The General Adversary Bound: A Survey

Lily Li Morgan Shirley

January 10, 2020

Contents

Abstract

A decade ago, Ben Reichardt showed that the general adversary bound of a function characterizes its quantum query complexity, a result which spanned several papers and drew inspiration from several more. This survey seeks to aggregate the background and definitions necessary to understand the proof.

Notable among these are the lower bound proof and the definition of span programs, witness size, and semi-definite programs. These definitions, in addition to examples and a detailed exposition, serve to give the reader a better intuition of the graph-theoretic nature of the upper bound. We also include an applications of this result to lower bounds on DeMorgan formula size.

1 Introduction

Given a function f, the quantum query complexity of f, denoted $Q(f)$ is the number of quantum oracle queries necessary to evaluate f. It is typically used as a lower bound on the complexity of a quantum algorithm as the amount of computation allowed between queries is unbounded. The *polynomial method* [Bea+01] and the *adversary bound method* [Amb02; BSS03] are the primary techniques used to show lower bounds on quantum query complexity. However, these techniques are currently incomparable. On the n-input collision problem, the adversary method only achieves an $O(1)$ while the polynomial method achieves the optimal $\Theta(n^{1/3})$ bound. On Ambainis' total function f^k on 4^k bits, the polynomial method achieves at most a 2^k lower bound which is strictly weaker than the adversarial bound of 2.5^k [Amb06].

The adversary bound was originally proposed by Ambainis [Amb02]. Given a boolean function f, the adversary bound of f, denoted $\mathsf{Adv}(f)$, captures the intuition that, in order to compute f, one must be able to distinguish between any two inputs w and x where $f(w) \neq f(x)$. Specifically, if $|\phi_w\rangle$ and $|\phi_x\rangle$ are the final state of a quantum query algorithm after running with inputs w and x respectively, then $|\phi_w\rangle$ and $|\phi_x\rangle$ must be far apart in our measurement basis if $f(w) \neq f(x)$. There are several equivalent formulations of the bound. This survey uses the spectral norm formulation of Barnum, Saks, and Szegedy [BSS03].

Definition 1.1. *An* adversary matrix *for $f : \{0,1\}^n \to \{0,1\}$ is a 2^n-by-2^n Hermitian matrix Γ where $\langle x|\Gamma|y\rangle = 0$ whenever $f(x) = f(y)$.*

Definition 1.2. *The matrix D_i is the 2^n-by-2^n matrix where $\langle x|D_i|y\rangle = 0$ if $x_i = y_i$ and $\langle x|D_i|y\rangle = 1$ if $x_i \neq y_i$.*

Definition 1.3. *The* adversary bound *on a function $f : \{0,1\}^n \to \{0,1\}$ is*

$$\mathsf{Adv}(f) = \max_{\substack{\Gamma \geq 0 \\ \Gamma \neq 0}} \frac{\|\Gamma\|}{\max_i \|\Gamma \circ D_i\|}$$

where Γ is an adversary matrix for f.

Theorem 1.4 ([BSS03])**.** $\mathsf{Adv}(f) = \Omega(Q_2(f))$.

Furthermore, Laplante, Lee, and Szegedy show that the adversary bound of a function is a lower bound on the square root of the function's De Morgan formula size.

Theorem 1.5 ([LLS06])**.** $\mathsf{Adv}(f) \leq \sqrt{\mathcal{L}_f}$

Høyer, Lee, and Špalek [HLS07] removed the non-negativity requirement from the adversary bound and showed that it remained a lower bound on quantum query complexity and formula size. In fact this generalization only strengthened the lower bound – indeed, it is a tight bound on quantum query complexity, although this was not shown until later by Reichardt [Rei11].

Definition 1.6. *The* general adversary bound *on a function $f : \{0,1\}^n \to \{0,1\}$ is*

$$\mathsf{Adv}^{\pm}(f) = \max_{\Gamma \neq 0} \frac{\|\Gamma\|}{\max_i \|\Gamma \circ D_i\|}$$

where Γ is an adversary matrix for f.

Unlike the matching lower bound, which has a relatively simple proof, this upper bound is proved using deceptively simple algorithms with complicated analyses. Key to this analysis is the span program model of computation [KW93]. Reichardt's main result is the following:

Theorem 1.7. (General Adversary Bound Characterize Quantum Query Complexity.) *For any n-ary boolean function f,*

$$Q(f) = \Theta(\mathsf{Adv}^{\pm}(f)).$$

Reichardt's result also been used more recently in entirely classical settings. From the work of Ambainis et. al [Amb+10], any function with de Morgan formula size ℓ has a quantum query algorithm which makes at most $O(\sqrt{\ell})$ queries. In combination with the polynomial method, this implies the existence of a polynomial p with degree $O(\sqrt{\ell})$ such that $p(x)$ approximates f up to a constant factor. More recently, Tal used the result to show a $\tilde{\Omega}(n^2)$ lower bound for the bipartite formula size of the Inner-Product function [Tal17].

2 Preliminaries

Let f be an n-ary boolean function. Then $F_0 = f^{-1}(0)$ and $F_1 = f^{-1}(1)$ are sets of strings which evaluate to 0 and 1 on f respectively.

We assume basic familiarity with quantum computation and bra-ket notation. Given a vector $|v\rangle$, let $\||v\rangle\|$ represent the ℓ_2-norm of $|v\rangle$. Given a matrix M, let $\|M\|$ represent the *spectral norm* of the matrix, defined as $\max_{|u\rangle} \|M|u\rangle\|$ where the maximum is over unit vectors $|u\rangle$. In this survey, we use the fact that $\|M\|$ is the largest singular value of M. Furthermore, let $\|M\|_{\mathrm{Tr}} = \max_B |\langle M, B\rangle| / \|B\|$, where the maximization is over complex matrices with the same dimensions as M.

We say that a matrix $A \in \mathbb{L}(U, V)$ if A is a linear transformation from vectors in $\mathbb{C}^U$ to vectors in $\mathbb{C}^V$. In this case A has $|U|$ columns and $|V|$ rows. Let $\mathbb{L}(U) = \mathbb{L}(U, U)$. $\mathbb{I}_k$ is the $k \times k$ identity matrix. When the dimensions are clear from context, we omit the subscript. For $u \in U$ and $v \in V$, we will use $|u\rangle$ and $|v\rangle$ to denote the indicator vectors for the relevant column and row of A respectively. In particular, $\langle v|A|u\rangle$ is the entry of A in row v and column u.

Readers will also require some familiarity with positive semi-definite matrix (PSD) and semi-definite programs (SDP). If X is PSD, we write $X \succeq 0$. When we write $X \succeq Y$, we mean that $X - Y \succeq 0$.

3 The General Adversary Bound

In which we show the following properties of the general adversary bound.

Theorem 3.1 ([HLS07])**.** $\mathsf{Adv}^{\pm}(f) = \Omega(Q_2(f))$.

Theorem 3.2 ([HLS07])**.** $\mathsf{Adv}^{\pm}(f) \leq \sqrt{\mathcal{L}_f}$

The corresponding upper bound, $\mathsf{Adv}^{\pm}(f) = O(Q_2(f))$, will be left to a later section.

3.1 $\mathsf{Adv}^{\pm}$ is a Lower Bound for Quantum Query Complexity

Consider a quantum query algorithm that computes f in T steps with error at most 1/3. Without loss of generality, the quantum query algorithm is of the form $U_T V_{\mathsf{IND}} U_{T-1} V_{\mathsf{IND}} \dots U_1 V_{\mathsf{IND}} U_0$ where each U_t is a

unitary that does not depend on the input x and V_{IND} is the standard phase oracle unitary on the *index function* IND:

$$V_{\mathsf{IND}}|i\rangle = (-1)^{x_i}|i\rangle$$

It will be helpful to divide the state of this quantum query algorithm into three sets of qubits: (1) the *input set* I holds the input x and remains unchanged throughout the execution of the algorithm, (2) the *query set* Q that is used by each V_{IND} to specify a coordinate of x, and (3) a *workspace set* W that can be acted upon arbitrarily. The qubits in Q and W are measured at the end of the algorithm to obtain an output in $\{0,1\}$. These measurements can be viewed as orthogonal projectors Π_0, Π_1. Let the combined state of Q and W on input x at step t be $|\psi_x^t\rangle$. Then the probability that we will measure outcome b on input x is $\left\|\Pi_b|\psi_x^T\rangle\right\|^2$. Note that for all x we have $\left\|\Pi_{f(x)}|\psi_x^T\rangle\right\|^2 \geq 2/3$. Three other important properties of the projectors are that $\Pi_0 + \Pi_1 = \mathbb{I}$ (the projectors are complete), $\Pi_b^2 = \Pi_b$ (performing a projection twice has no more effect than applying it once), and $\Pi_0\Pi_1 = \Pi_1\Pi_0 = 0$ (the projections are orthogonal).

The main observation that Høyer, Lee, and Špalek [HLS07] use in their proof of Theorem 3.1 is that the combined state of Q and W must be very different when the algorithm is run on x compared to when it is run on y if $f(x) \neq f(y)$: otherwise, any measurement would be unable to distinguish these states with high enough fidelity. We present their argument here.

Proof of Theorem 3.1. Let Γ be an adversary matrix. Note that $\|\Gamma\|$ is the largest absolute value of any eigenvalue of Γ, as Γ is Hermitian. Assume that $\|\Gamma\| = \lambda_1$ where λ_1 is the largest eigenvalue of Γ: this can be done without loss of generality by replacing Γ with $(-1)\Gamma$, which does not affect the value of $\|\Gamma\|$. Let $|D\rangle$ be the unit eigenvector corresponding to λ_1.

Consider running our quantum query algorithm for f with an input in a superposition defined by $|\delta\rangle$: the state of the input qubits will be $\sum_{x\in\{0,1\}^n}\langle x|\delta\rangle|x\rangle$. Then the state of Q and W at step t will be $\sum_{x\in\{0,1\}^n}\langle x|\delta\rangle|\psi_x^t\rangle$. Let $\mathsf{W}^{(t)}$ be the 2^n-by-2^n density matrix defined by $\langle x|\mathsf{W}^{(t)}|y\rangle = \langle x|\delta\rangle\langle\delta|y\rangle\langle\psi_y^t|\psi_x^t\rangle$. We measure the progress of the algorithm by comparing $\mathsf{W}^{(t)}$ to Γ. Define the progress measure $\mathsf{M}^{(t)} = \langle\Gamma, \mathsf{W}^{(t)}\rangle$. To prove the lower bound, it suffices to show that this progress measure changes by an amount bounded above by $2\max_i \|\Gamma \circ D_i\|$ at each step of the algorithm, but must change by at least a constant multiple of $\|\Gamma\|$ over the course of the entire algorithm. The following three claims show this.

Claim 3.3. $\mathsf{M}^{(0)} = \|\Gamma\|$

Proof. Before any executions of the phase oracle, the state cannot depend on the input: for all x and y, $|\psi_x^0\rangle = |\psi_y^0\rangle$, and so $\mathsf{W}^{(0)} = |\delta\rangle\langle\delta|$. Then $\mathsf{M}^{(0)} = \langle\Gamma, |\delta\rangle\langle\delta|\rangle = \mathrm{Tr}(\Gamma^*|\delta\rangle\langle\delta|) = \mathrm{Tr}(\lambda_1|\delta\rangle\langle\delta|) = \lambda_1 \cdot 1 = \|\Gamma\|$. □

Claim 3.4. $\mathsf{M}^{(T)} \leq (\frac{2}{3}\sqrt{2})\,\|\Gamma\|$

Proof. First note that $\Gamma = \Gamma \circ F$ where F is the 0/1 adversary matrix:

$$\langle x|F|y\rangle = \begin{cases} 0 & f(x) = f(y) \\ 1 & f(x) \neq f(y) \end{cases}$$

Thus $\mathsf{M}^{(T)} = \langle\Gamma \circ F, \mathsf{W}^{(T)}\rangle = \langle\Gamma, F \circ \mathsf{W}^{(T)}\rangle$. By the definition of the trace norm, this gives us $\mathsf{M}^{(T)} \leq \|\Gamma\| \left\|F \circ \mathsf{W}^{(T)}\right\|_{\mathrm{Tr}}$. To prove the claim we simply need to upper-bound $\left\|F \circ \mathsf{W}^{(T)}\right\|_{\mathrm{Tr}}$.

