# Quantum query complexity of state conversion

Troy Lee Rajat Mittal Ben W. Reichardt Robert Špalek Mario Szegedy

**Abstract**

State conversion generalizes query complexity to the problem of converting between two input-dependent quantum states by making queries to the input. We characterize the complexity of this problem by introducing a natural information-theoretic norm that extends the Schur product operator norm. The complexity of converting between two systems of states is given by the distance between them, as measured by this norm.

In the special case of function evaluation, the norm is closely related to the general adversary bound, a semi-definite program that lower-bounds the number of input queries needed by a quantum algorithm to evaluate a function. We thus obtain that the general adversary bound characterizes the quantum query complexity of any function whatsoever. This generalizes and simplifies the proof of the same result in the case of boolean input and output. Also in the case of function evaluation, we show that our norm satisfies a remarkable composition property, implying that the quantum query complexity of the composition of two functions is at most the product of the query complexities of the functions, up to a constant. Finally, our result implies that discrete and continuous-time query models are equivalent in the bounded-error setting, even for the general state-conversion problem.

## 1 Introduction

A quantum query algorithm for evaluating a function $f$ attempts to compute $f(x)$ with as few queries to the input $x$ as possible. Equivalently, the algorithm begins in a state $|0\rangle$, and should approach $|f(x)\rangle \otimes |0\rangle$. The *state-conversion* problem generalizes function evaluation to the case where the algorithm begins in a state $|\rho_x\rangle$ and the goal is to convert this to $|\sigma_x\rangle$. State-conversion problems arise naturally in algorithm design, generalizing classical subroutines (Figure 1). For example, the graph isomorphism problem can be reduced to creating a certain quantum state [Shi02].

We characterize the quantum query complexity of state conversion. We introduce a natural, information-theoretic norm, which extends the Schur product operator norm. The complexity of state conversion depends only on the Gram matrices of the sets of vectors $\{|\rho_x\rangle\}$ and $\{|\sigma_x\rangle\}$, and is characterized as the norm of the difference between these Gram matrices. For example, in function evaluation, the initial Gram matrix is the all-ones matrix, $J$, and the target Gram matrix is $F = \{\delta_{f(x),f(y)}\}_{x,y}$, so the query complexity depends only on the norm of $F - J$. Characterizing query complexity in terms of a norm-induced metric has interesting consequences. For example, it follows that if one can design an optimal algorithm for going from $J$ to $\frac{99}{100}J + \frac{1}{100}F$, then one also obtains an optimal algorithm for evaluating $f$.

The norm we introduce is related to the general adversary bound [HLŠ07], a strengthening of the popular adversary method for showing lower bounds on quantum query complexity [Amb02]. A recent sequence of works [FGG08, CCJY09, ACR+10, RŠ08] has culminated in showing that the general adversary bound gives, up to a constant factor, the bounded-error quantum query

complexity of any function with boolean output and binary input alphabet [Rei09, Rei11]. Our more general state-conversion result completes this picture by showing that the general adversary bound characterizes the bounded-error quantum query complexity of any function whatsoever:

**Theorem 1.1.** *Let $f : \mathcal{D} \to E$, where $\mathcal{D} \subseteq D^n$, and $D$ and $E$ are finite sets. Then the bounded-error quantum query complexity of $f$, $Q(f)$, is characterized by the general adversary bound, $\mathrm{Adv}^{\pm}(f)$:*

$$Q(f) = \Theta\big(\mathrm{Adv}^{\pm}(f)\big) \ . \tag{1.1}$$

The general adversary bound is a semi-definite program (SDP). When phrased as a minimization problem, $\mathrm{Adv}^{\pm}(f)$ only has constraints on $x, y$ pairs where $f(x) \neq f(y)$. In contrast, we consider an SDP that places constraints on *all* input pairs $x, y$. Fortunately, these extra constraints increase the optimal value by at most a factor of two. The extra constraints, however, are crucial for the construction of the algorithm, and for any extension to state conversion. They also lead to a new conceptual understanding of the adversary bound.

The modified SDP defines a norm, and we define the *query distance* as the metric induced by this new norm. Our main algorithmic theorem states that there is a quantum algorithm that converts $|\rho_x\rangle$ to a state with high fidelity to $|\sigma_x\rangle$, and that makes a number of queries of order the query distance between $\rho$ and $\sigma$, the respective Gram matrices of $\{|\rho_x\rangle\}$ and $\{|\sigma_x\rangle\}$.

The correctness of our algorithm has a direct and particularly simple proof. Though more general, it simplifies the previous characterization of boolean function evaluation. At its mathematical heart is a lemma that gives an "effective" spectral gap for the product of two reflections.

The query distance also gives lower bounds on the query complexity of state conversion. It is straightforward to argue that the query distance between $\rho$ and $\sigma$ lower bounds the number of queries needed to reach $\sigma$ *exactly*. To deal with the bounded-error case, we look at the minimum over all $\sigma'$ of the query distance between $\rho$ and $\sigma'$, where $\sigma'$ is a valid Gram matrix for the final states of a successful algorithm. We show that a simpler necessary and sufficient condition for the latter is that $\sigma$ and $\sigma'$ are close in the distance induced by the Schur product operator norm.

As the query distance remains a lower bound on the continuous-time query complexity, a corollary of our algorithm is that the continuous-time and discrete query models are related up to a constant factor in the bounded-error setting. Previously, this equivalence was known up to a sub-logarithmic factor [CGM+09].

Since the general adversary bound characterizes quantum query complexity, its properties immediately carry over thereto. We show that the general adversary bound satisfies a remarkable composition property, that $\mathrm{Adv}^{\pm}$ of a composed function $f \circ (g, g, \ldots, g)$ is $O\big(\mathrm{Adv}^{\pm}(f)\mathrm{Adv}^{\pm}(g)\big)$. Previously this was known in the boolean case [Rei09], and, again, having constraints on all input pairs turns out to be crucial in the extension to non-boolean functions. When the input of $f$ is boolean, we can show a matching lower bound, extending [HLŠ07].

## 2 Background

For a natural number $n$, let $[n] = \{1, 2, \ldots, n\}$. For two matrices $A$, $B$ of the same size, $A \circ B$ denotes their entrywise product, also known as Schur or Hadamard product. Let $\langle A, B\rangle = \mathrm{Tr}(A^\dagger B)$. Denote by $\|A\|$ the spectral norm of $A$. We will use $\mathbf{1}$ and $J$ for the identity and all-ones matrices, respectively, where size can be inferred from the context. Let $\delta_{a,b}$ be the Kronecker delta function.

