \right]^m \tag{4.3}$$

for t/m small. Thus the evolution operator $e^{-it\hat{H}}$ can be approximated as a product of pm unitary operators each of which acts on a fixed number of bits. As a function of n the largest times t that interest us are, say, n^A. Taking $m = n^{2A}$ allows us to obtain $e^{-it\hat{H}}$ with a number of elementary unitary operators that only grows polynomially with n, as long as p also grows only polynomially with n.

We now show two cases where the Hamiltonian $\hat{H}$ given by (2.1) can be written as a sum of $\hat{H}_k$ where each $\hat{H}_k$ acts on a fixed number of bits. Consider first the underlying branching tree, Fig. 1 and its associated $\hat{H}$. Start with $\ell = 2n + 1$ bits that we group for convenience as

$$(yx) = (y_0 y_1 \cdots y_n x_1 \cdots x_n) \quad . \tag{4.4}$$

The y bits indicate the level of the node. The states we use will have a single $y_i = 1$ and the rest 0 to indicate that the node is at level i. The $x_1 \cdots x_i$ will indicate the particular node at the i^{th} level; these nodes will also have $x_{i+1} = x_{i+2} = \cdots = x_n = 0$. We now define the following one bit operators through their action on the basis vectors $|yx\rangle$:

$$\begin{aligned}
\hat{y}_j |yx\rangle &= y_j |yx\rangle \\
\hat{x}_j |yx\rangle &= x_j |yx\rangle \\
\hat{\rho}_j |yx\rangle &= \hat{\rho}_j |y_0 \cdots y_j \cdots y_n x\rangle = \bar{y}_j |y_0 \cdots \bar{y}_j \cdots y_n x\rangle \\
\hat{\sigma}_j |yx\rangle &= \hat{\sigma}_j |y x_1 \cdots x_j \cdots x_n\rangle = \bar{x}_j |y x_1 \cdots \bar{x}_j \cdots x_n\rangle
\end{aligned} \tag{4.5}$$

where $\bar{y}_j = 1 - y_j$ and $\bar{x}_j = 1 - x_j$. We see that $\hat{x}_j$ and $\hat{y}_j$ are diagonal in this basis. The operator $\hat{\rho}_i^\dagger \hat{\rho}_{i+1}$ acting on a state at level i brings it to level $i+1$ whereas $\hat{\rho}_i \hat{\rho}_{i+1}^\dagger$ moves from level $i+1$ to level i.

The Hamiltonian (2.1) defined on the underlying branching tree is

$$\hat{H} = 2\hat{y}_0 + 3\sum_{i=1}^{n-1} \hat{y}_i + \hat{y}_n - \sum_{i=0}^{n-1} (\hat{\rho}_i^\dagger \hat{\rho}_{i+1} + \hat{\rho}_i \hat{\rho}_{i+1}^\dagger)(1 - \hat{x}_{i+1}) \\ - \sum_{i=0}^{n-1} (\hat{\rho}_i^\dagger \hat{\rho}_{i+1} \hat{\sigma}_{i+1} + \hat{\rho}_i \hat{\rho}_{i+1}^\dagger \hat{\sigma}_{i+1}^\dagger) \quad . \tag{4.6}$$

The first three terms give the diagonal matrix elements. The fourth term connects the nodes $x_1 \cdots x_i$ at level i with the nodes $x_1 \cdots x_i 0$ at level $i+1$ whereas the last term connects $x_1 \cdots x_i$ at level i with $x_1 \cdots x_i 1$ at level $i+1$. Thus we see that $\hat{H}$ can be written as a sum of $\hat{H}_k$ each of which acts on at most three bits.

We have built a Hilbert space with 2^{2n+1} states whereas the underlying branching tree has only $2^{n+1} - 1$ nodes. However, if we start in the state corresponding to the topmost node, that is, $y_0 = 1$ and all other bits 0, then if we act with $e^{-i\hat{H}t}$ with $\hat{H}$ given by (4.6) we only ever reach states in the subspace corresponding to the underlying branching tree. The 2^{2n+1}-dimensional Hilbert space may not be the most economical choice to describe the tree but it suffices for our purpose of showing that $\hat{H}$ can be built as a sum of local Hamiltonians.