Let X_0 (respectively X_1) be the $2^n \times 2^n$ matrix where the xth row is $\Pi_{f(x)}\delta_x|\psi_x^T\rangle$ (respectively $\Pi_{1-f(x)}\delta_x|\psi_x^T\rangle$). Think of X_0 as the projection onto correct answers and X_1 as the projection onto incorrect answers.

First, we observe that $F \circ \mathsf{W}^{(T)} = X_0X_1^* + X_1X_0^*$:

$$\langle x|(X_0X_1^* + X_1X_0^*)|y\rangle = \langle x|\delta\rangle\langle\delta|y\rangle\left(\langle\psi_y^T|\Pi_{1-f(y)}\Pi_{f(x)}|\psi_x^T\rangle + \langle\psi_y^T|\Pi_{1-f(x)}\Pi_{f(y)}|\psi_x^T\rangle\right)$$

If $f(x) = f(y)$, then the expression on the right is 0, as $\Pi_0\Pi_1 = 0$. Otherwise, $\Pi_b\Pi_b = \Pi_b$ and $\Pi_0 + \Pi_1 = \mathbb{I}$, so the expression on the right is $\langle x|\delta\rangle\langle\delta|y\rangle\langle\psi_y^T|\psi_x^T\rangle = \langle x|\mathsf{W}^{(T)}|y\rangle$.

We now need to upper-bound $\|X_0X_1^* + X_1X_0^*\|_{\mathrm{Tr}}$.

$$\begin{aligned}\|X_0X_1^* + X_1X_0^*\|_{\mathrm{Tr}} &\leq \|X_0X_1^*\|_{\mathrm{Tr}} + \|X_1X_0^*\|_{\mathrm{Tr}} && \text{by the triangle inequality}\\ &\leq 2\,\|X_0\|_F\,\|X_1\|_F && \text{by Hölder's Inequality}\end{aligned}$$

We upper-bound this final expression by noting the following two facts:

$$\|X_0\|_F^2 + \|X_1\|_F^2 = \sum_{x\in\{0,1\}^n} |\langle x|\delta\rangle|^2 \left(\left\|\Pi_{f(x)}|\psi_x^T\rangle\right\|^2 + \left\|\mathbb{I} - \Pi_{f(x)}|\psi_x^T\rangle\right\|^2\right) = \|\delta\|^2 = 1$$

$$\|X_0\|_F^2 = \sum_{x\in\{0,1\}^n} |\langle x|\delta\rangle|^2 \left\|\Pi_{f(x)}|\psi_x^T\rangle\right\|^2 \geq \frac{2}{3}\|\delta\|^2 = \frac{2}{3}$$

Therefore, $2\,\|X_0\|_F\,\|X_1\|_F$ is maximized at $2\sqrt{2/3}\sqrt{1/3} = \frac{2}{3}\sqrt{2}$. □

From the first two claims, we know that $\mathsf{M}^{(0)} - \mathsf{M}^{(T)} \geq (1 - \frac{2}{3}\sqrt{2})\,\|\Gamma\|$. The last step in the proof is to give an upper bound on $\mathsf{M}^{(t)} - \mathsf{M}^{(t+1)}$ for all t.

Claim 3.5. $\mathsf{M}^{(t)} - \mathsf{M}^{(t+1)} \leq 2\max_i \|\Gamma \circ D_i\|$

Proof. To help us prove this claim, we will define a new density matrix $\mathsf{W}_\star^{(t)}$ that is similar to $\mathsf{W}^{(t)}$. Whereas $\mathsf{W}^{(t)}$ is indexed by the basis states of the input qubits I and has entries defined by the state of the query qubits Q and the workspace qubits W, $\mathsf{W}_\star^{(t)}$ will be indexed by I and Q and have entries defined by the state of W.

$$\langle x, i|\mathsf{W}_\star^{(t)}|y, j\rangle = \langle x|\delta\rangle\langle\delta|y\rangle\langle\psi_y^t|j\rangle\langle i|\psi_x^t\rangle$$

Here, i and j are basis states of Q.

Let G and D be the following block-diagonal $(n\cdot 2^n)$-by-$(n\cdot 2^n)$ matrices:

$$G = \Gamma \otimes \mathbb{I}_n = \bigoplus_{i=1}^{n} \Gamma \qquad D = \bigoplus_{i=1}^{n} D_i$$

Note that $\mathsf{M}^{(t)} = \langle\Gamma, \mathsf{W}^{(t)}\rangle = \langle G, \mathsf{W}_\star^{(t)}\rangle$. We proceed with the proof. First observe that because the unitary U_{t+1} does not depend on the input qubits, we can ignore it for the purposes of our progress measure: $\langle\psi_y^t|U_{t+1}^*U_{t+1}|\psi_x^t\rangle = \langle\psi_y^t|\psi_x^t\rangle$, and so $\mathsf{W}^{(t)}$ does not change after the application of the unitary. This means that $\mathsf{M}^{(t+1)} = \langle G, \mathsf{W}_\star^{(t+1)}\rangle = \langle G, V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*\rangle$.

$$\mathsf{M}^{(t)} - \mathsf{M}^{(t+1)} = \langle G, \mathsf{W}_\star^{(t)}\rangle - \langle G, V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*\rangle = \langle G, \mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*\rangle = \langle G, (\mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*)\circ D\rangle$$

This last equality is true because $\langle x, i|(\mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*)|y, i\rangle = (1 - (-1)^{x_i+y_i})\langle x, i|\mathsf{W}_\star^{(t)}|y, i\rangle$ (which is 0 when $x_i = y_i$) and G is block-diagonal ($\langle x, i|G|y, j\rangle = 0$ if $i \neq j$).

$$\begin{aligned}\mathsf{M}^{(t)} - \mathsf{M}^{(t+1)} &= \langle G, (\mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*)\circ D\rangle\\ &= \langle G\circ D, (\mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*)\rangle\\ &\leq \|G\circ D\|\cdot\left\|\mathsf{W}_\star^{(t)} - V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*\right\|_{\mathrm{Tr}} && \text{(by the definition of the trace norm)}\\ &\leq \|G\circ D\|\cdot\left(\left\|\mathsf{W}_\star^{(t)}\right\|_{\mathrm{Tr}} + \left\|V_{\mathsf{IND}}\mathsf{W}_\star^{(t)}V_{\mathsf{IND}}^*\right\|_{\mathrm{Tr}}\right) && \text{(by the triangle inequality)}\\ &= 2\,\|G\circ D\|\left\|\mathsf{W}_\star^{(t)}\right\| && (V_{\mathsf{IND}}\text{ is unitary})\\ &= 2\,\|G\circ D\| && (\mathsf{W}_\star^{(t)}\text{ is a density matrix})\\ &= 2\max_i\|\Gamma\circ D_i\|\end{aligned}$$

In the above we used the following facts: the trace norm is invariant under conjugation with a unitary, the trace norm of a density matrix is 0, and the spectral norm of a block-diagonal matrix is the maximum of the spectral norms of the blocks. □

Putting it all together, we have that a quantum query algorithm for f requires at least $\frac{1-\frac{2}{3}\sqrt{2}}{2}\mathsf{Adv}^{\pm}f = \Omega(\mathsf{Adv}^{\pm}f)$ rounds. □

3.2 $\mathsf{Adv}^{\pm}$ is a Lower Bound for the Square Root of Formula Size

The lower bound on $\sqrt{\mathcal{L}(f)}$ using $\mathsf{Adv}^{\pm}(f)$ makes use of the *Karchmer-Wigderson game on* f.

Definition 3.6 ([KW90])**.** *Given a Boolean function f, the* Karchmer-Wigderson game on f ($\mathsf{KW}(f)$) *is a two-player communication game in which one party receives an input $x \in f^{-1}(0)$, one party receives an input $y \in f^{-1}(1)$, and the parties must collectively determine some coordinate i on which $x_i \neq y_i$.*

A useful fact is that the minimum number of leaves in a De Morgan formula that computes a function f – denoted $\mathcal{L}(f)$ – is exactly the minimum number of leaves in a communication protocol that successfully solves $\mathsf{KW}(f)$ – denoted $C^P(\mathsf{KW}(f))$.

Theorem 3.7 ([KW90])**.** $\mathcal{L}(f) = C^P(\mathsf{KW}(f))$

We give a brief sketch of the proof.

Proof (Sketch). Given a formula for f, we can use induction on the depth of the formula to produce a communication protocol for $\mathsf{KW}(f)$. If the formula is a single leaf, then no communication is required (and so the protocol is also a single leaf). If the formula is the logical And of two subformulae, then y must evaluate to 1 on both subformulae but x must evaluate to 0 on at least one, so the player holding x can report which. The parties continue with the protocol for that subformula, so the number of leaves in the communication protocol is (by induction) the sum of the number of leaves in the subformulae, which is just the number of leaves in the entire formula. A similar situation holds when the formula is the logical Or of two subformulae, but with the player holding y speaking.

A communication protocol for $\mathsf{KW}(f)$ can be used to construct a formula for f in an analogous fashion. □

A communication protocol for $\mathsf{KW}(f)$ partitions $f^{-1}(0) \times f^{-1}(1)$ into $C^P(\mathsf{KW}(f))$ combinatorial rectangles, where each rectangle is *monochromatic* in terms of $\mathsf{KW}(f)$: that is, each rectangle is associated with some i where $x_i \neq y_i$ for all (x, y) in the rectangle. Let $C^D(\mathsf{KW}(f))$ be the minimum number of monochromatic combinatorial rectangles required to partition $f^{-1}(0) \times f^{-1}(1)$. Clearly, $C^D(\mathsf{KW}(f)) \leq C^P(\mathsf{KW}(f))$.

In order to prove Theorem 3.2, we will exploit two properties of the spectral norm. The first is that the spectral norm (indeed, any matrix norm) is monotone with respect to submatrices: if A is a submatrix of B, then $\|A\| \leq \|B\|$. The second is that the *square* of the spectral norm is subadditive over rectangles. For a $|X| \times |Y|$ matrix A and a combinatorial rectangle in $X \times Y$, let A_R be defined by:

$$\langle x|A_R|y\rangle = \begin{cases} \langle x|A|y\rangle & (x,y) \in R \\ 0 & \text{otherwise} \end{cases}$$

Lemma 3.8 ([LLS06])**.** *If A is an $|X| \times |Y|$ matrix and $\mathcal{R}$ partitions $X \times Y$ into combinatorial rectangles, then $\|A\|^2 \leq \sum_{R \in \mathcal{R}} \|A_R\|^2$.*

Proof. Note that $\|A\| = \max_{u,v} |\langle u|A|v\rangle| / \||u\rangle\| \||v\rangle\|$. In the following, let $|u\rangle$ and $|v\rangle$ be the unit vectors that achieve the maximum in this expression.

For any $R \in \mathcal{R}$ where $R = X_R \times Y_R$ for $X_R \subseteq X, Y_R \subseteq Y$, define $|u_R\rangle$ and $|v_R\rangle$ as follows:

$$\langle u_R|x\rangle = \begin{cases} \langle u|x\rangle & x \in X_R \\ 0 & \text{otherwise} \end{cases} \qquad \langle v_R|y\rangle = \begin{cases} \langle v|y\rangle & y \in Y_R \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \|A\| = |\langle u|A|v\rangle| &= \left|\langle u| \left(\sum_{R\in\mathcal{R}} A_R\right) |v\rangle\right| = \left|\sum_{R\in\mathcal{R}} \langle u|A_R|v\rangle\right| = \left|\sum_{R\in\mathcal{R}} \langle u_R|A_R|v_R\rangle\right| \\ &\leq \sum_{R\in\mathcal{R}} |\langle u_R|A_R|v_R\rangle| \leq \sum_{R\in\mathcal{R}} \|A_R\| \, ||u_R\rangle| ||v_R\rangle| \\ &\leq \left(\sum_{R\in\mathcal{R}} \|A_R\|^2\right)^{1/2} \left(\sum_{R\in\mathcal{R}} ||u_R\rangle|^2 ||v_R\rangle|^2\right)^{1/2} \qquad \text{(by the Cauchy-Schwarz inequality)} \end{aligned}$$

Note that the second term here simplifies:

$$\begin{aligned} \sum_{R\in\mathcal{R}} ||u_R\rangle|^2 ||v_R\rangle|^2 &= \sum_{R\in\mathcal{R}} \sum_{(x,y)\in R} (\langle u|x\rangle)^2 (\langle v|y\rangle)^2 \\ &= ||u\rangle|^2 ||v\rangle|^2 \quad \text{(as } \mathcal{R} \text{ partitions } X \times Y) \end{aligned}$$

To conclude, note that as $|u\rangle$ and $|v\rangle$ are unit vectors, $||u\rangle|^2 ||v\rangle|^2 = 1$: therefore, $\|A\| \leq \left(\sum_{R\in\mathcal{R}} \|A_R\|^2\right)^{1/2}$ and so $\|A\|^2 \leq \sum_{R\in\mathcal{R}} \|A_R\|^2$. □

Now we can prove Theorem 3.2.