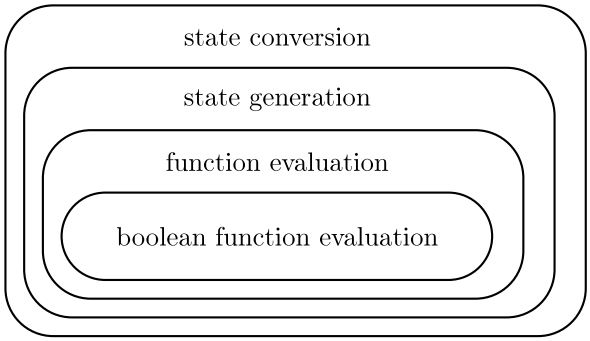


Figure 1: The state-conversion problem generalizes the state-generation problem, studied by Ambainis et al. [AMRR11], which in turn generalizes function evaluation. The quantum query complexity of evaluating boolean functions has been characterized by [Rei11].

### 2.1 Coherent and non-coherent state-conversion problems

The *quantum query complexity* of a function $f$, $Q(f)$, is the number of input queries needed to evaluate $f$ with error at most $1/3$ [BW02]. The *state-conversion* problem generalizes function evaluation to the case where the aim is to transform one quantum state into another, using queries to the input. The problem is parameterized by sets of states $\{|\rho_x\rangle\}$ and $\{|\sigma_x\rangle\}$. On input $x$, we begin in a state $|\rho_x\rangle$ and wish to create the state $|\sigma_x\rangle$, using as few queries to $x$ as possible. State conversion is a slight generalization of the *state-generation* problem, in which in the initial state $|\rho_x\rangle$ is independent of $x$. This problem was introduced by Shi [Shi02], and recently studied systematically by Ambainis et al. [AMRR11].

State conversion has two variants, coherent and non-coherent. Both versions allow the algorithm workspace.

**Definition 2.1** (Coherent output condition)**.** *An algorithm solves the coherent state-conversion problem with error $\epsilon$ if for every $x$ it replaces the initial state $|\rho_x\rangle \otimes |0\rangle$ by a state $|\sigma'_x\rangle$ such that* $\mathrm{Re}(\langle\sigma'_x|(|\sigma_x\rangle \otimes |0\rangle)) \geq \sqrt{1-\epsilon}$.

**Definition 2.2** (Non-coherent output condition)**.** *An algorithm solves the non-coherent state-conversion problem with error $\epsilon$ if for every $x$ it replaces the initial state $|\rho_x\rangle \otimes |0\rangle$ by a state $|\sigma'_x\rangle$ such that $|\langle\sigma'_x|(|\sigma_x\rangle \otimes |s_x\rangle)| \geq \sqrt{1-\epsilon}$ for some state $|s_x\rangle$ that may depend on $x$.*

The query complexity of state conversion only depends on the Gram matrices of the initial and target states, i.e., on $\rho = \{\langle\rho_x|\rho_y\rangle\}_{x,y}$ and $\sigma = \{\langle\sigma_x|\sigma_y\rangle\}_{x,y}$. (In function evaluation and state generation, $\rho$ is the all-ones matrix.) Let $Q_\epsilon(\rho,\sigma)$ and $Q^{nc}_\epsilon(\rho,\sigma)$ be the minimum number of queries required to solve the coherent and non-coherent state-conversion problems, respectively, with error $\epsilon$.

For evaluating functions, coherent and non-coherent complexities are equal up to constant factors. An example that illustrates the difference between these output conditions is computing a boolean function in the phase, that is where the target state $|\sigma_x\rangle = (-1)^{f(x)}|0\rangle$. In this case the non-coherent complexity is trivial, while the bounded-error coherent complexity is equal to $Q(f)$, up to constant factors.

### 2.2 The $\gamma_2$ norm

We will make use of the $\gamma_2$ norm, also known as the Schur product operator norm [Bha07, LSŠ08]. This norm has been introduced recently to complexity theory by Linial et al. [LMSS07], and has proven to be very useful in quantum information. It is currently the best lower bound known on quantum communication [LS09], and Tsirelson has shown that its dual norm characterizes the bias of a quantum XOR game [Tsi87, Ung08].

**Definition 2.3.** *Let $A$ be matrix with rows labeled by $\mathcal{D}_1$ and columns by $\mathcal{D}_2$. Define*

$$\gamma_2(A) = \min_{\substack{m\in\mathbf{N},\\ |u_x\rangle,|v_y\rangle\in\mathbf{C}^m}} \left\{ \max\left\{ \max_{x\in\mathcal{D}_1} \||u_x\rangle\|^2, \max_{y\in\mathcal{D}_2} \||v_y\rangle\|^2 \right\} : \forall x\in\mathcal{D}_1, y\in\mathcal{D}_2,\ A_{x,y} = \langle u_x|v_y\rangle \right\} . \quad (2.1)$$

The following fact plays a key role in the design of our algorithm and in relating our new norm to the general adversary bound:

**Fact 2.4.** *For any $k\in\mathbf{N}$, let $\mathbf{1}$ and $J$ be the $k$-by-$k$ identity and all-ones matrices, respectively. Then $\gamma_2(J-\mathbf{1}) \le 2(1-1/k)$.*

*Proof.* We demonstrate unit vectors $\{|\mu_i\rangle\}_{i\in[k]}$, and $\{|\nu_i\rangle\}_{i\in[k]}$ such that $\langle\mu_i|\nu_j\rangle = \frac{1}{2}\frac{k}{k-1}(1-\delta_{i,j})$: let $|\mu_i\rangle = -\alpha|i\rangle + \frac{\sqrt{1-\alpha^2}}{\sqrt{k-1}}\sum_{j\neq i}|j\rangle$, $|\nu_i\rangle = \sqrt{1-\alpha^2}|i\rangle + \frac{\alpha}{\sqrt{k-1}}\sum_{j\neq i}|j\rangle$, for $\alpha = \sqrt{\frac{1}{2}-\frac{\sqrt{k-1}}{k}}$. □

## 3 Filtered $\gamma_2$ norm and query distance

We define a natural generalization of the $\gamma_2$ norm, in which the factorization is filtered through certain matrices:

**Definition 3.1** (Filtered $\gamma_2$ norm)**.** *Let $A$ be a matrix with rows indexed by elements of $\mathcal{D}_1$ and columns by $\mathcal{D}_2$, and let $Z = \{Z_1,\ldots,Z_n\}$ be a set of $|\mathcal{D}_1|$-by-$|\mathcal{D}_2|$ matrices. Define $\gamma_2(A|Z)$ by*

$$\begin{aligned}\gamma_2(A|Z) = \min_{\substack{m\in\mathbf{N},\\ |u_{xj}\rangle,|v_{yj}\rangle\in\mathbf{C}^m}} & \max\left\{ \max_{x\in\mathcal{D}_1}\sum_j \||u_{xj}\rangle\|^2, \max_{y\in\mathcal{D}_2}\sum_j \||v_{yj}\rangle\|^2 \right\} \\ & \forall x\in\mathcal{D}_1, y\in\mathcal{D}_2,\ A_{x,y} = \sum_j (Z_j)_{x,y}\langle u_{xj}|v_{yj}\rangle \ .\end{aligned} \quad (3.1)$$

The filtered $\gamma_2$ norm $\gamma_2(\,\cdot\,|Z)$ is a norm. Among its many properties (see Appendix A) are that $\gamma_2(A) = \gamma_2(A|\{J\})$, where $J$ is the all-ones matrix, and $\gamma_2(A|\{A\}) = 1$ if $A\neq 0$. We use below the general inequality

$$\gamma_2(A|\{Z_j\}) \le \gamma_2(A|\{Z_j\circ B\})\gamma_2(B) \ . \quad (3.2)$$

The query distance is the metric induced when the filter matrices are related to the query process. Let $\mathcal{D}\subseteq D^n$ be a finite set, and let $\Delta = \{\Delta_1,\ldots,\Delta_n\}$, where $\Delta_j = \{1-\delta_{x_j,y_j}\}_{x,y\in\mathcal{D}}$. Thus $\Delta_j$ encodes when a query to index $j$ distinguishes input $x$ from input $y$.