Of course we also want to construct $\hat{H}$ as a sum of Hamiltonians acting on a fixed number of bits for interesting trimmed decision trees. There are families of trimmed trees whose Hamiltonians we cannot represent in this way. But for many interesting problems we can write $\hat{H}$ as a sum of Hamiltonians that act on at most B bits, where B does not grow with n. For example, we now show how to do this for a version of the exact cover problem discussed in the introduction. We restrict the matrix A, which defines an instance of the exact cover problem, to have exactly three 1's in any row and three or fewer 1's in any column. Even with this restriction, the problem is NP-complete.

Consider first the question of whether the i^{th}-level node $x_1 \cdots x_i$ connects to the $(i+1)^{\text{th}}$-level node $x_1 \cdots x_i 1$. We assume that $x_1 \cdots x_i$ is in the tree and we need to be consistent with (1.1) so we know that for each j, $\sum_{k=1}^{i} A_{jk} x_k$ is 0 or 1. If for some j this sum is 1 and also $A_{j,i+1} = 1$, then $x_1 \cdots x_i 1$ is eliminated as a node. Consider the function

$$C_i^1(x_1 \cdots x_i) = \prod_{j=1}^{m} \Big\{ \big[1 - \sum_{k=1}^{i} A_{jk} x_k\big] A_{j,i+1} + \big[1 - A_{j,i+1}\big] \Big\} \quad . \tag{4.7}$$

Given that $x_1 \cdots x_i$ is an allowed node, then this function is 1 if $x_1 \cdots x_i 1$ is allowed and 0 if $x_1 \cdots x_i 1$ is excluded. Furthermore, given the restriction that A has three 1's in any row and three or fewer in any column, C_i^1 has at most six x_k's appearing.

Now we ask if $x_1 \cdots x_i$ at level i connects to $x_1 \cdots x_i 0$ at level $i+1$. This connection will be allowed unless for some j with $A_{j,i+1} = 1$, there is a $k \le i$ and a distinct $k' \le i$ such that

$A_{jk} = A_{jk'} = 1$ and $x_k = x_{k'} = 0$. The reason the node $x_1 \cdots x_i 0$ would be eliminated in this case is that there are exactly three 1's in any row and (1.1) could not be satisfied if the three bits x_k, $x_{k'}$, and x_{i+1} are all 0. Now consider the function

$$d_i^j(x_1 \cdots x_i) = \sum_{k=1}^{i} A_{jk}(1 - x_k) \quad . \tag{4.8}$$

For any j with $A_{j,i+1} = 1$, d_i^j can be 0, 1, or 2. Only if $d_i^j(x_1 \cdots x_i) = 2$ is $x_1 \cdots x_i 0$ eliminated. Let

$$C_i^0(x_1 \cdots x_i) = \prod_{j=1}^{m} \left\{ \left[\tfrac{1}{2} d_i^j(1 - d_i^j) + 1\right] A_{j,i+1} + (1 - A_{j,i+1}) \right\} \quad . \tag{4.9}$$

Then this function is 0 if $x_1 \cdots x_i 0$ excluded and it is 1 if $x_1 \cdots x_i 0$ is allowed. Again because of the restrictions placed on A, this function has only six x_k's appearing.

The functions C_i^0 and C_i^1 can be promoted to operators simply by replacing their arguments by the bit operators $\hat{x}_k$ defined in (4.5), that is, we have $C_i^0(\hat{x}_1 \cdots \hat{x}_i)$ and $C_i^1(\hat{x}_1 \cdots \hat{x}_i)$. If we multiply the last term in (4.6) by C_i^1 and the fourth term by C_i^0, the Hamiltonian has off-diagonal elements only where the tree has connections. Similarly we can write the diagonal term as

$$\hat{H}_{\text{diagonal}} = 2\hat{y}_0 + \sum_{i=1}^{n-1} \hat{y}_i(1 + C_i^0 + C_i^1) + \hat{y}_n \quad . \tag{4.10}$$

Thus we have written the Hamiltonian for the trees trimmed by A in the form (4.2) with $B = 9$.