Proof of Theorem 3.2. Let A be any $f^{-1}(0) \times f^{-1}(1)$ matrix. Let $\mathcal{R}_f$ be an optimal rectangle partition in terms of $\mathsf{KW}(f)$.

$$\|A\|^2 \leq \sum_{R\in\mathcal{R}_f} \|A_R\|^2 \leq C^D(\mathsf{KW}(f)) \cdot \max_{R\in\mathcal{R}_f} \|A_R\|^2$$

Let A_i be the $f^{-1}(0) \times f^{-1}(1)$ matrix defined by:

$$\langle x|A_i|y\rangle = \begin{cases} \langle x|A|y\rangle & x_i \neq y_i \\ 0 & \text{otherwise} \end{cases}$$

Then A_R is a submatrix of A_i, so by the monotonicity with respect to rectangles:

$$C^D(\mathsf{KW}(f)) \cdot \max_{R\in\mathcal{R}_f} \|A_R\|^2 \leq C^D(\mathsf{KW}(f)) \cdot \max_{i\in[n]} \|A_i\|^2$$

Rearranging, we get:

$$\mathcal{L}(f) \geq C^D(\mathsf{KW}(f)) \geq \max_{A\neq 0} \frac{\|A\|^2}{\max_i \|A_i\|^2}$$

We conclude by taking the square root of the above expression and noting that for any matrix $A \in f^{-1}(0) \times f^{-1}(1)$, letting A' be the matrix of the form $A' = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$, we have that A' is an adversary matrix for f and $\|A'\| = \|A\|$, so maximizing over matrices A on the right-hand side is equivalent to maximizing over adversary matrices A'. □

4 Span Programs

Given a function f, we will define a span program and a variant known as the canonical span program for f. Using SDP duality, we show that a complexity measure of the canonical span program, known as the witness size, is equivalent to the general adversary bound of f.

Definition 4.1. (Span Programs, [KW93].) *A* span program P_f *for a n-ary boolean function f consists of a matrix $A \in \mathbb{L}(\mathbb{C}^I, \mathbb{C}^{[m]})$ and a target vector $|t\rangle \in \mathbb{C}^{[m]}$, where I is the disjoint union of $2n$ index sets $I_{1,0}$, $I_{1,1}$, ..., $I_{n,0}$, $I_{n,1}$ one for each setting of each entry of a boolean string $s \in \{0,1\}^n$. Given s, $\Pi(s) \in \mathbb{L}(\{0,1\}^I)$ is the diagonal matrix whose diagonal entry $(j,b) \times (j,b)$ indicates $s_j = b$ i.e.*

$$\Pi(s) = \mathbb{I} - \sum_{j \in [n]} |j, \overline{s}_j\rangle\langle j, \overline{s}_j|. \tag{1}$$

The span program P_f evaluates to false if there exists a negative witness $|y\rangle \in \mathbb{C}^{[m]}$ i.e. $\langle y|A\Pi(s) = 0$ but $\langle y|t\rangle > 0$; without loss of generality assume that $\langle y|t\rangle = 1$ by scaling. Conversely, P_f evaluates to true if there exists a positive witness $|z\rangle \in \mathbb{C}^I$ i.e. $|t\rangle$ is in the span of $A\Pi(s)$ and $A\Pi(s)|z\rangle = |t\rangle$.

For inputs which evaluate to true on P_f, i.e. $x \in F_1$ for which there exists $|z\rangle$ such that $A\Pi(x)|z\rangle = |t\rangle$, let $\text{wsize}(P_f, x) = \||z\rangle\|^2$. *For inputs which evaluate to false on P_f, i.e. $w \in F_0$ for which there exists $|y\rangle$ such that $\langle y|A\Pi(w) = 0$ and $\langle y|t\rangle = 1$, let* $\text{wsize}(P_f, w) = \|\langle y|A\|^2$.[1] *The* witness size *of P_f is then*

$$\text{wsize}(P_f) = \max_{s \in \{0,1\}^n} \text{wsize}(P_f, s).$$

Example 4.2. *In the following we consider the span programs for several simple functions. Note that there can be many different span programs for the same function. All omitted column index sets are assumed to be empty.*

1. *For the n-ary logical or function,* OR_n, *let $|t\rangle = [1]$ and*

$$A = \begin{array}{c} \begin{matrix} I_{1,1} & I_{2,1} & \cdots & I_{n-1,1} & I_{n,1} \end{matrix} \\ \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \end{array}.$$

   Observe that $\text{wsize}(P_{\text{OR}_n}) = \max_{s \in \{0,1\}^n} \text{wsize}(P_{\text{OR}_n}, s) = n^2$ *is achieved by the input string $s = [0, ...0]^\top$ with the witness $|y\rangle = [1]$.*

2. *For the parity function $\oplus_2$, let $|t\rangle = [1,1]^\top$ and*

$$A = \begin{array}{c} \begin{matrix} I_{1,0} & I_{1,1} & I_{2,0} & I_{2,1} \end{matrix} \\ \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \end{array}.$$

   Observe that $\text{wsize}(P_{\oplus_2}) = 2$. *This can be achieved by a string which evaluates to false e.g. $w = 00$ with witness $|y\rangle = [0,1]^\top$ and* $\text{wsize}(P_{\oplus_2}, w) = \||y\rangle A\|^2$ *or by a string which evaluates to true e.g. $x = 01$ with witness $|z\rangle = [1,1]^\top$ and* $\text{wsize}(P_{\oplus_2}, w) = \||z\rangle\|^2$.

4.1 Canonical Span Programs

In order to relate the complexity of the span program of a given function f to its query complexity, we put it in *canonical span program* form. Every span program can be transformed into a canonical span program with at most a polynomial blow-up in size [KW93].

[1] Note that this value is equivalent to $\|\langle z|A(\mathbb{I} - \Pi(w))\|$ by the first condition.

Definition 4.3. (Canonical Span Program.) *The input matrix A and target vector $|t\rangle$ of the canonical span program will be as follows. Define $|t\rangle \in \mathbb{C}^{F_0}$ to be a scalar multiple of the all ones vector. Let $A \in \mathcal{L}(\mathbb{C}^I, \mathbb{C}^{F_0})$ where $I = [n] \times \{0,1\} \times [m]$ for a yet-to-be-determined m. Each row of A corresponds to an input w evaluating to zero on f. Divide this row further into $2n$ row vectors of length m one for each setting of each entry in the input. In the following, if $w_j = b$, then denote each length-m vector corresponding to $I_{j,b}(w)$ by $|v'_{w,j}\rangle$ and corresponding to $I_{j,\bar{b}}$ by $|v_{w,j}\rangle$.*

Define $|v'_{w,j}\rangle$ to be the all zeroes vector for all $w \in F_0$ and $j \in [m]$. Observe that $|t\rangle$ cannot be in the span of $A\Pi(w)$ since the row of $A\Pi(w)$ corresponding to w consists entirely of zeros. Further, since the indicator vector $|w\rangle \in \mathbb{C}^{F_0}$ for w is a witness for A,[2]

$$\text{wsize}(P_f, w) = \|\langle w|A\|^2 = \sum_{j\in[n]} \||v_{w,j}\rangle\|^2 .$$

Each $x \in F_1$ will be assign an input vector of length mn. These will not *appear in A, but will be used to ensure that the vectors $|v_{w,j}\rangle$ in A satisfy certain constraints. Each vector will be divided into n length m vectors corresponding to the n entries of x. These will be denoted by $|v_{x,j}\rangle$. Since $|t\rangle$ needs to be in the span of $A\Pi(x)$, we require that for all $w \in F_0$, $\sum_{w_j \neq x_j} \langle v_{w,j}|v_{x,j}\rangle = 1$. Observe that the witness size is again of the form*

$$\text{wsize}(P_f, x) = \sum_{j\in[n]} \||v_{x,j}\rangle\|^2 .$$

The smallest m for which there exists such vectors $|v_{w,j}\rangle$ and $|v_{x,j}\rangle$ will suffice.

In the following let W be the witness size of the canonical span program.

Example 4.4. *The canonical span program for $\oplus_2$ is as follows. Let the target vector be $|t\rangle = c[1,1]^\top$ where $c = 1/(3\sqrt{W})$. Then for $\{w_1 = 00, w_2 = 11\} = F_0$ with vector $\langle v_{w_i,j}| \in \mathbb{C}^{[m]}$ corresponding to the length m vector of the j^{th} bit of w_i, we have*

$$A = \begin{array}{c} \begin{matrix} I_{1,0} & I_{1,1} & I_{2,0} & I_{2,1} \end{matrix} \\ \begin{bmatrix} 0 & \langle v_{w_1,1}| & 0 & \langle v_{w_1,2}| \\ \langle v_{w_2,1}| & 0 & \langle v_{w_2,2}| & 0 \end{bmatrix} \end{array} \begin{matrix} \\ w_1 = 00 \\ w_2 = 11 \end{matrix}$$

Further, to each string x_i in $\{x_1 = 10, x_2 = 01\} = F_1$ we assign a vector $|x_i\rangle = [|v_{x_i,1}\rangle, |v_{x_i,2}\rangle]^\top$ where $|v_{x_i,j}\rangle \in \mathbb{C}^{[m]}$ corresponds to the length m vector of the j^{th} bit of x_i. Note that $m = 1$ suffices, since the matrix A where

$$A = \begin{array}{c} \begin{matrix} I_{1,0} & I_{1,1} & I_{2,0} & I_{2,1} \end{matrix} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \end{array} \begin{matrix} \\ w^{(1)} = 00 \\ w^{(2)} = 11 \end{matrix}$$

and the pair of vectors $|x_1\rangle = |x_2\rangle = [1,1]^\top$ satisfies the condition $\sum_{w_j \neq x_j} \langle v_{w,j}|v_{x,j}\rangle = 1$.

4.2 The Dual of $\mathsf{Adv}^{\pm}$ is Span Program Witness Size

From the canonical span program above we write the witness size as the following optimization problem:

$$\text{wsize}(P_f) = \min_{\{|v_{x,j}\rangle\}} \max_{s\in\{0,1\}^n, j\in[n]} \||v_{s,j}\rangle\|^2$$

[2] Since $\langle w|t\rangle = 1$ while $\langle w|A\Pi(w) = 0$.

subject to the constraint that for all pairs $(w,x) \in F_0 \times F_1$, $\sum_{w_j \neq x_j} \langle v_{w,j}|v_{x,j}\rangle = 1$. Let X be a positive semi-definite (PSD) matrix such that entry $\langle w,i|X|x,j\rangle = \langle v_{w,i}|v_{x,j}\rangle$ for all $w \in F_0$ and $x \in F_1$. Write $\text{wsize}(P_f)$ as the following equivalent semi-definite program (SDP)

$$\text{wsize}(P_f) = \min_{X \succeq 0} \max_{s \in \{0,1\}^n} \sum_{j \in [n]} \langle s,j|X|s,j\rangle$$

subject to the constraint that for all $(x,w) \in F_0 \times F_1$, $\sum_{w_j \neq x_j} \langle w,j|X|x,j\rangle = 1$.