**Definition 3.2.** *The* query distance *between $\rho$ and $\sigma$, two $|\mathcal{D}|$-by-$|\mathcal{D}|$ matrices, is $\gamma_2(\rho-\sigma|\Delta)$.*

Theorem 4.9 below shows that the query distance characterizes the quantum query complexity of state conversion. Furthermore, as we show now, it is closely related to the general adversary bound, a lower bound on the quantum query complexity for function evaluation introduced by [HLŠ07]. Let $f : \mathcal{D} \to E$ and let $F = \{\delta_{f(x),f(y)}\}_{x,y}$.

**Definition 3.3.** *The general adversary bound for $f$ is given by*

$$\mathrm{Adv}^{\pm}(f) = \max\left\{ \|\Gamma\| : \forall j \in [n], \|\Gamma \circ \Delta_j\| \leq 1 \right\} , \tag{3.3}$$

*where the maximization is over $|\mathcal{D}|$-by-$|\mathcal{D}|$ real, symmetric matrices $\Gamma$ satisfying $\Gamma \circ F = 0$.*

By taking the dual of this SDP, we obtain a bound that is the same as $\gamma_2(J - F|\Delta)$, except without any constraints on pairs $x, y$ with $f(x) = f(y)$. In other words, $\mathrm{Adv}^{\pm}(f) = \gamma_2(J - F|\{\Delta_j \circ (J - F)\})$.

**Theorem 3.4.** *The values of the general adversary bound and $\gamma_2(J - F|\Delta)$ differ by at most a factor of two, and are equal in the case that the function has boolean output:*

$$\mathrm{Adv}^{\pm}(f) \leq \gamma_2(J - F|\Delta) \leq 2\big(1 - 1/|E|\big)\mathrm{Adv}^{\pm}(f) .$$

*Proof.* Since $\mathrm{Adv}^{\pm}(f)$ has fewer constraints as a minimization problem, $\mathrm{Adv}^{\pm}(f) \leq \gamma_2(J - F|\Delta)$.

For the other direction, use Eq. (3.2) with $Z_j = \Delta_j$, $A = B = J - F = A \circ B$. As it is readily seen that $\gamma_2$ is invariant under adding or removing duplicate rows or columns, Fact 2.4 implies $\gamma_2(J - F) \leq 2(1 - 1/|E|)$. □

# 4 Characterization of quantum query complexity

In this section, we show that the query complexities for function evaluation, and coherent and non-coherent state conversion are characterized in terms of $\gamma_2$ and $\gamma_2(\cdot|\Delta)$. We begin with the upper bounds.

## 4.1 Quantum query algorithm for state conversion

**Theorem 4.1.** *Consider the problem of converting states $\{|\rho_x\rangle\}$ to $\{|\sigma_x\rangle\}$, for $x \in \mathcal{D} \subseteq D^n$. Let $\rho$ and $\sigma$ be the states' Gram matrices. For any $\epsilon \in (0, \gamma_2(\rho - \sigma|\Delta))$, this problem has query complexity*

$$Q_\epsilon(\rho, \sigma) = O\Big(\gamma_2(\rho - \sigma|\Delta)\frac{\log(1/\epsilon)}{\epsilon^2}\Big) .$$

Theorem 1.1 follows from Theorems 3.4 and 4.1, together with the lower bound from [HLŠ07].

The mathematical heart of our analysis is to study the spectrum of the product of two reflections. The following lemma gives an "effective" spectral gap for two reflections applied to a vector. It is closely related to [Rei09, Theorem 8.7], but has a significantly simpler statement and proof.

**Lemma 4.2** (Effective spectral gap lemma)**.** *Let $\Pi$ and $\Lambda$ be projections, and let $R = (2\Pi - \mathbf{1})(2\Lambda - \mathbf{1})$ be the product of the reflections about their ranges. Let $\{|\beta\rangle\}$ be a complete orthonormal set of eigenvectors of $R$, with respective eigenvalues $e^{i\theta(\beta)}$, $\theta(\beta) \in (-\pi, \pi]$.*

*For any $\Theta \geq 0$, let $P_\Theta = \sum_{\beta:|\theta(\beta)|\leq\Theta} |\beta\rangle\langle\beta|$. If $\Lambda|w\rangle = 0$, then*

$$\|P_\Theta \Pi|w\rangle\| \leq \tfrac{\Theta}{2}\||w\rangle\| .$$

*Proof.* The claim can be shown via Jordan's Lemma [Jor75]; we give a direct proof. Let $|v\rangle = P_\Theta \Pi |w\rangle$, $|v'\rangle = (2\Lambda - \mathbf{1})|v\rangle$ and $|v''\rangle = (2\Pi - \mathbf{1})|v'\rangle = R|v\rangle$. When $\Theta$ is small, $|v\rangle$ and $|v''\rangle$ are close:

$$\||v\rangle - |v''\rangle\|^2 = \Big\| \sum_{\beta : |\theta(\beta)| \leq \Theta} (1 - e^{i\theta(\beta)}) \langle \beta | v \rangle |\beta\rangle \Big\|^2 \leq 2(1 - \cos\Theta) \||v\rangle\|^2 \leq \Theta^2 \||v\rangle\|^2 \ .$$

Notice that $|v\rangle + |v'\rangle$ is fixed by $\Lambda$. Similarly, $\Pi$ fixes $|v'\rangle + |v''\rangle$ and $\overline{\Pi} = \mathbf{1} - \Pi$ fixes $|v'\rangle - |v''\rangle$. Hence $0 = \langle v + v' | w \rangle = \langle v + v' | \Pi | w \rangle + \langle v + v' | \overline{\Pi} | w \rangle = \langle v + v'' | \Pi | w \rangle + \langle v - v'' | \overline{\Pi} | w \rangle$. Therefore, $\||v\rangle\|^2 = |\langle v | \Pi | w \rangle| = \frac{1}{2} |\langle v - v'' | \Pi | w \rangle + \langle v + v'' | \Pi | w \rangle| = \frac{1}{2} |\langle v - v'' | (\Pi - \overline{\Pi}) | w \rangle|$. We conclude

$$\||v\rangle\|^2 \leq \tfrac{1}{2} \||v\rangle - |v''\rangle\| \|(\Pi - \overline{\Pi})|w\rangle\| \leq \tfrac{\Theta}{2} \||v\rangle\| \||w\rangle\| \ . \qquad \square$$

