CS 880: Quantum Information Processing 10/18/10

Lecture 17: Classical vs. Quantum Hitting Time

Instructor: Dieter van Melkebeek Scribe: Cong Han Lim

In the previous lecture, we introduced the idea of using quantum walks in search algorithms to obtain quadratic speedups over classical random walks. This was done by using a theorem that related the spectral gap of a transition matrix with the classical and quantum hitting times:

Theorem 1. *Let G be a regular graph with spectral gap δ. Let ϵ represent the fraction of good vertices. The expected number of steps to first reach a good vertex is $O(\frac{1}{\delta\epsilon})$ in the classical random walk, and $O(\frac{1}{\sqrt{\delta\epsilon}})$ in the quantum random walk (based on Grover's diffusion operator).*

Here we will prove a weaker form of the theorem. For a proof of Theorem 1, refer to [4].

1 Main Theorem and Proof

The version of the theorem we will prove in this lecture will allow us to distinguish between the cases where our graph G has either no good vertices or at least an ϵ-fraction of good vertices. We obtain the proof of this by performing a simpler analysis on the same algorithm used in the full version.

This weaker formulation still allows us to obtain certain desired results. For example, in the element distinctness problem, we can still detect whether we have duplicate elements but we cannot return an offending pair.

1.1 Set-up for Analysis

Let A be the matrix describing the random walk as a Markov process where, without loss of generality, the vertices are ordered such that all the bad vertices occur before the good ones. Since we want to compute the expected time to first hit a good vertex, we will consider the variation where we never leave a good vertex once we reach it.

Let A' be the matrix obtained by setting the transitions for the good vertices to be the identity transition. We can write this matrix as

$$\begin{pmatrix} A_{BB} & 0 \\ A_{GB} & I \end{pmatrix},$$

where A_{BB} denotes the transitions between the bad states and A_{GB} denotes the transitions from the bad to the good states. Given the transition matrix A', we will never leave from a good vertex. (In the literature, A' is known as the transition matrix for an *absorbing* Markov chain, and the bad and good vertices are known as *leaking* and *absorbing* states respectively.)

1.2 The Classical Random Walk

We begin the classical random walk with a uniform distribution over the states. After taking one step, if we hit a good vertex we are done. Otherwise, by conditional probability, we are left with a uniform distribution over the bad vertices. From this position we analyze the probability of success (ie. hitting a good vertex) within k steps:

$$\begin{aligned}
\Pr[\text{we do not succeed in the next } k \text{ steps}] &= \frac{1}{|B|}\sum_{v,w\in B}\left((A')^k\right)_{vw} \\
&= \left|\left(\frac{1}{\sqrt{B}}\cdot 1_B\right)^T (A')^k \left(\frac{1}{\sqrt{B}}\cdot 1_B\right)^T\right| \\
&\leq \left\|(A_{BB})^k\right\| \\
&\leq \left\|A_{BB}\right\|^k .
\end{aligned}$$

This gives us

$$\begin{aligned}
\text{E}[\text{number of steps needed to first success}] &= \sum_{k=0}^{\infty}\Pr[\text{first success occurs after the } k\text{-th step}] \\
&\leq \sum_{k=0}^{\infty}\left\|A_{BB}\right\|^k \\
&= \frac{1}{1-\|A_{BB}\|}. \qquad (1)
\end{aligned}$$

Lemma 1. $\|A_{BB}\| \leq 1-\frac{\delta\epsilon}{2}$

Proof. (Outline) Consider a vector x and the norm $\|A_{BB}x\|$. We can represent this norm as $\|A\tilde{x}\|^2$, where the vector $\tilde{x}$ is simply the vector x padded with zeroes for coordinates corresponding to the good vertices.

By the spectral theorem, we can decompose $\tilde{x}$ into two components - the component parallel to the uniform distribution (all-ones vector which has eigenvalue 1) and component consisting of linear combinations of eigenvectors perpendicular to the uniform distribution. The spectral gap gives a lower bound on the reduction in the perpendicular component each time we apply the matrix A, and repeated applications of A would reduce the parallel component as certain states get weeded out. The analysis here is similar to the analysis done in Lecture 16. □

1.3 Analysis of Quantum Random Walk

Recall that for the quantum random walk, we act on directed edges $|v,w\rangle$ instead of vertices. Every step consists of two stage:

1. Coin flip stage
2. Swap stage

We will begin by analysing the coin flip stage. Note that we can write the coin flip (Grover diffusion) operator on vertex v as

$$C_v = \frac{1}{\sqrt{|N_v|}}\sum_{(v,w)\in E}|w\rangle$$

where N_v denote the set of neighbors of v. Let π_v denote the projection onto N_v. Then we can write $C_v = 2\pi_v - I$. The general coin flip operator for every vertex can be denoted by

$$C : |v, w\rangle \rightarrow (I \otimes C_v) |v, w\rangle$$

As for the swap stage, we have the swap operator S that gives

$$S |v, w\rangle = |w, v\rangle .$$

By introducing the operator $\pi : |v, w\rangle \rightarrow (I \otimes \pi_v) |v, w\rangle = |v, \pi_v w\rangle$, we can combine the two stages as a single operator:

$$U = SC = S(2\pi - I).$$

We want to make an observation after every application of U to check if we have hit a good vertex. Directly observing the register describing our superposition over the edges will cause the states will collapse, reducing our algorithm to the classical random walk. This problem can be avoided by using an oracle query gate with an ancilla qubit to store the query result. If we observe a 1 in the ancilla qubit, we cause our superposition to collapse to a superposition over only the good vertices, and the opposite happens when we observe a 0.

We will now describe the quantum walk algorithm. We initialize the qubits to represent a uniform superposition over the graph. Let $|\psi_v\rangle = \frac{1}{\sqrt{N_v}} \sum_w A_{wv} |w\rangle$. Our initial state is

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_v |v, \psi_v\rangle .$$

We first make an oracle query over this superposition and observe the ancilla qubit. If we observe a 1, we are done. Otherwise, we are left with a uniform superposition over the bad vertices:

$$|\psi'\rangle = \frac{1}{\sqrt{B}} \sum_{v \in B} |v, \psi_v\rangle .$$

We now consider the separate cases where our graph either has no good vertices or an ϵ-fraction of them. If our graph has no good vertices, note that applying U has no effect on $|\psi'\rangle$, so the resulting vector always corrresponds to the eigenvalue 1. If our graph has at least one good vertex, then $|\psi'\rangle$ will have no component corresponding to the eigenvalue 1. This allows us to distinguish between the two cases by applying eigenvalue estimation on U.

When we compute the phase of the eigenvalues of U, in the case with no good vertices we will obtain a phase of zero, whereas in the other case we will obtain a non-zero phase. Hence, we can apply phase estimation sufficiently precise (depending on the spectral gap) to differentiate between the cases. In order to perform this process, it is enough to have a non-zero lower bound for the eigenvalues of U. The following lemma gives us a relation between the eigenvalues of U and A_{BB}.

Lemma 2. *The eigenvalues of U other than ± 1 are given by*

$$\lambda \pm i\sqrt{1 - \lambda^2} = e^{\pm \arccos \lambda}$$

where λ is an eigenvalue of the discriminant matrix D given by $D_{vw} = \sqrt{(A')_{vw} \cdot (A')_{wv}}$

We will defer our proof of the lemma to the end of the section.

The matrix D can be viewed as a symmetrized version of A', and note that it has the form

$$D = \begin{pmatrix} A_{BB} & 0 \\ 0 & I \end{pmatrix}.$$

The eigenvectors that have eigenvalue $\neq \pm 1$ have to correspond to the A_{BB} portion of the matrix. By applying the lemma we obtain

$$\text{Phase of any eigenvalue of } U \geq |\arccos \|A_{BB}\| \, |$$

and by Taylor expansion, we know that

$$\begin{aligned} \cos x &\geq 1 - \frac{x^2}{2} \\ x &\geq \sqrt{2}\sqrt{1 - \cos x}, \end{aligned}$$

giving us

$$\begin{aligned} \text{Phase of any eigenvalue of } U &\geq \sqrt{2}\sqrt{\|A_{BB}\|} \\ &\geq \sqrt{\delta\epsilon}. \end{aligned}$$

Hence, when we perform eigenvalue estimation, in the ϵ-fraction case it suffices to observe a phase between 0 and $\sqrt{\delta\epsilon}$. The number of operations we need to perform are $O\left(\frac{1}{\sqrt{\delta\epsilon}}\right)$, as desired.

While this might seem like a roundabout process to obtain the result, it is actually a trick that allows us to obtain the weaker theorem in a fairly simple manner. The proof for the full version is significantly more involved.

1.4 Proof Outline of Lemma 2

One thing to note about the operator U is that it acts on two registers. Let T be the operator such that $T\,|v\rangle = |v, \psi_v\rangle$.