We will turn the above SDP into the general adversarial bound. First introduce a variable ξ in order to eliminate the inner maximization function. For adversary matrix Γ let $\Gamma_j = \Gamma \circ D_j$.

$$\text{wsize}(P_f) = \min_{\substack{X \succeq 0,\ \xi \geq 0, \\ \forall (w,x) \in F_0 \times F_1 : \sum_{w_j \neq x_j} \langle w,j|X|x,j\rangle = 1, \\ \forall s \in \{0,1\}^n : \xi \geq \sum_{j \in [n]} \langle s,j|X|s,j\rangle}} \xi \tag{2}$$

$$= \max_{\substack{\{\alpha_{w,x}\}, \\ \beta_s \geq 0,\ \sum_s \beta_s = 1, \\ \sum_s \beta_s |s\rangle\langle s| \succeq \sum_{w,x \in F_0 \times F_1, w_j \neq x_j} \alpha_{w,x} |w\rangle\langle x|}} \sum \alpha_{w,x} \qquad \text{(SDP duality; see Appendix A)} \tag{3}$$

$$= \max_{\substack{\{\alpha_{w,x}\}, \\ \beta_s \geq 0,\ \sum_s \beta_s = 1, \\ \sum_{s \in \{0,1\}^n} |s'\rangle\langle s'| \succeq \sum_{w,x \in F_0 \times F_1, w_j \neq x_j} \frac{\alpha_{w,x}}{\sqrt{\beta_w \beta_x}} |w'\rangle\langle x'|}} \sum \alpha_{w,x} \qquad \left(\text{substitute } |s'\rangle = \frac{1}{\sqrt{\beta_s}}|s\rangle\right) \tag{4}$$

$$= \max_{\substack{\{\alpha'_{w,x}\}, \\ \beta_s \geq 0,\ \sum_s \beta_s = 1, \\ \sum_{s \in \{0,1\}^n} |s'\rangle\langle s'| \succeq \sum_{w,x \in F_0 \times F_1, w_j \neq x_j} \alpha'_{w,x} |w'\rangle\langle x'|}} \alpha'_{w,x}\sqrt{\beta_w \beta_x} \quad \left(\text{substitute } \alpha'_{w,x} = \alpha_{w,x}/\sqrt{\beta_w \beta_x}\right) \tag{5}$$

$$= \max_{\substack{\Gamma_{w,x} = \alpha'_{w,x}, \\ |\beta\rangle_s = \sqrt{\beta_s},\ \||\beta\rangle\| = 1, \\ \mathbb{I} - \Gamma_i \succeq 0}} \langle \beta|\Gamma|\beta\rangle \qquad (\langle w|\Gamma_j|x\rangle = 0 \text{ if } w_j = x_j) \tag{6}$$

$$= \max_{\substack{\Gamma_{w,x} = \alpha'_{w,x}, \\ \|\Gamma_i\| \leq 1}} \|\Gamma\| = \mathsf{Adv}^{\pm}(f) \tag{7}$$

4.3 Span Programs as Graphs

The canonical span program matrix A of f can be transformed into the biadjacency matrices of two bipartite graphs. These graphs capture the evaluation of a string s on f [Rei10; Rei11; RS12]. Let $B_{G(s)} \in \mathbb{C}^{(F_0 \cup I') \times (\{\mu_0\} \cup I)}$ and $B_{G'(s)} \in \mathbb{C}^{(F_0 \cup I') \times I}$ be the true and false biadjacency matrices corresponding to the bipartite graph G of the span program respectively. In particular,

$$B_{G(s)} = \begin{array}{c} \begin{matrix} \mu_0 & I \end{matrix} \\ \begin{bmatrix} |t\rangle & A \\ 0 & \overline{\Pi}(s) \end{bmatrix} \end{array} \begin{array}{c} F_0 \\ I' \end{array} \qquad B_{G'(s)} = \begin{array}{c} \begin{matrix} I \end{matrix} \\ \begin{bmatrix} A \\ \overline{\Pi}(s) \end{bmatrix} \end{array} \begin{array}{c} F_0 \\ I' \end{array} \tag{8}$$

where $|t\rangle$ and A are defined as

$$|t\rangle = \frac{1}{3\sqrt{W}} \sum_{w \in F_0} |w\rangle \text{ and } A = \sum_{w \in F_0, j \in [n]} |w\rangle\langle j, \overline{w}_j| \otimes |v_{w,j}\rangle \tag{9}$$

and $\overline{\Pi}(s) = \mathbb{I} - \Pi(s) \in \mathbb{L}(\mathbb{C}^I)$ where $\Pi(s)$ is defined in Equation 1.

Matrix-vector products $B_{G(s)}|\psi\rangle$ and $B^*_{G'(s)}|\psi'\rangle$ can be interpreted as operating on the sets of column vectors separately. That is, let $|\psi\rangle = |\psi_1\rangle + |\psi_2\rangle$ where $|\psi_1\rangle := \alpha|0\rangle$ operates on column μ_0 and $|\psi_2\rangle$ operates on columns I. Similarly, let $|\psi'\rangle = |\psi'_1\rangle + |\psi'_2\rangle$ where $|\psi'_1\rangle$ and $|\psi'_2\rangle$ operates on columns F_0 and I'.

Example 4.5. *Let us turn the canonical span program of the parity function, shown in Example 4.4, into its corresponding bipartite graphs. The matrices $B_{G(x)}$ and $B_{G'(w)}$ are then defined as follows for strings $x = 10$ and $w = 00$ which evaluates to true and false respectively on $\oplus$.*

$$B_{G(x)} = \begin{array}{c} \begin{matrix} \mu_0 & \multicolumn{4}{c}{I} \end{matrix} \\ \left[\begin{array}{c|cccc} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{matrix} \\ \\ F_0 \\ \\ \\ I' \end{matrix} \end{array} \qquad B^*_{G'(w)} = \begin{array}{c} \begin{matrix} & \multicolumn{2}{c}{F_0} & \multicolumn{4}{c}{I'} \end{matrix} \\ \left[\begin{array}{cc|cccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] I \end{array} .$$

These corresponds to the bipartite graphs shown Figure 1 and Figure 2.

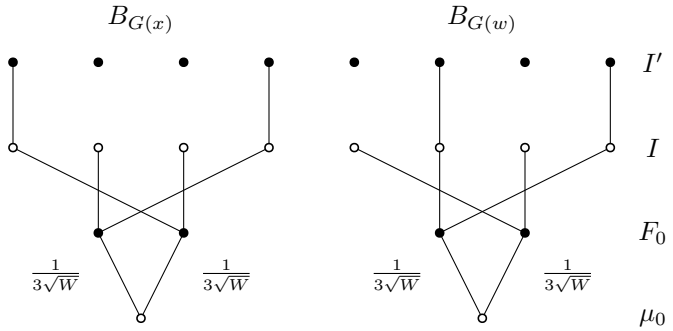

Figure 1: The bipartite graphs corresponding to the true biadjacency matrix. All unmarked edges have weight one. The matix vector product $B_{G(x)}|\phi\rangle$ is equivalent to assigning weights to the open dots in the picture. In order to find an eigenvalue zero eigenvector $|\phi\rangle$ of $B_{G(x)}$, the assignment of weights must ensure the neighbours of every solid dot sums to zero. Observe that $B_{G(x)}$, with $\oplus(x) = 1$, has an eigenvalue zero eigenvector while $B_{G(w)}$, with $\oplus(w) = 0$, does not.

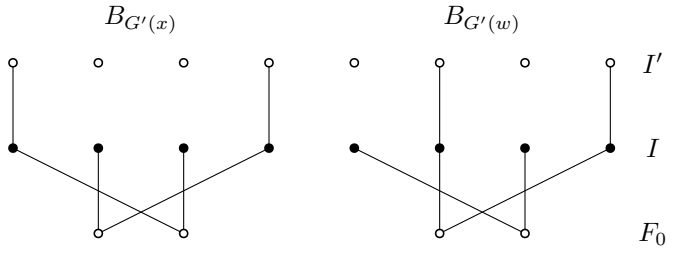

Figure 2: The bipartite graphs corresponding to the false biadjacency matrix. As opposed to the above, $B_{G'(x)}$, with $\oplus(x) = 1$, does not have an eigenvalue zero eigenvector while $B_{G'(w)}$, with $\oplus(w) = 0$, does.

Lemma 4.6. (Spectral Gap of Eigenvalue Zero Eigenvectors.) *If $f(x) = 1$, then the vector*

$$|\psi\rangle = |\psi_1\rangle + |\psi_2\rangle \text{ where } |\psi_1\rangle = -3\sqrt{W}|0\rangle \text{ and } |\psi_2\rangle = \sum_{j\in[n]} |j, x_j\rangle \otimes |v_{x,j}\rangle$$

is an eigenvalue zero eigenvector of $B_{G(x)}$*. Further,* $|\psi\rangle$ *satisfies* $|\langle 0|\psi\rangle|^2 \geq 9\left\||\psi\rangle\right\|^2/10$*.*

If instead $f(w) = 0$*, then the vector*

$$|\psi'\rangle = |\psi'_1\rangle + |\psi'_2\rangle \text{ where } |\psi'_1\rangle = -|w\rangle \text{ and } |\psi'_2\rangle = \sum_{j\in[n]} |j, \overline{w}_j\rangle \otimes |v_{w,j}\rangle$$

is an eigenvalue zero eigenvector of $B_{G'(w)}$*. Further,* $|\psi'\rangle$ *satisfies* $|\langle t|\psi'\rangle|^2 \geq \left\||\psi\rangle\right\|^2/(9W(W+1))$*.*

Proof. Let $x \in F_1$ and $w \in F_0$. Observe that $B_{G(x)}|\psi_1\rangle$ is the vector with $|F_0|$ non-zero entries followed by $|I'|$ zeros. Since $f(x) = 1$, there exists a linear combination of the columns of A which sum to $|t\rangle$. Choosing this set of columns will also ensure that the rows indexed by I' sum to zero as an entry of $|\psi_2\rangle$ is non-zero only when the associated column of $\overline{\Pi}(x)$ is. Further $|\langle 0|\psi\rangle|^2 = \left\||\psi_1\rangle\right\|^2 = 9W$, while $\left\||\psi\rangle\right\|^2 = \left\||\psi_1\rangle\right\|^2 + \left\||\psi_2\rangle\right\|^2 = 9W + W$ by definition. Similarly, observe that $B^*_{G'(w)}|\psi'_1\rangle$ multiplies the column associated with w among the rows of A by negative one, while $B^*_{G'(w)}|\psi'_2\rangle$ is exactly this same column. Further, $|\langle t|\psi'\rangle|^2 = 1/(9W)$ and $\left\||\psi'\rangle\right\|^2 = \left\||\psi'_1\rangle\right\|^2 + \left\||\psi'_2\rangle\right\|^2 = 1 + W$. □

5 Optimal Quantum Query Algorithms for Span Programs

Let $|t\rangle$ and A, as shown in Equation 9, be the matrix of the canonical span program. Further let G be the associated bipartite graph with biadjacency matrix B_G and adjacency matrix A_G as follows

$$B_G = \begin{matrix} \mu_0 & I \\ [|t\rangle & A] \end{matrix} \begin{matrix} \\ F_0 \end{matrix} \quad \text{and } A_G = \begin{matrix} F_0 & \mu_0 & I \\ \end{matrix} \begin{bmatrix} 0 & |t\rangle & A \\ \langle t| & 0 & 0 \\ A^* & 0 & 0 \end{bmatrix} \begin{matrix} F_0 \\ \mu_0 \\ I \end{matrix} \tag{10}$$

Let $\Delta \in \mathbb{L}(\mathbb{C}^{F_0 \cup \{\mu_0\} \cup I})$ be the orthogonal projection onto the span of all eigenvalue zero eigenvectors of A_G. For a string $s \in \{0,1\}^n$, let $\Pi_s \in \mathbb{L}(\mathbb{C}^{F_0 \cup \{\mu_0\} \cup I})$ be

$$\Pi_s = \mathbb{I} - \sum_{j\in[n], k\in[m]} |j, \overline{s}_j, k\rangle\langle j, \overline{s}_j, k|.$$

The graph $G(s)$ has biadjacency matrix $B_{G(s)}$ (from Equation 8) and adjacency matrix $A_{G(s)}$.

$$B_{G(s)} = \begin{matrix} \mu_0 & I \\ \end{matrix} \begin{bmatrix} |t\rangle & A \\ 0 & \overline{\Pi}(s) \end{bmatrix} \begin{matrix} F_0 \\ I' \end{matrix} \quad \text{and } A_{G(s)} = \begin{matrix} F_0 & I' & \mu_0 & I \\ \end{matrix} \begin{bmatrix} 0 & 0 & |t\rangle & A \\ 0 & 0 & 0 & \overline{\Pi}(s) \\ \langle t| & 0 & 0 & 0 \\ A^* & \overline{\Pi}(s) & 0 & 0 \end{bmatrix} \begin{matrix} F_0 \\ I' \\ \mu_0 \\ I \end{matrix} \tag{11}$$

Note that $A_{G(s)} \in \mathbb{L}(\mathbb{C}^{F_0 \cup I' \cup \{\mu_0\} \cup I})$ contains A_G and the additional vertices of I'. Further $\mathbb{I} - \Pi_s \in \mathbb{L}(\mathbb{C}^{F_0 \cup \{\mu_0\} \cup I})$ contains $\overline{\Pi}(s) \in \mathbb{L}(\mathbb{C}^I)$ as a subgraph and is everywhere else all zeros.