Generally, we think of decision trees as associated with functions f_i that impose constraints: $f_i(x_1 \cdots x_i) = 1$ if the $(i-1)^{\text{th}}$ level node $x_1 \cdots x_{i-1}$ is connected to the i^{th} level node $x_1 \cdots x_i$; otherwise $f_i = 0$. The exact cover example above makes clear that as long as there is a fixed B such that $f_i(x_1 \cdots x_i)$ depends on only B bits for each i (*which* bits can vary with i, of course) then the problem can be implemented within the usual quantum computing paradigm – we only need to replace $C_{i-1}^x(\hat{x}_1 \cdots \hat{x}_{i-1})$ in (4.10) by $f_i(\hat{x}_1 \cdots \hat{x}_{i-1}, x)$ and also to multiply the appropriate connection terms in (4.6) by $f_i(\hat{x}_1 \cdots \hat{x}_{i-1}, x)$.

Note that our example in Section 3 for which the quantum algorithm achieved exponential speed-up does not meet this fixed-B requirement. We do have, however, similar examples that achieve exponential speed-up and that do meet this requirement. These problems also rely on even-length, very structured bushes, and also can be solved quickly by other classical algorithms.

5 Conclusions

There is great interest in devising quantum algorithms that improve on classical algorithms, and there have been some notable successes. For example, the well-known Shor [6] and

Grover [5] algorithms demonstrate remarkable ingenuity. Each uses quantum interference, the necessary ingredient for quantum speed-up, in what appears to be a problem-specific way. So far these methods have not been successfully applied to problems very different from the ones for which they were originally devised.

In this paper, we have considered a single time-dependent Hamiltonian that evolves a quantum state through the nodes of a decision tree. This is in contrast to the usual setup consisting of a sequence of unitary operators each acting on a fixed number of bits. (For many problems, including NP-complete ones, our algorithm can be rewritten in the conventional language of quantum computation.) Studying Hamiltonian evolution on decision trees is facilitated by the technique of calculating energy-dependent transmission coefficients. The example in Section Three shows explicitly how interference allows a class of trees to be penetrated exponentially faster by quantum evolution than by classical random walk. However, this example can be quickly solved by a different classical algorithm.

The particular Hamiltonian we chose allowed us to prove, in Section Two, that the quantum algorithm succeeded in polynomial time whenever the corresponding classical random walk on the decision trees succeeded in polynomial time. In searching for more examples where the quantum algorithm outperforms the classical algorithm, one is not restricted to this Hamiltonian. We can imagine trying any Hamiltonian with nonzero off-diagonal elements where there are links between the nodes on the decision tree. With this flexibility, we hope that the class of trees that can be penetrated quickly by the quantum algorithm is large enough to include classically difficult problems.

Acknowledgment

We thank Francis Low and Mike Sipser for their help and insights. We also thank Rachel Cohen and Cindy Lewis for LaTeX assistance, and Martin Stock for creating Figs. 1–9 and final formatting.

References

[1] "Computational Complexity", D.S. Johnson and C.H. Papadimitriou, p. 37 of *The Traveling Salesman Problem*, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, eds. John Wiley & Sons, 1985.

[2] A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J.A. Smolin, and H. Weinfurter, Phys. Rev. **A52**, 3457 (1995) and references therein.

[3] S. Lloyd, Science **273**, 1073 (1996).

[4] This is known as Gerschgorin's theorem; c.f. C.G. Cullen, *Matrices and Linear Transformations*, 2^{nd} edition, Addison-Wesley (1972), p. 283.

[5] L.K. Grover, Proceedings, 28^{th} Annual ACM Symposium on the Theory of Computing (STOC) 1996, pp. 212–218 and quant-ph/9605043.

[6] P.W. Shor, "Polynomial Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer", quant-ph/9508027