We will also use a routine that, roughly, reflects about the eigenvalue-one eigenspace of a unitary:

**Theorem 4.3** (Phase detection [Kit95, MNRS07])**.** *For any $\Theta, \delta > 0$, there exists $b = O(\log \frac{1}{\delta} \log \frac{1}{\Theta})$ and, for any unitary $U \in \mathcal{L}(\mathcal{H})$, a quantum circuit $R(U)$ on $\mathcal{H} \otimes (\mathbf{C}^2)^{\otimes b}$ that makes at most $O\big(\frac{\log(1/\delta)}{\Theta}\big)$ controlled calls to $U$ and $U^{-1}$, and such that for any eigenvector $|\beta\rangle$ of $U$, with eigenvalue $e^{i\theta}$, $\theta \in (-\pi, \pi]$,*

- *If $\theta = 0$, then $R(U)|\beta\rangle \otimes |0^b\rangle = |\beta\rangle \otimes |0^b\rangle$.*
- *If $|\theta| > \Theta$, then $R(U)|\beta\rangle \otimes |0^b\rangle = -|\beta\rangle \otimes (|0^b\rangle + |\delta_\beta\rangle)$ for some vector $|\delta_\beta\rangle$ with $\||\delta_\beta\rangle\| < \delta$. Thus, if $|\gamma\rangle \in \mathcal{H}$ is orthogonal to all eigenvectors of $U$ with eigenvalues $e^{i\theta}$ for $|\theta| \leq \Theta$, then $\big\|(R(U) + \mathbf{1})|\gamma\rangle \otimes |0^b\rangle\big\| < \delta$.*

*$R(U)$ is constructed uniformly in the parameters $\Theta$ and $\delta$, and its structure does not depend on $U$.*

The phase-detection procedure can be constructed using, for example, standard phase estimation [Kit95, CEMM98, NWZ09]. Phase detection is a common subroutine in quantum algorithms, used implicitly or explicitly in, e.g., [Sze04, ACR+10, MNRS07, RŠ08, MNRS09, Rei09, Rei11].

Now we are ready to construct the algorithm to prove Theorem 4.1. Let $W = \gamma_2(\rho - \sigma|\Delta)$. Let $\{|u_{xj}\rangle\}$ and $\{|v_{xj}\rangle\}$, vectors in $\mathbf{C}^m$, be a solution to Eq. (3.1) for $\gamma_2(\rho - \sigma|\Delta)$. The first step is to turn this solution into a more natural geometric object. If the input alphabet size is $|D| = k$, let $|\mu_i\rangle, |\nu_i\rangle$ be the vectors given in Fact 2.4. Notice that we can rewrite the sum from Eq. (3.1) $\sum_{j \in [n]} (\Delta_j)_{x,y} \langle u_{xj} | v_{yj} \rangle = \sum_{j \colon x_j \neq y_j} \langle u_{xj} | v_{yj} \rangle$ as simply the inner product between the vectors $\sum_j |j\rangle |u_{xj}\rangle |\mu_{x_j}\rangle$ and $\frac{2(k-1)}{k} \sum_j |j\rangle |v_{yj}\rangle |\nu_{y_j}\rangle$. Our algorithm is based on these combined vectors.

Let $\mathcal{H}$ be the Hilbert space for the states $|\rho_x\rangle$ and $|\sigma_x\rangle$. For $y \in \mathcal{D}$, let $\Pi_y = \mathbf{1} - \sum_j |j\rangle\langle j| \otimes |\mu_{y_j}\rangle\langle \mu_{y_j}| \otimes \mathbf{1}_{\mathbf{C}^m}$. Also, define vectors $|t_{y\pm}\rangle, |\psi_y\rangle \in (\mathbf{C}^2 \otimes \mathcal{H}) \oplus (\mathbf{C}^n \otimes \mathbf{C}^k \otimes \mathbf{C}^m)$ by

$$|t_{y\pm}\rangle = \tfrac{1}{\sqrt{2}} \big( |0\rangle \otimes |\rho_y\rangle \pm |1\rangle \otimes |\sigma_y\rangle \big)$$
$$|\psi_y\rangle = \frac{\epsilon}{\sqrt{W}} |t_{y-}\rangle - \sum_{j \in [n]} |j\rangle \otimes |\mu_{y_j}\rangle \otimes |u_{yj}\rangle \ .$$

Let $\Lambda$ be the projection onto the orthogonal complement of the span of the vectors $\{|\psi_y\rangle\}_{y \in \mathcal{D}}$. Then our algorithm is given by:

> **Algorithm:** On input $x$, let $U_x = (2\Pi_x - \mathbf{1})(2\Lambda - \mathbf{1})$. Apply the phase-detection circuit $R(U_x)$, from Theorem 4.3, with precision $\Theta = \epsilon^2/W$ and error $\delta = \epsilon$, on input state $|0\rangle \otimes |\rho_x\rangle \otimes |0^b\rangle$. Output the result.

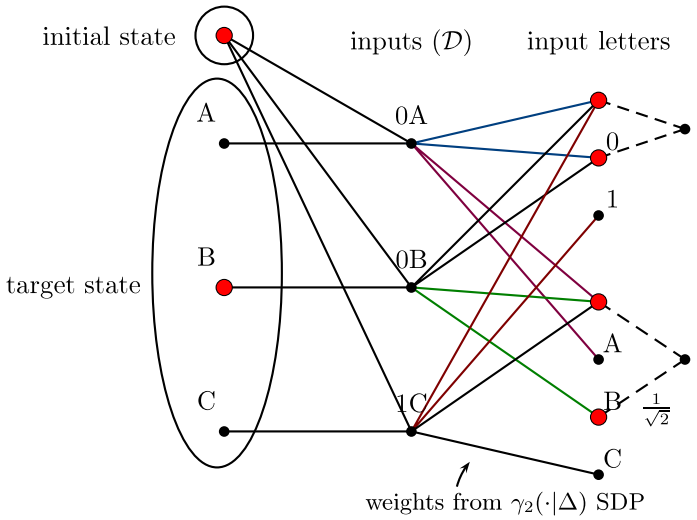


Figure 2: The algorithm can be interpreted as running a quantum walk on a bipartite graph with weighted biadjacency matrix $\Pi_x + \sum_y |y\rangle\langle\psi_y|$. Shown above is an example graph for the function $f(x_1x_2) = x_2$ mapping $\mathcal{D} = \{0A, 0B, 1C\} \subset \{0,1\} \times \{\mathrm{A},\mathrm{B},\mathrm{C}\}$ to $E = \{\mathrm{A},\mathrm{B},\mathrm{C}\}$.