Define $U_s \in \mathbb{L}(\mathbb{C}^{F_0 \cup \{\mu_0\} \cup I})$ as

$$U_s = (2\Pi_s - \mathbb{I})(2\Delta - \mathbb{I}),$$

the matrix which reflects a vector across Δ then across Π_s. Observe that Δ is independent of the input s, while Π_s requires one query of the quantum f-oracle. The following are three different quantum query

algorithms which computes $f(s)$ with query complexity W.

Algorithm 1: Phase Estimation

Initialize state $|0\rangle \in \mathbb{C}^{F_0 \cup \mu_0 \cup I}$
$\delta_p \leftarrow \frac{1}{100W}$
$\delta_e \leftarrow \frac{1}{10}$
Run phase estimation on U_s with precision δ_p and error δ_e
Return 1 if phase estimation returns zero, otherwise return 0

Algorithm 2: Quantum Search

Initialize state $|+\rangle \otimes |0\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^{F_0 \cup \mu_0 \cup I}$
$T \leftarrow$random integer in $\{1, ..., \lceil 100W \rceil\}$
Apply $|0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes U_x^T$ to initial state
Measure the first qubit in the Hadamard basis
Return 1 if the value is $|+\rangle$, otherwise return 0

Algorithm 3: Quantum Search without Register

Initialize state $|0\rangle \in \mathbb{C}^{F_0 \cup \mu_0 \cup I}$
$T \leftarrow$random integer in $\{1, ..., \lceil 100W \rceil\}$
Apply U_x^T to $|0\rangle$
Measure $U_x^T|0\rangle$ in the standard basis
Return 1 if the value is $|0\rangle$, otherwise return 0

We will omit the analysis for Algorithm 3 here as is quite complex and the algorithm itself has equivalent quantum query complexity to the other two. Instead, we will focus our attention on the analysis for the first two algorithms.

The following lemma about the "effective spectral" gap of $A_{G(s)}$ will be necessary for the analysis. Its intuition and proof can be found in Appendix B.

Lemma 5.1. (Effective Spectral Gap.) *If $f(s) = 1$ then $A_{G(s)}$ has an eigenvalue zero eigenvector $|\psi\rangle$ with $|\langle 0|\psi\rangle|^2 \geq 9 \||\psi\rangle\|^2 / 10$.*

If $f(w) = 0$ and $\{|\alpha\rangle\}$ is the set of all orthonormal eigenvectors with corresponding eigenvalues $\rho(\alpha)$ of $A_{G(s)}$, then for any $c \geq 0$

$$\sum_{\alpha:|\rho(\alpha)|\leq c/W} |\langle \alpha|0\rangle|^2 \leq 72c^2 \left(1 + \frac{1}{W}\right).$$

5.1 Spectral Gap for U_s

Using Lemma 5.1, we prove a spectral gap on the eigenvectors of the matrix U_s.

Lemma 5.2. *If $f(s) = 1$ then U_s has an eigenvalue one eigenvector $|\varphi\rangle$ with $|\langle 0|\varphi\rangle|^2 / \||\varphi\rangle\|^2 \geq 9/10$.*

If $f(s) = 0$ and $\{|\beta\rangle\}$ is a set of orthonormal eigenvectors of U_s with corresponding eigenvalues $e^{i\theta(\beta)}$, where $\theta(\beta) \in (-\pi, \pi]$. Then for any $\Theta \geq 0$

$$\sum_{\beta:|\theta(\beta)|\leq\Theta} |\langle \beta|0\rangle|^2 \leq \left(2\sqrt{6\Theta W} + \frac{\Theta}{2}\right)^2$$

A key tool used to prove Lemma 5.2 is the fact that we can rotate the basis of U_s so that it becomes represented by a block-diagonal matrix, where the blocks are of maximum dimension two[3]. This was proved by Szegedy [Sze04]. Nagaj, Wocjan, and Zhang [NWZ09] gave a different proof that follows from a Lemma of Jordan [Jor75]:

Lemma 5.3. ([Jor75]) *Given projections Π_s and Δ in Hilbert space $\mathcal{H}$, there exists a decomposition of $\mathcal{H}$ into orthogonal one-dimensional and two-dimensional subspaces invariant under Π_s and Δ. On the two-dimensional subspaces, Π_s and Δ are rank-one projectors.*

Lemma 5.3 implies that $\mathcal{H}$ can be decomposed into a set of one-dimensional subspaces $\{T_i\}$ and a set of two-dimensional subspaces $\{S_i\}$. Each one-dimensional subspace T_i is spanned by a vector $|v_i\rangle$ for which there exists $b, c \in \{0,1\}$ such that $\Delta|v_i\rangle = b|v_i\rangle$ and $\Pi_s|v_i\rangle = c|v_i\rangle$: that is, each of Δ and Π_s either act as the identity on T_i or are orthogonal to T_i. Each two-dimensional subspace S_i is spanned by vectors $|v_i\rangle, |v_i^\perp\rangle$ such that $\Delta|v_i\rangle = |v_i\rangle$ and $\Delta|v_i^\perp\rangle = 0$. Also, S_i is spanned by vectors $|w_i\rangle, |w_i^\perp\rangle$ such that $\Pi_s|w_i\rangle = |w_i\rangle$ and $\Pi_s|w_i^\perp\rangle = 0$. Let $\theta_i = 2\arccos|\langle v_i|w_i\rangle|$. Then,

$$|w_i\rangle = \cos\frac{\theta_i}{2}|v_i\rangle + \sin\frac{\theta_i}{2}|v_i^\perp\rangle \qquad\qquad |w_i^\perp\rangle = -\sin\frac{\theta_i}{2}|v_i\rangle + \cos\frac{\theta_i}{2}|v_i^\perp\rangle$$

Theorem 5.4. ([Sze04; NWZ09]) *Let $\{S_i\}, \{T_i\}$ be the decomposition of Π_s and Δ given by Lemma 5.3. Then U_s has eigenvalues $e^{\mp i\theta_i}$ corresponding to $\frac{|v_i\rangle \pm i|v_i^\perp\rangle}{\sqrt{2}}$ on each two-dimensional subspace S_i, and eigenvalue either 1 or -1 on each one-dimensional subspace T_i.*

Proof. On one-dimensional subspace T_i, each individual reflection multiplies a vector by ± 1, so both reflections in succession do as well. For the rest of the proof, consider a two-dimensional subspace S_i. By the above relationship between $\{|v_i\rangle, |v_i^\perp\rangle\}$ and $\{|w_i\rangle, |w_i^\perp\rangle\}$, we get the following:

$$\begin{bmatrix} |w_i\rangle \\ |w_i^\perp\rangle \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta_i}{2} & \sin\frac{\theta_i}{2} \\ -\sin\frac{\theta_i}{2} & \cos\frac{\theta_i}{2} \end{bmatrix} \begin{bmatrix} |v_i\rangle \\ |v_i^\perp\rangle \end{bmatrix}$$

Recall that $\sigma_y = \left[\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right]$ is the Pauli Y matrix, which has eigenvalues 1 and -1 corresponding to eigenvectors $|\phi_y^+\rangle = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right]$ and $|\phi_y^-\rangle = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -i \end{smallmatrix}\right]$, respectively.

$$\begin{aligned}
\begin{bmatrix} \cos\frac{\theta_i}{2} & \sin\frac{\theta_i}{2} \\ -\sin\frac{\theta_i}{2} & \cos\frac{\theta_i}{2} \end{bmatrix} &= \cos\frac{\theta_i}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + i\sin\frac{\theta_i}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= \frac{\cos\frac{\theta_i}{2}}{2}\left(\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} + \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}\right) + \frac{i\sin\frac{\theta_i}{2}}{2}\left(\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} - \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}\right) \\
&= \left(\frac{\cos\frac{\theta_i}{2}}{2} + \frac{i\sin\frac{\theta_i}{2}}{2}\right)\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} + \left(\frac{\cos\frac{\theta_i}{2}}{2} - \frac{i\sin\frac{\theta_i}{2}}{2}\right)\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \\
&= e^{i\theta_i/2}\frac{1}{2}\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} + e^{-i\theta_i/2}\frac{1}{2}\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \\
&= e^{i\theta_i/2}|\phi_y^+\rangle\langle\phi_y^+| + e^{-i\theta_i/2}|\phi_y^-\rangle\langle\phi_y^-| = e^{(i\theta/2)\sigma_y}
\end{aligned}$$

In the basis $\{|v_i\rangle, |v_i^\perp\rangle\}$, we have that $(2\Delta - \mathbb{I}) = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right] - \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right] = \sigma_z$, where σ_z is the Pauli Z matrix. Similarly, in the basis $\{|w_i\rangle, |w_i^\perp\rangle\}$, we have that $(2\Pi(s) - \mathbb{I}) = \sigma_z$. Then, in the basis $\{|v_i\rangle, |v_i^\perp\rangle\}$, $U_s = (2\Pi(s) - \mathbb{I})(2\Delta - \mathbb{I}) = e^{(-i\theta_i/2)\sigma_y}\sigma_z e^{(i\theta_i/2)\sigma_y}\sigma_z = e^{(-i\theta_i/2)\sigma_y}e^{(-i\theta_i/2)\sigma_y}\sigma_z\sigma_z = e^{(-i\theta_i)\sigma_y}$, where we used the fact that σ_y and σ_z anticommute ($\sigma_y\sigma_z = -\sigma_y\sigma_z$) and the fact that $\sigma_z\sigma_z = \mathbb{I}$. Therefore, the eigenvalues are $e^{-i\theta_i}$ and $e^{i\theta_i}$, corresponding to eigenvectors $|\phi_y^+\rangle = \frac{|v_i\rangle + i|v_i^\perp\rangle}{\sqrt{2}}$ and $|\phi_y^-\rangle = \frac{|v_i\rangle - i|v_i^\perp\rangle}{\sqrt{2}}$ respectively. □

Now we can prove Lemma 5.2.

[3]We can do this for any unitary made up of two reflections

Proof. Let $\{|\beta\rangle\}$ be the set of eigenvectors given by the decomposition of Theorem 5.4. Since Δ is the projection into the nullspace of A_G, $A_G\Delta = 0$. Thus $A_{G(s)}(\Delta \oplus \mathbb{I}) = T(\mathbb{I} - \Pi_s) \otimes (\mathbb{I}_2)$ for a permutation matrix T since G is a subgraph of $A_{G(s)}$ and $\overline{\Pi}(s)$ is a submatrix of $\mathbb{I} - \Pi_s$ from Equation 10 and Equation 11.

First consider the case where $f(s) = 1$. Take $|\psi\rangle$ to be the eigenvalue zero eigenvector of $A_{G(s)}$ such that $|\langle 0|\psi\rangle|^2 \geq 9\,\||\psi\rangle\|^2/10$ from Lemma 5.1. Obtain $|\phi\rangle$ from $|\psi\rangle$ by restricting to the entries corresponding to the index sets $F_0 \cup \{\mu_0\} \cup I$. Since $|\psi\rangle$ is an eigenvalue zero eigenvector of $A_{G(s)}$ (see Lemma 4.6), it is not supported on the removed entries so $\||\psi\rangle\| = \||\phi\rangle\|$ and $|\phi\rangle$ is an eigenvalue zero eigenvector of A_G. Thus $\Delta|\phi\rangle = |\phi\rangle$. Since Π_s is the identity matrix on the support of $|\psi\rangle$, $\Pi_s|\phi\rangle = |\phi\rangle$. Together $U_s|\phi\rangle = |\phi\rangle$.