Exercise 1. *Prove that T satisfies*

1. $TT^\dagger = \pi$,
2. $T^\dagger T = I$,
3. $T^\dagger S T = D$.

Let $|v\rangle$ be an eigenvector of D and let λ be the corresponding eigenvalue. We can "extend" $|v\rangle$ to a two register version by letting $|v'\rangle = T\,|v\rangle$. We will show that the eigenvectors of U that do not have eigenvalues ± 1 are given by linear combinations of $|v'\rangle$ and $S\,|v'\rangle$.

Note that we have

$$\begin{aligned} U\,|v'\rangle &= S\,|v'\rangle \\ US\,|v'\rangle &= 2\lambda S\,|v'\rangle - |v'\rangle \end{aligned}$$

so the space spanned by the $|v'\rangle$ and $S|v'\rangle$ is invariant under the operator U. This allows us to break up the problem into two parts: considering eigenvectors contained in this space and those that have a component in the orthogonal space.

We first consider the case of an eigenvector $|w\rangle$ of U, which we can assume to be in the form $|w\rangle = |v'\rangle - \mu S|v'\rangle$ for some μ. Applying the operator U, we obtain

$$\begin{aligned} U|w\rangle &= U\left|v'\right\rangle - \mu US\left|v'\right\rangle \\ &= S\left|v'\right\rangle - \mu(2\lambda S\left|v'\right\rangle - \left|v'\right\rangle) \\ &= \mu\left|v'\right\rangle + (1-2\mu\lambda)S\left|v'\right\rangle. \end{aligned}$$

Observing the coefficients, we know that $-\mu^2 = 1 - 2\mu\lambda$. So

$$\mu^2 - 2\lambda\mu + 1 = 0 \Rightarrow \mu = \lambda \pm i\sqrt{1-\lambda^2}$$

which is our desired result. Solving for μ, this gives us two eigenvectors for each choice of eigenvector v of D. This is ideal, since we have doubled the dimension in considering U instead of D and we are doubling the number of eigenvectors accordingly.

As for eigenvectors that have a component in the orthogonal space, note that U simply peforms $-S$ on the orthogonal component, so these have eigenvalues ± 1.

2 General Framework

This brings us to the general framework for using quantum walks. There are three main steps:

1. Set-up
2. Update
3. Check

The set-up is performed once at the start, and we loop the update and check process until we hit a good vertex. Let S, U, C denote the respective costs of these processes. The overall cost of such a quantum walk algorithm would be $S + O\left(\frac{1}{\sqrt{\delta\epsilon}}(U+C)\right)$.

One could reduce the overall cost of the algorithm by notiing that the checking cost C tends to be significantly higher than the update cost U. We could check less frequently, the trade-off being lowering the probability of hitting a good vertex. The optimal algorithm one could obtain has cost $S + O\left(\frac{1}{\sqrt{\epsilon}}\left(\frac{1}{\sqrt{\delta}}U + C\right)\right)$. For a full description of how to obtain refer to [3].

We will briefly describe how the element distinctness problem fits in this framework. For the quantum walk algorithm for this problem, we were performing a walk on the Johnson graph $J_{n,m}$ where each vertex represented a subset of m elements out of the original of n elements and two vertices were adjacent if and only if the subsets they represent shared exactly $m-1$ elements. Each step of the random walk can be seen as replacing an element in our chosen subset. Our goal is to hit a vertex representing a subset containing a collision. In this framework we can describe the associated costs as:

1. S: cost of collecting the initial m vertices
2. U: cost of replacing an element in the subset
3. C: cost of checking the subset to see if there are any collisions.

3 Further Reading

The paper by Szegedy [4] presents similar material in a slightly different form. Magniez et al. [3] elaborates on the general framework and discusses how to optimize the algorithm with regards to the update and checking processes. The papers by Ambainis [1] and Kempe [2] provide an introductory overview to quantum walks in general.

References

[1] A. Ambainis. Quantum walks and their algorithmic applications. *arXiv:quant-ph/0403120v3*, 2004.

[2] J. Kempe. Quantum random walks - an introductory overview. *Contemporary Physics*, Vol. 44, pages 307-327, 2004.

[3] F. Magniez, A. Nayak, J. Roland, M. Santha. Search via Quantum Walks. *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 575-584, 2007.

[4] M. Szegedy. Quantum speed-up of Markov chain based algorithms. *Proceedings of the 45th Symposium on Foundations of Computer Science*, pages 32-41, 2004.