Note that the reflection $2\Pi_x - \mathbf{1}$ can be computed with a query to the input oracle and its inverse—compute $x_j$, reflect in $|\mu_{x_j}\rangle$, then uncompute $x_j$. Therefore, the algorithm uses $O(W/\epsilon^2 \cdot \log \frac{1}{\epsilon})$ input queries. The algorithm is based on repeated reflections. As sketched in Figure 2, the algorithm can also be interpreted as a quantum walk on the bipartite graph with biadjacency matrix $\Pi_x + \sum_y |y\rangle\langle\psi_y|$.

To get some intuition for why this algorithm works, observe that the initial state $|0\rangle \otimes |\rho_x\rangle$ is $\frac{1}{\sqrt{2}}(|t_{x+}\rangle + |t_{x-}\rangle)$. The vector $|t_{x+}\rangle$ has large overlap with an eigenvalue-one eigenvector of $U_x$ (Claim 4.4 below), whereas the cumulative squared overlap of $|t_{x-}\rangle$ with eigenvectors of $U_x$ with small angle is small (Claim 4.5). The phase-detection procedure therefore approximately reflects the $|t_{x-}\rangle$ term, giving roughly $\frac{1}{\sqrt{2}}(|t_{x+}\rangle - |t_{x-}\rangle) = |1\rangle \otimes |\sigma_x\rangle$—our target state.

Now we give the formal analysis of the algorithm. Let $\{|\beta\rangle\}$ be a complete set of eigenvectors of $U_x$ with corresponding eigenvalues $e^{i\theta(\beta)}$, $\theta(\beta) \in (-\pi, \pi]$. For an angle $\Theta \geq 0$, let $P_\Theta = \sum_{\beta:|\theta(\beta)|\leq\Theta} |\beta\rangle\langle\beta|$, and $\overline{P}_\Theta = \mathbf{1} - P_\Theta$.

**Claim 4.4.** $\|P_0|t_{x+}\rangle\|^2 \geq 1 - \epsilon^2$.

*Proof.* We give a state $|\varphi\rangle$ such that $U_x|\varphi\rangle = |\varphi\rangle$ and $|\langle t_{x+}|\varphi\rangle|^2/\||\varphi\rangle\|^2 \geq 1 - \epsilon^2$. Let

$$|\varphi\rangle = |t_{x+}\rangle + \frac{1}{2}\frac{\epsilon}{\sqrt{W}}\frac{2(k-1)}{k}\sum_{j\in[n]} |j\rangle \otimes |\nu_{x_j}\rangle \otimes |v_{xj}\rangle \ .$$

Then $|\varphi\rangle$ is orthogonal to all $|\psi_y\rangle$, since $\langle t_{y-}|t_{x+}\rangle = \frac{1}{2}\left(\langle\rho_y|\rho_x\rangle - \langle\sigma_y|\sigma_x\rangle\right)$ implies

$$\langle\psi_y|\varphi\rangle = \frac{\epsilon}{\sqrt{W}}\langle t_{y-}|t_{x+}\rangle - \frac{1}{2}\frac{\epsilon}{\sqrt{W}}\sum_{j:x_j\neq y_j}\langle u_{yj}|v_{xj}\rangle = 0 \ .$$

Thus $\Lambda|\varphi\rangle = |\varphi\rangle$. Since $\Pi_x|t_{x+}\rangle = |t_{x+}\rangle$ and $\langle\mu_{x_j}|\nu_{x_j}\rangle = 0$ for all $j \in [n]$, also $\Pi_x|\varphi\rangle = |\varphi\rangle$. □

**Claim 4.5.** *For all* $\Theta \geq 0$, $\|P_\Theta|t_{x-}\rangle\|^2 \leq \frac{\Theta^2}{4}\left(\frac{W^2}{\epsilon^2}+1\right)$.

*Proof.* Apply Lemma 4.2 with $\Pi = \Pi_x$ and $|w\rangle = \frac{\sqrt{W}}{\epsilon}|\psi_x\rangle$. Then $\Delta|w\rangle = 0$ and $|t_{x-}\rangle = \Pi_x|w\rangle$. □

The following proposition completes the proof of Theorem 4.1:

**Proposition 4.6.** *If* $W \geq \epsilon$, *then with the parameters* $\delta = \epsilon$ *and* $\Theta = \epsilon^2/W$,

$$\left\|R(U_x)|0\rangle \otimes |\rho_x\rangle \otimes |0^b\rangle - |1\rangle \otimes |\sigma_x\rangle \otimes |0^b\rangle\right\| < 4\epsilon \ .$$

*Proof.* We have

$$\begin{aligned}\left\|R(U_x)|0\rangle|\rho_x\rangle|0^b\rangle - |1\rangle|\sigma_x\rangle|0^b\rangle\right\| &= \frac{1}{\sqrt{2}}\left\|R(U_x)(|t_{x+}\rangle + |t_{x-}\rangle)|0^b\rangle - (|t_{x+}\rangle - |t_{x-}\rangle)|0^b\rangle\right\| \\ &\leq \frac{1}{\sqrt{2}}\left\|(R(U_x)-\mathbf{1})|t_{x+}\rangle|0^b\rangle\right\| + \frac{1}{\sqrt{2}}\left\|(R(U_x)+\mathbf{1})|t_{x-}\rangle|0^b\rangle\right\| \ .\end{aligned}$$

By Theorem 4.3, the first term equals $\frac{1}{\sqrt{2}}\|(R(U_x)-\mathbf{1})\overline{P}_0|t_{x+}\rangle|0^b\rangle\| \leq \sqrt{2}\|\overline{P}_0|t_{x+}\rangle\| \leq \sqrt{2}\epsilon$, by Claim 4.4. The second term is at most $\frac{1}{\sqrt{2}}\|(R(U_x)+\mathbf{1})\overline{P}_\Theta|t_{x-}\rangle|0^b\rangle\| + \sqrt{2}\|P_\Theta|t_{x-}\rangle\| < \frac{\delta}{\sqrt{2}} + \frac{\Theta}{\sqrt{2}}\sqrt{\frac{W^2}{\epsilon^2}+1}$, by Theorem 4.3 and Claim 4.5. Now substitute our choices of parameters and use $\frac{W^2}{\epsilon^2}+1 \leq 2W^2/\epsilon^2$ to conclude the proof. □

Notice that the constant factor hidden by the big-$O$ notation in Theorem 4.1 is the same as the constant hidden in Theorem 4.3 for the number of calls to $U$ and $U^{-1}$, and is less than 100.

## 4.2 Lower bound for state conversion

We now show how $\gamma_2(\rho-\sigma|\Delta)$ can be used to show query complexity lower bounds for the state-conversion problem. The argument has two parts. First, we show that $\gamma_2(\rho-\sigma|\Delta)$ lower-bounds the complexity of *exactly* converting $\rho$ to $\sigma$. Second, we develop an output condition constraining those $\sigma'$ that are viable final Gram matrices of a successful algorithm with error $\epsilon$. These two parts have been present in all previous adversary arguments, but the separation has not been fully recognized.