Now consider the case where $f(s) = 0$. Let $|\zeta\rangle = \sum_{\beta:|\theta(\beta)|\leq\Theta} |\beta\rangle\langle\beta|0\rangle$: this is the projection of $|0\rangle$ onto low-angle subspaces of U_s. We want to bound $\sum_{\beta:|\theta(\beta)|\leq\Theta} |\langle\beta|0\rangle|^2 = \sum_{\beta:|\theta(\beta)|\leq\Theta} \langle 0|\beta\rangle\langle\beta|0\rangle = \langle 0|\zeta\rangle$. We will find it more convenient to bound $|\langle 0|\hat{\zeta}\rangle|^2 = \langle 0|\zeta\rangle$, where $|\hat{\zeta}\rangle$ is the normalized vector $|\zeta\rangle/\,\||\zeta\rangle\|$.

Observe that $|\hat{\zeta}\rangle$ is not supported on any eigenvectors $|\beta\rangle$ where $\theta(\beta) = 0$. Without loss of generality, $\theta(\beta) = 0$ only when $|\beta\rangle$ is in a one-dimensional subspace T_i with eigenvalue one. Then $(2\Pi_s - \mathbb{I})$ and $(2\Delta - \mathbb{I})$ either both reflect $|\beta\rangle$ or they both don't. In the first case, $\Pi_s|\beta\rangle = \Delta|\beta\rangle = 0$, so $\langle 0|\beta\rangle = \langle 0|\Pi_s|\beta\rangle = 0$ because $\Pi_s|0\rangle = |0\rangle$. In the second case, $\Pi_s|\beta\rangle = \Delta|\beta\rangle = |\beta\rangle$ and so $A_{G(x)}|\beta\rangle = A_{G(x)}\Delta|\beta\rangle = T(\mathbb{I} - \Pi_s)|\beta\rangle = T(\beta - \beta) = 0$, so by the $f(x) = 0$ case of Lemma 5.1 with $c = 0$ we have that $\langle 0|\beta\rangle = 0$.

The observation above implies that if we consider $\Theta < \pi$ (the Lemma is trivial otherwise), $e^{i\theta\beta} \neq \pm 1$ for the $|\beta\rangle$ in the support of $|\hat{\zeta}\rangle$, and so we can restrict our analysis to just the two-dimensional subspaces of U_s. We now split $\langle 0|\hat{\zeta}\rangle$:

$$\begin{aligned}
\langle 0|\hat{\zeta}\rangle &= \langle 0|\Delta + (\mathbb{I} - \Delta)|\hat{\zeta}\rangle \\
&= \langle 0|\Delta|\hat{\zeta}\rangle + \langle 0|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle && (\Pi_s|0\rangle = |0\rangle) \\
&\leq |\langle 0|\Delta|\hat{\zeta}\rangle| + |\langle 0|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle| && \text{(by the triangle inequality)} \\
&\leq |\langle 0|\Delta|\hat{\zeta}\rangle| + \|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle\|
\end{aligned}$$

Now our goal is to bound both of the values in the last expression. First we bound $\|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle\|$.

Given an eigenvector $|\beta\rangle$ in the support of $|\hat{\zeta}\rangle$, let $|-\beta\rangle$ be the other eigenvector in the two-dimensional subspace containing $|\beta\rangle$. Note that $\theta(\beta) = -\theta(-\beta)$. Let $|\hat{\zeta}\rangle = \sum_\beta c_\beta|\beta\rangle$, where here the sum is over all eigenvectors[4]. Then $\|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle\|^2 = \|\sum_\beta \Pi_s(\mathbb{I} - \Delta)c_\beta|\beta\rangle\|^2$. Thanks to Theorem 5.4, we can break this summation up into pairs.

$$\begin{aligned}
\|\Pi_s(\mathbb{I} - \Delta)|\hat{\zeta}\rangle\|^2 &= \sum_{\beta:\theta(\beta)>0} \|\Pi_s(\mathbb{I} - \Delta)(c_\beta|\beta\rangle + c_{-\beta}|-\beta\rangle)\|^2 \\
&= \sum_{S_i:\theta_i\neq 0} \left\| \frac{i}{\sqrt{2}}(c_{-\beta} - c_\beta)\Pi_s|v_i^\perp\rangle \right\|^2 && \left(\text{rewrite } |\beta\rangle, |-\beta\rangle \text{ in terms of } |v_i\rangle, |v_i^\perp\rangle\right) \\
&= \sum_{S_i:\theta_i\neq 0} \left\| \frac{i}{\sqrt{2}}(c_{-\beta} - c_\beta)\sin\frac{\theta_i}{2}|w_i\rangle \right\|^2 && \text{(change of basis)} \\
&= \sum_{\beta:\theta(\beta)>0} \left(\sin\frac{\theta(\beta)}{2}\right)^2 \left\| \frac{i}{\sqrt{2}}(c_{-\beta} - c_\beta)|w_i\rangle \right\|^2 \\
&\leq \sum_{\beta:\theta(\beta)>0} \left(\sin\frac{\theta(\beta)}{2}\right)^2 \leq \left(\frac{\Theta}{2}\right)^2 && (\sin\theta \leq \theta \text{ for the values considered})
\end{aligned}$$

[4]Not just the ones in the support of $|\hat{\zeta}\rangle$

Next we bound the term $|\langle 0|\Delta|\hat{\zeta}\rangle$ which we will write as $|\langle 0|w\rangle|\||\Delta|\hat{\zeta}\rangle\|$ where $|w\rangle = \Delta|\hat{\zeta}\rangle/\|\Delta|\hat{\zeta}\rangle\|$ is the normalized projection of the vector $|\hat{\zeta}\rangle$ onto span of the eigenvalue zero eigenvectors of A_G. We will work exclusively with $|w\rangle$. First we bound the magnitude of the vector $\left\|A_{G(x)}|w\rangle\right\|$, then decompose $|w\rangle$ into its components in the space of "small" and "large" eigenvalue eigenvectors of $A_{G(x)}$ for particular choices of "small" and "large".

$$\begin{aligned}
\|A_{G(x)}\Delta|\hat{\zeta}\rangle\|^2 &= \|(\mathbb{I}-\Pi_s)\Delta|\hat{\zeta}\rangle\|^2 \\
&= \sum_{S_i:\theta_i\neq 0}\left\|\frac{i}{\sqrt{2}}(c_{-\beta}-c_\beta)\Delta|v_i^\perp\rangle\right\|^2 && \left(\text{rewrite } |\beta\rangle, |-\beta\rangle \text{ in terms of } |v_i\rangle, |v_i^\perp\rangle\right) \\
&= \sum_{\beta:\theta(\beta)\geq 0}\left(\sin\frac{\theta(\beta)}{2}\right)^2\left\|\frac{i}{\sqrt{2}}(c_{-\beta}-c_\beta|w_i\rangle\right\|^2 && \text{(change of basis)} \\
&\leq \left(\frac{\Theta}{2}\right)^2\|\Delta|\hat{\zeta}\rangle\|^2.
\end{aligned}$$

By the definition of $|w\rangle$, we have

$$\left\|A_{G(x)}|w\rangle\right\|^2 = \frac{\|A_{G(x)}\Delta|\hat{\zeta}\rangle\|^2}{\|\Delta|\hat{\zeta}\rangle\|^2} \leq \frac{\Theta}{2}.$$

For a fixed d, to be determined later, let $|w\rangle = |w_{\text{small}}\rangle + |w_{\text{big}}\rangle$ where

$$|w_{\text{small}}\rangle = \sum_{\alpha:|\rho(\alpha)|\leq d\Theta/2}|\alpha\rangle\langle\alpha|w\rangle \text{ and } |w_{\text{big}}\rangle = \sum_{\alpha:|\rho(\alpha)|>d\Theta/2}|\alpha\rangle\langle\alpha|w\rangle.$$

Thus we have

$$|\langle 0|\Delta|\hat{\zeta}\rangle = |\langle 0|w\rangle|\left\|\Delta|\hat{\zeta}\rangle\right\| \leq |\langle 0|w\rangle| \leq |\langle 0|w_{\text{small}}\rangle| + |\langle 0|w_{\text{big}}\rangle|$$

where the equality is by definition, the first inequality is due to the fact that the projection of the unit vector $|\hat{\zeta}\rangle$, and the second is by triangle inequality.

Bound $|\langle 0|w\rangle|$ as follows:

$$\begin{aligned}
|\langle 0|w_{\text{small}}\rangle|^2 &= \left(\sum_{\alpha:|\rho(\alpha)|\leq d\Theta/2}\langle 0|\alpha\rangle\langle\alpha|w\rangle\right)^2 \\
&\leq \left(\sum_{\alpha:|\rho(\alpha)|\leq d\Theta/2}|\langle 0|\alpha\rangle|^2\right)\cdot\left(\sum_{\alpha:|\rho(\alpha)|\leq d\Theta/2}|\langle\alpha|w\rangle|^2\right) && \text{(Cauchy-Schwartz)} \\
&= \left(\sum_{\alpha:|\rho(\alpha)|\leq d\Theta/2}|\langle 0|\alpha\rangle|^2\right)\||w_{\text{small}}\rangle\|^2 && \text{(definition of } |w_{\text{small}}\rangle) \\
&\leq 72c^2\left(1+\frac{1}{W}\right)\||w_{\text{small}}\rangle\|^2 && \left(\text{Lemma 5.1 with } c=\frac{d\Theta W}{2}\right) \\
&\leq 6d\Theta W && (W\geq 1 \text{ and } |w\rangle \text{ is normalized})
\end{aligned}$$

We further have $A_{G(x)}|w\rangle = \sum_\alpha \rho(\alpha)|\alpha\rangle\langle\alpha|w\rangle$ so

$$\begin{aligned}
\left(\frac{\Theta}{2}\right)^2 &\geq \left\|A_{G(x)}|w\rangle\right\|^2 \\
&= \left\|A_{G(x)}|w_{\text{small}}\rangle\right\|^2 + \left\|A_{G(x)}|w_{\text{big}}\rangle\right\|^2 && \text{(orthogonality of } |\alpha\rangle) \\
&\geq d^2\left(\frac{\Theta}{2}\right)^2\||w_{\text{big}}\rangle\|^2
\end{aligned}$$

Thus $\||w_{\text{big}}\rangle\| \leq 1/d$. Since $|0\rangle$ is a column of the identity matrix, $\langle 0|w_{\text{big}}\rangle \leq \||w_{\text{big}}\rangle\|$. Together we have

$$\sqrt{\sum_{\beta:|\theta(\beta)|\leq\Theta} |\langle\beta|0\rangle|^2} = \langle 0|\hat{\zeta}\rangle \leq |\langle 0|\Delta|\hat{\zeta}\rangle| + \|\Pi_s(\mathbb{I}-\Delta)|\hat{\zeta}\rangle\| \leq |\langle 0|w_{\text{small}}\rangle| + |\langle 0|w_{\text{big}}\rangle| + \frac{\Theta}{2} \leq 6d\Theta W + \frac{1}{d} + \frac{\Theta}{2}.$$

Choosing $d = 1/\sqrt{6\Theta W}$, we find the bound to be $2\sqrt{6\Theta W} + \Theta/2$. □

5.2 Analysis of the Algorithms

Given the spectral gap for U_s given in Lemma 5.2, we can analyze the algorithms.

Algorithm 1 measures the phase of U_s with input $|0\rangle$, which is in general a superposition of eigenvectors of U_s. If $f(s) = 1$ then by Lemma 5.2 most of the amplitude of $|0\rangle$ is in the direction of an eigenvector with phase zero, and so the likelihood of measuring phase zero is at least $9/10$ minus the error δ_e, which gives a probability of at least $4/5$. If $f(s) = 0$, then if we set Θ to be the precision δ_p then only a very small amount of the amplitude of $|0\rangle$ is in the direction of eigenvectors with phase zero: by Lemma 5.2, the algorithm will measure of a phase of zero with probability at most $\delta_e + (2\sqrt{6\delta_p W} + \delta_p/2)^2 < 2/5$.

Algorithm 2 prepares the state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle U_s^T|0\rangle)$ and measures the first qubit in the basis $\{|+\rangle, |-\rangle\}$, which is equivalent to measuring the first qubit of $H|\varphi\rangle = \frac{1}{2}(|0\rangle|0\rangle + |1\rangle 0 + |0\rangle U_s^T|0\rangle - |1\rangle U_x^T|0\rangle)$ in the standard basis. The first qubit of $H|\varphi\rangle$ has amplitude $\frac{1}{2} + \frac{1}{2}\langle 0|U_s^T|0\rangle$ in the $|0\rangle$ direction, and so we will measure $|0\rangle$ with probability $\frac{1}{4}\|(\mathbb{I} + U_s^T)|0\rangle\|^2$. When $f(s) = 1$, this probability will be at least $9/10$ regardless of T. When $f(s) = 0$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{T\in[\tau]}\left[\frac{1}{4}\left\|\left(\mathbb{I}+U_s^T\right)|0\rangle\right\|^2\right] &= \mathop{\mathbb{E}}_{T\in[\tau]}\left[\frac{1}{4}\sum_\beta |1+\exp(i\theta(\beta)T)|^2 |\langle 0|\beta\rangle|^2\right] \\
&= \frac{1}{4}\sum_\beta |\langle 0|\beta\rangle|^2 \sum_{T=1}^{\tau} \frac{(2+2\exp(i\theta(\beta)T))}{\tau} \\
&= \frac{1}{4}\sum_\beta |\langle 0|\beta\rangle|^2 \left(2 + \frac{1}{\tau}\sum_{T=1}^{\tau} 2\exp(i\theta(\beta)T)\right) \\
&= \frac{1}{4}\sum_\beta |\langle 0|\beta\rangle|^2 \left(2 + \frac{1}{\tau}\left(\sum_{T=-\tau}^{\tau} \exp(i\theta(\beta)T) - \exp(i\theta(\beta)\cdot 0)\right)\right) \\
&= \frac{1}{4}\sum_\beta |\langle 0|\beta\rangle|^2 \left(2 + \frac{1}{\tau}\left(\frac{\exp(i\theta(\beta)(\tau+1)) - \exp(-i\theta(\beta)\tau)}{e^{i\theta(\beta)}-1} - 1\right)\right)
\end{aligned}$$

We let $\Theta = 1/(50W)$ and define $\nu = (2\sqrt{6\Theta W} + \Theta/2)^2$. Divide the $|\beta\rangle$ by their eigenvalues. For $\theta(\beta) \leq \Theta$,

we use Lemma 5.2 to bound the terms in the sum by ν. Next consider those $|\beta\rangle$ such that $\theta(\beta) > \Theta$.

$$\begin{aligned}
&\sum_{|\beta\rangle:\theta(\beta)>\Theta} |\langle 0|\beta\rangle|^2 \left(\frac{1}{2} + \frac{1}{4\tau}\left(\frac{\exp(i\theta(\beta)(\tau+1)) - \exp(-i\theta(\beta)\tau)}{e^{i\theta(\beta)}-1} - 1\right)\right) \\
&\quad \leq (1-\nu)\cdot\left(\frac{1}{2} + \frac{1}{4\tau}\left(\frac{\exp(i\theta(\beta)(\tau+1)) - \exp(-i\theta(\beta)\tau)}{e^{i\theta(\beta)}-1} - 1\right)\right) \\
&\quad = (1-\nu)\cdot\left(\frac{1}{2} + \frac{1}{4\tau}\left(\frac{\exp(i\Theta(\tau+1)) - \exp(-i\Theta\tau) - \exp(i\Theta) + 1}{\exp(i\Theta)-1}\right)\right) \\
&\quad = (1-\nu)\cdot\left(\frac{1}{2} + \frac{1}{4\tau}\left(\frac{\sin(\Theta(\tau+1/2)) - \sin(\Theta/2)}{\sin(\Theta/2)}\right)\right) \\
&\quad = (1-\nu)\cdot\left(\frac{1}{2} + \frac{1}{4\tau\sin(\Theta/2)}\right) \qquad (\Theta \in (0,\pi])
\end{aligned}$$

Thus algorithm two outputs 1 with probability at most $\nu + (1-\nu)\cdot(1/2 + 1/(4\tau\sin(\Theta/2))$. When $\tau = \lceil 100W \rceil$ and $W > 1$ this probability is at most 88%.

5.2.1 Analysis of Algorithm 3

The analysis of Algorithm 3 requires a bit more work so we will devote the remainder of the survey to this task. We want to show a separation between the probability that Algorithm 3 outputs 1 when $f(x) = 1$ and when $f(x) = 0$. In particular we show that the probability is greater than 64% and 61% in the former and latter case respectively.

Let $\tau = \lceil 10^5 W \rceil$. The probability that the algorithm outputs one is

$$p = \mathop{\mathbb{E}}_{T\in[\tau]}\left[|\langle 0|U_x^T|0\rangle|^2\right] = \mathop{\mathbb{E}}_{T\in[\tau]}\left|\sum_\beta e^{i\theta(\beta)T}\langle\beta|0\rangle|^2\right|^2. \tag{12}$$

A Lagrangian Duality

Consider the following objective function:

$$\begin{aligned} \text{Minimize} \quad & f_0(|x\rangle) \\ \text{Subject to} \quad & f_i(|x\rangle) \leq 0 \text{ for } i \in [m] \\ & h_j(|x\rangle) = 0 \text{ for } j \in [p] \end{aligned}$$

for x in some domain $\mathcal{D} \subset \mathbb{R}^n$. Then the associated *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^j$ is the function

$$L(|x\rangle, |\lambda\rangle, |\nu\rangle) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(|x\rangle) + \sum_{j=1}^{p} \nu_j h_j(|x\rangle).$$

Further, the *Lagrangian dual function* is

$$g(|\lambda\rangle, |\nu\rangle) = \inf_{|x\rangle \in \mathcal{D}} F(|x\rangle, |\lambda\rangle, |\nu\rangle).$$

Observe that $g(|\lambda\rangle, |\nu\rangle)$ is a lower bound for the optimal value p^* of the objective function above when $|\lambda\rangle \geq 0$. Let $|x\rangle$ be any feasible solution, then $f_i(|x\rangle) \leq 0$ and $h_j(|x\rangle) = 0$. Thus

$$p^* \geq f_0(|x\rangle) \geq f_0(|x\rangle) + \sum_{i=1}^{m} \lambda_i f_i(|x\rangle) + \sum_{j=1}^{p} \nu_j h_j(|x\rangle) = L(|x\rangle, |\lambda\rangle, |\nu\rangle) \geq \inf_{|x\rangle \in \mathcal{D}} F(|x\rangle, |\lambda\rangle, |\nu\rangle) = g(|\lambda\rangle, |\nu\rangle).$$

The best lower bound is obtained by maximizing over the dual function.

In our case

$$\begin{aligned} \text{Minimize} \quad & t \\ \text{Subject to} \quad & \sum_{w_j \neq x_j} \langle w, j|X|x, j\rangle = 1 \text{ for all } (x, w) \in \Delta, w_j \neq x_j \\ & \sum_{j \in [n]} \langle s, j|X|s, j\rangle \leq t \text{ for all } s \in \{0,1\}^n \end{aligned}$$

where $X \succeq 0$. The Lagrangian has one variable for every constraint. Let $Y \succeq 0$ be the variable for the constraint $X \succeq 0$, $\alpha_{w,x}$ and $\beta_s \geq 0$ be the variables for the equality and inequality constraints respectively. Then

$$L = L(Y, |\alpha\rangle, |\beta\rangle; X, t) = t - \langle X|Y\rangle + \sum_{(x,w) \in \Delta, w_j \neq x_j} \alpha_{x,w} \left(1 - \langle w, j|X|x, j\rangle\right) - \sum_{|s\rangle \in \{0,1\}^n} \beta_s \left(t - \langle s, j|X|s, j\rangle\right)$$

with dual function

$$g(Y, |\alpha\rangle, |\beta\rangle) = \inf_{X,t} L(Y, |\alpha\rangle, |\beta\rangle; X, t).$$

Since the infimum is taken over all $X \succeq 0$ and values t, there exists choices of Y, $|\alpha\rangle$ and $|\beta\rangle$ such that $\inf_{X \succeq 0, t} L(Y, |\alpha\rangle, |\beta\rangle) = -\infty$. To remove these values from consideration, we find the implicit constraints.

Fix $Y, |\alpha\rangle, |\beta\rangle, X$ and rewrite L interms of t.

$$L = t\left(1 - \sum_{|s\rangle \in \{0,1\}^n} \beta_s\right) - \langle X|Y\rangle + \sum_{(x,w) \in \Delta, w_j \neq x_j} \alpha_{x,w} \left(1 - \langle w, j|X|x, j\rangle\right) + \sum_{|s\rangle \in \{0,1\}^n} \beta_s \langle s, j|X|s, j\rangle.$$

Since the last three terms are fixed, by taking $t \to -\infty$, $L \to -\infty$. Thus we require

$$1 = \sum_{|s\rangle \in \{0,1\}^n} \beta_s.$$

Similarly, fix $Y, |\alpha\rangle, |\beta\rangle, t$ and rewrite L interms of X.

$$L = \langle X|Z - Y\rangle + t + \sum_{(x,w)\in\Delta} \alpha_{x,w} - t \sum_{|s\rangle\in\{0,1\}^n} \beta_s$$

where $Z = \sum_{|s\rangle\in\{0,1\}^n} \beta_s |s\rangle\langle s| - \sum_{(x,w)\in\Delta, w_j \neq x_j} \alpha_{x,w} |w\rangle\langle x|$. Again, if $\langle X|Z-Y\rangle \neq 0$, then X can be chosen such that $L \to -\infty$. Thus $Z = Y$. Since $Y \succeq 0$, we can simplify this to $Z \succeq 0$.

B Spectral Analysis of Adjacency and Biadjacency Matrices

Let G be a weighted bipartite graph with biadjacency matrix $B_G \in \mathcal{L}(\mathbb{C}^U, \mathbb{C}^T)$ and weighted adjacency matrix $A_G \in \mathcal{L}(\mathbb{C}^{T\cup U})$. Further let $|t\rangle \in \mathbb{C}^T$ and G' be the graph with biadjacency matrix

$$B_{G'} = \begin{matrix} \mu_0 & U \\ [|t\rangle & B_G] \end{matrix} \quad \begin{matrix} \\ T \end{matrix}$$

and adjacency graph $A_{G'}$. To understand the eigenvectors of the modified adjacency matrix $A_{G'}$, we need the following theorem about eigenvectors of a PSD matrix.

Theorem B.1. (Spectral Bounds for PSD Matrices, Theorem 8.9 [RS12].) *Let $X \in \mathbb{L}(V)$ with $X \succeq 0$, $|t\rangle \in V$, and $X' = X + |t\rangle\langle t|$. Further, let $\{|\beta\rangle\}$ be the eigenvectors of X' with corresponding eigenvalue $\theta(\beta) \geq 0$. If there exists a vector $|\psi\rangle$ in the null-space of X with $|\langle t|\psi\rangle|^2 \geq \delta \||\psi\rangle\|^2$, then for any $\gamma \geq 0$*

$$\sum_{\beta:|\theta(\beta)|\leq\gamma, \langle t|\beta\rangle\neq 0} \frac{|\langle t|\beta\rangle|^2}{\lambda(\beta)} \leq \frac{4\gamma}{\delta}.$$

Note that this sum is well defined since $\theta(\beta) \geq 0$ whenever $\langle t|b\rangle \neq 0$ then $\langle\beta|X'|\beta\rangle = \langle\beta|X|\beta\rangle + \|\langle t|\beta\rangle\|^2 > 0$.

Theorem B.2. (Spectral Properties of Small Eigenvalue Eigenvectors.) *Let G, B_G, A_G, G', $B_{G'}$, and $A_{G'}$ be as before. Suppose for some $\delta > 0$, A_G has an eigenvalue zero eigenvector such that*

$$|\langle t|\psi_T\rangle|^2 \geq \delta \||\psi\rangle\|^2 .$$

Let $\{|\alpha\rangle\}$ be the complete set of orthonormal eigenvectors of $A_{G'}$ with corresponding eigenvalues $\rho(\alpha)$. Further, let $|0\rangle$ be the vector $[0,1,0]^\top \in \mathbb{C}^{T\cup\{\mu_0\}\cup U}$. Then for all $\gamma > 0$, we have

$$\sum_{\alpha:|\rho(\alpha)|\leq\gamma} |\langle\alpha|0\rangle|^2 \leq \frac{8\gamma^2}{\delta}.$$

Proof. The structure of the proof is as follows. We begin by reviewing relationships between the eigenvectors and eigenvalues of the adjacency graph A_G and the biadjacency graph B_G. Given an eigenvector of A_G, we will relate this to the eigenvectors of the modified adjacency matrix $A_{G'}$ and modified biadjacency graph $B_{G'}$. Central to this analysis will be the study of PSD matrix $B_{G'}B_{G'}^*$.