Once these two steps are finished, the lower bound naturally becomes $\min_{\sigma' \approx_\epsilon \sigma} \gamma_2(\rho-\sigma'|\Delta)$, where the notion of approximation is given by the output condition. This paradigm follows the use of approximation norms for lower bounds in communication complexity [LS07].

We begin with the lower bound for exact state conversion:

**Lemma 4.7.** *Suppose that* $\sigma$ *can be reached from* $\rho$ *with one query. Then* $\gamma_2(\rho-\sigma|\Delta) \leq 2$.

*Proof.* Let $\Gamma_j$ project onto the query register containing index $j$. For $x \in \mathcal{D}$, let $O_x$ be the unitary query oracle. It satisfies $O_x^\dagger O_y \Gamma_j = \Gamma_j$ when $x_j = y_j$. By assumption, $\sigma_{x,y} = \langle\rho_x|O_x^\dagger O_y|\rho_y\rangle$. Then

$$(\rho-\sigma)_{x,y} = \sum_j \langle\rho_x|\Gamma_j|\rho_y\rangle - \langle\rho_x|O_x^\dagger O_y\Gamma_j|\rho_y\rangle = \sum_{j:x_j\neq y_j} \langle\rho_x|\Gamma_j|\rho_y\rangle - \langle\rho_x|O_x^\dagger O_y\Gamma_j|\rho_y\rangle \ .$$

Now define $|u_{xj}\rangle = (\Gamma_j|\rho_x\rangle, O_x\Gamma_j|\rho_x\rangle)$ and $|v_{xj}\rangle = (\Gamma_j|\rho_x\rangle, -O_x\Gamma_j|\rho_x\rangle)$. Then $\langle u_{xj}|v_{yj}\rangle = \langle\rho_x|(\mathbf{1} - O_x^\dagger O_y)\Gamma_j|\rho_y\rangle$, as desired. Furthermore, $\sum_j \||u_{xj}\rangle\|^2 = \sum_j \||v_{xj}\rangle\|^2 = 2$. □

Lower bounds for approximate query problems follow by combining this lemma with appropriate output conditions. For example, in the functional case, one can use a condition based on $\ell_\infty$ distance [Amb02], or the full output condition from [BSS03]. The output condition traditionally used for the general adversary method is based on the $\gamma_2$ norm [HLŠ07]. This condition has the advantage that it is an SDP, it extends to state conversion, and, as we now show, it is tight.

**Lemma 4.8.** *Let $\{|\rho_x\rangle\}, \{|\sigma_x\rangle\} \subset \mathcal{H}$ be finite sets of vectors with the same index set, and let $\rho, \sigma$ be their respective Gram matrices. Then*

- *If $\mathrm{Re}(\langle\rho_x|\sigma_x\rangle) \geq \sqrt{1-\epsilon}$ for every $x$, then $\gamma_2(\rho-\sigma) \leq 2\sqrt{\epsilon}$ [HLŠ07].*
- *If $\gamma_2(\rho-\sigma) \leq \epsilon$, then there exists a unitary $U$ such that $\langle\rho_x|U|\sigma_x\rangle \geq 1-\sqrt{2\epsilon}$ for all $x$.*

The second item has recently been improved by [LR11] to $\gamma_2(\rho-\sigma) \leq \epsilon$ implies there exists a unitary $U$ such that $\langle\rho_x|U|\sigma_x\rangle \geq 1-\epsilon/2$ for all $x$.

*Proof.* For the first part of the lemma, we can factorize $\rho-\sigma$ as

$$(\rho-\sigma)_{x,y} = \tfrac{1}{2}\left(\langle\rho_x+\sigma_x|\rho_y-\sigma_y\rangle + \langle\rho_x-\sigma_x|\rho_y+\sigma_y\rangle\right) \ .$$

Thus by a triangle inequality $\gamma_2(\rho-\sigma) \leq \max_{x,y} \|\rho_x+\sigma_x\|\|\rho_y-\sigma_y\| \leq 2\max_y \sqrt{2-2\mathrm{Re}(\langle\rho_y|\sigma_y\rangle)} \leq 2\sqrt{\epsilon}$. For the last inequality we used $\sqrt{1-\epsilon} \geq 1-\frac{\epsilon}{2}$.

To prove the second part, let $\{u_x\}$ and $\{v_x\}$ be arbitrary factorizations of $\rho$ and $\sigma$, respectively. As $\gamma_2(\rho-\sigma) \leq \epsilon$, there exists a factorization $(\rho-\sigma)_{x,y} = \langle\alpha_x|\beta_y\rangle$ with $\|\alpha_x\|, \|\beta_y\| \leq \sqrt{\epsilon}$. Then

$$\begin{aligned}\langle v_x|v_y\rangle &= \langle u_x|u_y\rangle - \langle\alpha_x|\beta_y\rangle \\ &= \langle u_x|u_y\rangle - \frac{1}{2}\langle\alpha_x|\beta_y\rangle - \frac{1}{2}\langle\beta_x|\alpha_y\rangle\end{aligned}$$

as $\rho-\sigma$ is Hermitian. Let $p_x = \frac{1}{2}(\alpha_x-\beta_x)$ and $q_x = \frac{1}{2}(\alpha_x+\beta_x)$. Then the previous equation implies

$$\langle(u_x,p_x)|(u_y,p_y)\rangle = \langle(v_x,q_x)|(v_y,q_y)\rangle \ .$$

By unitary freedom of square roots, if $AA^\dagger = BB^\dagger$ for any two matrices $A$ and $B$, then $AU = B$ for some unitary $U$. So there is a unitary $U$ such that $(u_x,p_x)U = (v_x,q_x)$. As $\langle\alpha_x|\beta_x\rangle = 0$ because $\rho-\sigma$ has zeros on the diagonal, we have $\|p_x\|^2 = \|q_x\|^2 \leq \epsilon/2$.

$$\begin{aligned}\langle(u_x,0)U|(v_x,0)\rangle &= \langle(u_x,p_x)U|(v_x,q_x)\rangle - \langle(u_x,0)U|(0,q_x)\rangle - \langle(0,p_x)U|(v_x,0)\rangle - \langle(0,p_x)U|(0,q_x)\rangle \\ &\geq 1 + \|q_x\|^2 - \|p_x\| - \|q_x\| - \|p_x\|\|q_x\| \\ &\geq 1-\sqrt{2\epsilon} \ .\end{aligned}$$

□

Based on Lemma 4.8, we immediately derive tight SDPs for the query complexities of state conversion:

**Theorem 4.9.** *For $\delta > 0$, let*

$$\begin{aligned}q_\delta(\rho,\sigma) &= \min_{\sigma'\succeq 0} \left\{\gamma_2(\rho-\sigma'|\Delta) : \gamma_2(\sigma'-\sigma) \leq \delta\right\} \\ q_\delta^{nc}(\rho,\sigma) &= \min_{\sigma',S\succeq 0} \left\{\gamma_2(\rho-\sigma'|\Delta) : \gamma_2(\sigma'-\sigma\circ S) \leq \delta,\ S\circ\mathbf{1} = \mathbf{1}\right\} \ .\end{aligned} \tag{4.1}$$

*Then the bounded-error coherent and non-coherent state-conversion query complexities satisfy*

$$\begin{aligned} \Omega\Big(q_{2\sqrt{2\epsilon}}(\rho,\sigma)\Big) \leq Q_\epsilon(\rho,\sigma) \leq O\Big(q_{\epsilon^4/16}(\rho,\sigma)\frac{\log(1/\epsilon)}{\epsilon^2}\Big) \\ \Omega\Big(q^{nc}_{2\sqrt{2\epsilon}}(\rho,\sigma)\Big) \leq Q^{nc}_\epsilon(\rho,\sigma) \leq O\Big(q^{nc}_{\epsilon^4/16}(\rho,\sigma)\frac{\log(1/\epsilon)}{\epsilon^2}\Big) \ . \end{aligned} \tag{4.2}$$

Thus for coherent state conversion, the output condition used is $\gamma_2(\sigma'-\sigma) \leq \delta$ for an appropriately chosen $\delta$, and in the non-coherent case, optimization over additional garbage states is allowed.

For well-behaved problems, i.e., problems satisfying $Q_{1/3}(\rho,\sigma) = O\big(Q_\epsilon(\rho,\sigma)\log(1/\epsilon)\big)$ as in the functional case, this is true characterization. General state-conversion problems, however, do not necessarily satisfy this robustness condition. Just as the complexity of a boolean function can have a precipitous change around error $1/2$, state-conversion problems can have non-continuous changes in complexity even around small values of $\epsilon$. For such problems Theorem 4.9 may not be a true characterization because of the gap in error parameters on the left- and right-hand sides. The gap in the error dependence arises from the looseness of the necessary and sufficient conditions in Lemma 4.8, plus the error from Theorem 4.1. We do not know if the $\epsilon$-dependence in Theorem 4.1 can be improved to polylogarithmic in $1/\epsilon$.

An advantage of using the $\gamma_2$ output condition is that the quantities $q_\delta$ and $q^{nc}_\delta$ are described by semi-definite programs. One could define analogous quantities with other output conditions, however, including the "true" output condition given by Definition 2.1 and Definition 2.2. In this case the only slack in the characterization would arise from Theorem 4.1 and thus the error parameters on left- and right-hand sides would agree up to constant factors.

Ambainis et al. [AMRR11], previously extended both the general adversary bound and the multiplicative adversary bound [Špa08] to the state-generation problem. A difference between our work and theirs is that we separate the bound for the exact problem from the output condition used to handle the bounded error case. [AMRR11] focus on the output condition introduced by [Špa08] with the multiplicative adversary method and show how to extend the additive adversary method with this output condition to the state generation problem, calling this the hybrid adversary method. They show that the hybrid adversary method dominates the general adversary method, and that the hybrid adversary method is dominated by the bound of [Špa08] extended to the case of state generation. We do not know if the hybrid adversary method also dominates the $q_\delta(J,\sigma)$ measure. The proof in [AMRR11] that the multiplicative method dominates the hybrid method actually shows that the multiplicative method for exact state generation dominates the $\gamma_2(J-\sigma|\Delta)$ measure, as explicitly shown by [LR11]. Thus the multiplicative method will dominate the $\gamma_2(J-\sigma|\Delta)$ bound whenever they are paired with the same output condition.

This line of research into discrete query complexity was launched by the discovery of a continuous-time query algorithm for evaluating AND-OR formulas [FGG08]. We now complete the circle:

**Theorem 4.10.** *The bounded-error continuous-time and discrete query models are equivalent.*

Cleve et al. [CGM+09] have shown that the models are equivalent up to a sub-logarithmic factor. The proof of Theorem 4.10, given in Appendix B, follows from our algorithm, Theorem 4.1, together with the observation that the general adversary bound remains a lower bound for continuous-time query algorithms. The latter result has been observed by Yonge-Mallo in 2007 [YM11] and, independently, Landahl (personal communication).

## 5 Function composition

In this section we show that the adversary method behaves well with respect to function composition, extending previous work for the boolean case [HLŠ07, Rei09]. Let $g : \mathcal{C} \to D$ where $\mathcal{C} \subseteq C^m$ and $f : \mathcal{D} \to E$ $(\mathcal{D} \subseteq D^n)$ for finite sets $C, D, E$. Define the composed function $f \circ g^n$ by

$$(f \circ g^n)(x) = f\big(g(x_1, \ldots, x_m), \ldots, g(x_{(n-1)m+1}, \ldots, x_{mn})\big) \ .$$

**Lemma 5.1.** *Letting* $G = \{\delta_{g(x),g(y)}\}_{x,y}$, $\mathrm{Adv}^{\pm}(f \circ g^n) \leq \mathrm{Adv}^{\pm}(f)\, \gamma_2(J - G|\Delta)$.

The proof of this lemma follows in the natural way. We take optimal solutions to the $\mathrm{Adv}^{\pm}(f)$ program and the $\gamma_2(J - G|\Delta)$ program, and form their tensor product to construct a solution to the composed program. This proof strategy does not directly work for the general adversary bound—we crucially use the extra constraints present in the $\gamma_2(J - G|\Delta)$ program. In fact, this lemma can be seen as a special case of the more general inequality $\gamma_2(A|Z) \leq \gamma_2(A|Y) \max_j \gamma_2(Y_j|Z)$ (Lemma A.2). The proof is given in Appendix C.

In the case where all the functions $f$ and $g$ have boolean inputs and outputs, a matching lower bound to Lemma 5.1 has been shown by Høyer et al. [HLŠ07]. In general, we cannot always show such a matching lower bound. For example, let $g$ be a function that only outputs even numbers, and let $f$ output the sum of its inputs modulo two; then $f \circ g^n$ is constant. We can, however, show a matching composition lower bound when the range of $g$ is boolean:

**Lemma 5.2.** *Let* $g : \mathcal{C} \to \{0, 1\}$ *and* $f : \{0, 1\}^n \to E$. *Then* $\mathrm{Adv}^{\pm}(f \circ g^n) \geq \mathrm{Adv}^{\pm}(f)\mathrm{Adv}^{\pm}(g)$.