Let G be a graph and A_G and B_G be its adjacency and biadjacency matrices as described in the theorem statement. Let $|\psi\rangle = (|\psi_T\rangle, |\psi_U\rangle) \in \mathbb{C}^{T\cup U}$ be an eigenvector of A_G with associated eigenvalue $\rho > 0$ i.e.

$$\begin{bmatrix} 0 & B_G \\ B_G^* & 0 \end{bmatrix} \cdot \begin{bmatrix} |\psi_T\rangle \\ |\psi_U\rangle \end{bmatrix} = \rho \begin{bmatrix} |\psi_T\rangle \\ |\psi_U\rangle \end{bmatrix}.$$

Then we obtain the identities $B_G|\psi_U\rangle = \rho|\psi_T\rangle$ and $B_G^*|\psi_T\rangle = \rho|\psi_U\rangle$. By negating these identities, we observe that $(|\psi_T\rangle, -|\psi_U\rangle)$ is also an eigenvector of A_G with associated eigenvalue $-\rho$. Observe further that $|\psi_T\rangle$, defined to be $\frac{1}{\rho}B_G|\psi_U\rangle$, is an eigenvector of $B_GB_G^*$ with eigenvalue ρ^2. Similarly $|\psi_U\rangle$, defined to be $\frac{1}{\rho}B_G^*|\psi_T\rangle$, is an eigenvector of $B_G^*B_G$ with eigenvalue ρ^2. If, instead, we begin with an eigenvector $|\phi\rangle \in \mathbb{C}^T$ of $B_GB_G^*$ with eigenvalue λ, then $B_G^*|\phi\rangle \in \mathbb{C}^U$ is an eigenvalue of $B_G^*B_G$ with eigenvalue λ then

$$B_G^*B_G\left(B_G^*|\phi\rangle\right) = \lambda B_G^*|\phi\rangle.$$

The pair $(|\phi\rangle, \frac{\pm 1}{\sqrt{\lambda}}B_G^*|\phi\rangle)$ are eigenvectors of A_G with eigenvalues $\pm\sqrt{\lambda}$ since, in the positive case for example,

$$B_G\left(\frac{1}{\sqrt{\lambda}}B_G^*|\phi\rangle\right) = \sqrt{\lambda}|\phi\rangle \text{ and } B_G^*|\phi\rangle = \sqrt{\lambda}\left(\frac{1}{\sqrt{\lambda}}B_G^*|\phi\rangle\right)$$

since $|\phi\rangle$ is an eigenvector of $B_GB_G^*$ with eigenvector λ for the former.

Let $(|\psi_T\rangle, 0)$, an eigenvector of A_G, be the input to our theorem. Note that $|\langle t|\psi_T\rangle|^2 \geq \delta \||\psi\rangle\|^2$ and $B_G^*|\psi_T\rangle = 0$. We would like to bound the magnitude of

$$\sum_{\alpha:|\rho(\alpha)|\leq\gamma} |\langle\alpha|0\rangle|^2$$

where $\{|\alpha\rangle\}$ is a complete set of orthonormal eigenvectors of $A_{G'}$ with associated eigenvalue $\rho(\alpha)$ and $|0\rangle$ is the indicator vector for entry corresponding to μ_0. First we show that the eigenvalue zero eigenvectors of $A_{G'}$ are unsupported on μ_0 so will not contribute to this sum. We bound $|\langle\alpha|0\rangle|^2$ for eigenvectors $|\alpha\rangle$ with $0 < \rho(\alpha) \leq \gamma$ using Theorem B.1 by considering the eigenvectors of $B_GB_G^*$.

Let $|\zeta\rangle = (|\zeta_T\rangle, \zeta_{\mu_0}, |\zeta_U\rangle)$ be an eigenvalue zero eigenvector of $A_{G'}$. Then modified biadjacency matrix $B_{G'}$ must satisfy

$$B_{G'}(\zeta_{\mu_0}, |\zeta_U\rangle) = \begin{bmatrix}|t\rangle & B_G\end{bmatrix} \cdot \begin{bmatrix}\zeta_{\mu_0} \\ |\zeta_U\rangle\end{bmatrix} = \zeta_{\mu_0}|t\rangle + B_G|\zeta_U\rangle = 0.$$

By multiplying both sides by $\langle\psi_T|$, we have

$$\zeta_{\mu_0}\langle\psi_T|t\rangle + \langle\psi_T|B_G|\zeta_T\rangle = 0,$$

since $B_G^*|\psi_T\rangle = 0$ and $\langle t|\psi_T\rangle > 0$, $\zeta_{\mu_0} = 0$. Thus eigenvalue zero eigenvectors of $A_{G'}$ are orthogonal to $|0\rangle$.

It remains to consider those eigenvectors $|\alpha\rangle = (|\alpha_T\rangle, \alpha_{\mu_0}, |\alpha_U\rangle)$ of $A_{G'}$ with $\rho(\alpha) > 0$. First, using the definition of eigenvectors and the property that $A_{G'}|0\rangle = |t\rangle$, we have

$$\rho(\alpha)\langle\alpha|0\rangle = \langle\alpha|A_{G'}|0\rangle = \langle\alpha_T|t\rangle.$$

Substituting this into our desired sum, we obtain

$$\sum_{\alpha:0<|\rho(\alpha)|\leq\gamma} |\langle\alpha|0\rangle|^2 = \sum_{\alpha:0<|\rho(\alpha)|\leq\gamma} \frac{|\langle\alpha_T|t\rangle|^2}{\rho(\alpha)^2}.$$

Let $B_{G'}B_{G'}^*$ be a matrix with eigenvectors $\{|\beta\rangle\}$ and corresponding eigenvalues $\theta(\beta)$. By the relationship between the eigenvalues and eigenvectors of $A_{G'}$ and $B_{G'}$ considered above, each $|\beta\rangle$ with $\theta(\beta)$ corresponds to two eigenvectors of $A_{G'}$ with eigenvalue $\left(|\beta\rangle, \frac{\pm 1}{\sqrt{\lambda(\beta)}}B_{G'}^*|\beta\rangle\right)$ with $\pm\sqrt{\theta(\beta)}$. Thus

$$\sum_{\alpha:0<|\rho(\alpha)|\leq\gamma} \frac{|\langle\alpha_T|t\rangle|^2}{\rho(\alpha)^2} = 2\sum_{\beta:\theta(\beta)\leq\gamma^2,\theta(\beta)\neq 0} \frac{|\langle\beta|t\rangle|^2}{\theta(\beta)}.$$

Using Theorem B.1 with $X = B_{G'}B_{G'}^* = B_GB_G^* - |t\rangle\langle t|$ and $|\psi_T\rangle$ gives us the bound

$$\sum_{\alpha:|\rho(\alpha)|\leq\gamma} |\langle\alpha|0\rangle|^2 = 2\sum_{\beta:\theta(\beta)\leq\gamma^2,\theta(\beta)\neq 0} \frac{|\langle\beta|t\rangle|^2}{\theta(\beta)} \leq \frac{8\gamma^2}{\delta}$$

as required. □

Applying Theorem B.2 with $\delta = 1/(9W(W+1))$ and $\gamma = c/W$ to Lemma 4.6 in the case where $f(s) = 0$, we obtain the following Lemma 5.1.

References

[Amb+10] Andris Ambainis et al. "Any AND-OR Formula of Size N Can Be Evaluated in Time $N^{1/2+o(1)}$ on a Quantum Computer". In: *SIAM J. Comput.* 39.6 (2010), pp. 2513–2530. DOI: 10.1137/080712167. URL: https://doi.org/10.1137/080712167.

[Amb02] Andris Ambainis. "Quantum Lower Bounds by Quantum Arguments". In: *J. Comput. Syst. Sci.* 64.4 (2002), pp. 750–767. DOI: 10.1006/jcss.2002.1826. URL: https://doi.org/10.1006/jcss.2002.1826.

[Amb06] Andris Ambainis. "Polynomial degree vs. quantum query complexity". In: *J. Comput. Syst. Sci.* 72.2 (2006), pp. 220–238. DOI: 10.1016/j.jcss.2005.06.006. URL: https://doi.org/10.1016/j.jcss.2005.06.006.

[Bea+01] Robert Beals et al. "Quantum lower bounds by polynomials". In: *J. ACM* 48.4 (2001), pp. 778–797. DOI: 10.1145/502090.502097. URL: https://doi.org/10.1145/502090.502097.

[BSS03] Howard Barnum, Michael E. Saks, and Mario Szegedy. "Quantum query complexity and semidefinite programming". In: *18th Annual IEEE Conference on Computational Complexity (Complexity 2003), 7-10 July 2003, Aarhus, Denmark*. 2003, pp. 179–193. DOI: 10.1109/CCC.2003.1214419. URL: https://doi.org/10.1109/CCC.2003.1214419.

[HLS07] Peter Høyer, Troy Lee, and Robert Spalek. "Negative weights make adversaries stronger". In: *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11-13, 2007*. 2007, pp. 526–535. DOI: 10.1145/1250790.1250867. URL: https://doi.org/10.1145/1250790.1250867.

[Jor75] Camille Jordan. "Essai sur la géométrie à n dimensions". In: *Bulletin de la Société mathématique de France* 3 (1875), pp. 103–174.

[KW90] Mauricio Karchmer and Avi Wigderson. "Monotone Circuits for Connectivity Require Super-Logarithmic Depth". In: *SIAM J. Discrete Math.* 3.2 (1990), pp. 255–265. DOI: 10.1137/0403021. URL: https://doi.org/10.1137/0403021.

[KW93] Mauricio Karchmer and Avi Wigderson. "On Span Programs". In: *Proceedings of the Eigth Annual Structure in Complexity Theory Conference, San Diego, CA, USA, May 18-21, 1993*. 1993, pp. 102–111. DOI: 10.1109/SCT.1993.336536. URL: https://doi.org/10.1109/SCT.1993.336536.

[LLS06] Sophie Laplante, Troy Lee, and Mario Szegedy. "The Quantum Adversary Method and Classical Formula Size Lower Bounds". In: *Computational Complexity* 15.2 (2006), pp. 163–196. DOI: 10.1007/s00037-006-0212-7. URL: https://doi.org/10.1007/s00037-006-0212-7.

[NWZ09] Daniel Nagaj, Pawel Wocjan, and Yong Zhang. "Fast amplification of QMA". In: *Quantum Information & Computation* 9.11 (2009), pp. 1053–1068. URL: http://www.rintonpress.com/xxqic9/qic-9-1112/1053-1068.pdf.

[Rei10] Ben Reichardt. "Span programs and quantum query algorithms". In: *Electronic Colloquium on Computational Complexity (ECCC)* 17 (2010), p. 110. URL: http://eccc.hpi-web.de/report/2010/110.

[Rei11] Ben Reichardt. "Reflections for quantum query algorithms". In: *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*. 2011, pp. 560–569. DOI: 10.1137/1.9781611973082.44. URL: https://doi.org/10.1137/1.9781611973082.44.

[RS12] Ben Reichardt and Robert Spalek. "Span-Program-Based Quantum Algorithm for Evaluating Formulas". In: *Theory of Computing* 8.1 (2012), pp. 291–319. DOI: 10.4086/toc.2012.v008a013. URL: https://doi.org/10.4086/toc.2012.v008a013.

[Sze04] Mario Szegedy. "Quantum Speed-Up of Markov Chain Based Algorithms". In: *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*. 2004, pp. 32–41. DOI: 10.1109/FOCS.2004.53. URL: https://doi.org/10.1109/FOCS.2004.53.

[Tal17] Avishay Tal. "Formula lower bounds via the quantum method". In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*. 2017, pp. 1256–1268. DOI: 10.1145/3055399.3055472. URL: https://doi.org/10.1145/3055399.3055472.