The proof is given in Appendix C. The above composition lemmas also lead to direct-sum results for quantum query complexity:

**Corollary 5.3.** *Let* $g : \mathcal{D} \to E$, *and let* $g^n : \mathcal{D}^n \to E^n$ *consist of* $n$ *independent copies of* $g$, *given by* $g^n(x^1, \ldots, x^n) = \big(g(x^1), \ldots, g(x^n)\big)$. *Then*

$$Q(g^n) = \Theta\big(n\, Q(g)\big) \ . \tag{5.1}$$

The lower bound $\mathrm{Adv}^{\pm}(g^n) \geq n\, \mathrm{Adv}^{\pm}(g)$ has been shown by [ACLT10]. The corresponding upper bound $\mathrm{Adv}^{\pm}(g^n) \leq n\, \mathrm{Adv}^{\pm}(g)$ is a special case of Lemma 5.1, with $f$ the identity function. Corollary 5.3 then follows from Theorem 1.1. Let us remark that when $E = \{0, 1\}$, the upper bound $Q(f^n) = O\big(n\, Q(f)\big)$ follows from the robust input recovery quantum algorithm [BNRW07, Theorem 3]. The same algorithm can be generalized to handle larger $E$.

### Acknowledgements

We thank Jérémie Roland and Miklos Santha for many helpful conversations on these topics, and thank Richard Cleve and Ronald de Wolf for useful comments on an earlier draft. Part of this work was done while T.L. was at Rutgers University, supported by an NSF postdoctoral fellowship and grant CCF-0728937. R.M. acknowledges support from NSERC and NSF grant CCF-0832787. B.R. acknowledges support from NSERC, ARO-DTO and MITACS.

## References


[ACLT10] Andris Ambainis, Andrew M. Childs, François Le Gall, and Seiichiro Tani. The quantum query complexity of certification. *Quantum Inf. Comput.*, 10:181–188, 2010, arXiv:0903.1291 [quant-ph].

[ACR+10] Andris Ambainis, Andrew M. Childs, Ben W. Reichardt, Robert Špalek, and Shengyu Zhang. Any AND-OR formula of size $N$ can be evaluated in time $N^{1/2+o(1)}$ on a quantum computer. *SIAM J. Comput.*, 39(6):2513–2530, 2010. Earlier version in FOCS'07.

[Amb02] Andris Ambainis. Quantum lower bounds by quantum arguments. *J. Comput. Syst. Sci.*, 64:750–767, 2002, arXiv:quant-ph/0002066. Earlier version in STOC'00.

[AMRR11] Andris Ambainis, Loïck Magnin, Martin Roetteler, and Jérémie Roland. Symmetry-assisted adversaries for quantum state generation. In *26th IEEE Conference on Computational Complexity*, pages 167–177, 2011, arXiv:1012.2112 [quant-ph].

[Bha07] Rajendra Bhatia. *Positive Definite Matrices*. Princeton University Press, Princeton, 2007.

[BNRW07] Harry Buhrman, Ilan Newman, Hein Röhrig, and Ronald de Wolf. Robust polynomials and quantum algorithms. *Theory Comput. Syst.*, 40(4):379–395, 2007, arXiv:quant-ph/0309220. Earlier version in STACS'05.

[BSS03] Howard Barnum, Michael Saks, and Mario Szegedy. Quantum query complexity and semidefinite programming. In *Proc. 18th IEEE Complexity*, pages 179–193, 2003.

[BW02] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.

[CCJY09] Andrew M. Childs, Richard Cleve, Stephen P. Jordan, and David Yeung. Discrete-query quantum algorithm for NAND trees. *Theory of Computing*, 5:119–123, 2009, arXiv:quant-ph/0702160.

[CEMM98] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. *Proc. R. Soc. London A*, 454(1969):339–354, 1998, arXiv:quant-ph/9708016.

[CGM+09] Richard Cleve, Daniel Gottesman, Michele Mosca, Rolando D. Somma, and David L. Yonge-Mallo. Efficient discrete-time simulations of continuous-time quantum query algorithms. In *Proc. 41st ACM STOC*, pages 409–416, 2009, arXiv:0811.4428 [quant-ph].

[FGG08] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum algorithm for the Hamiltonian NAND tree. *Theory of Computing*, 4:169–190, 2008, arXiv:quant-ph/0702144.

[HLŠ07] Peter Høyer, Troy Lee, and Robert Špalek. Negative weights make adversaries stronger. In *Proc. 39th ACM STOC*, pages 526–535, 2007, arXiv:quant-ph/0611054.

[Jor75] Camille Jordan. Essai sur la géométrie à $n$ dimensions. *Bulletin de la S. M. F.*, 3:103–174, 1875.

[Kit95] A. Yu. Kitaev. Quantum measurements and the Abelian stabilizer problem. 1995, arXiv:quant-ph/9511026.

[LMSS07] Nati Linial, Shahar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27:439–463, 2007.

[Lov03] László Lovász. Semidefinite programs and combinatorial optimization. In B. A. Reed and C. Linhares Sales, editors, *Recent Advances in Algorithms and Combinatorics*, volume 11 of *CMS Books Math.*, pages 137–194. Springer, 2003.

[LR11] Troy Lee and Jérémie Roland. A strong direct product theorem for quantum query complexity. 2011, arXiv:1104.4468 [quant-ph].

[LS07] Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–399, 2007.

[LS09] Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34:368–394, 2009. Earlier version in STOC'07.

[LSŠ08] Troy Lee, Adi Shraibman, and Robert Špalek. A direct product theorem for discrepancy. In *Proc. 23rd Conference on Computational Complexity*, pages 71–80, 2008.

[MNRS07] Frédéric Magniez, Ashwin Nayak, Jérémie Roland, and Miklos Santha. Search via quantum walk. In *Proc. 39th ACM STOC*, pages 575–584, 2007, arXiv:quant-ph/0608026.

[MNRS09] Frédéric Magniez, Ashwin Nayak, Peter C. Richter, and Miklos Santha. On the hitting times of quantum versus random walks. In *Proc. 20th ACM-SIAM SODA*, pages 86–95, 2009, arXiv:0808.0084 [quant-ph].

[NWZ09] Daniel Nagaj, Pawel Wocjan, and Yong Zhang. Fast amplification of QMA. *Quantum Inf. Comput.*, 9:1053–1068, 2009, arXiv:0904.1549 [quant-ph].

[Rei09] Ben W. Reichardt. Span programs and quantum query complexity: The general adversary bound is nearly tight for every boolean function. 2009, arXiv:0904.2759 [quant-ph]. Extended abstract in *Proc. 50th IEEE FOCS*, pages 544–551, 2009.

[Rei11] Ben W. Reichardt. Reflections for quantum query algorithms. In *Proc. 22nd ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 560–569, 2011, arXiv:1005.1601 [quant-ph].

[RŠ08] Ben W. Reichardt and Robert Špalek. Span-program-based quantum algorithm for evaluating formulas. In *Proc. 40th ACM STOC*, pages 103–112, 2008, arXiv:0710.2630 [quant-ph].

[Shi02] Yaoyun Shi. Quantum lower bounds for the collision and the element distinctness problems. In *Proc. 43rd IEEE FOCS*, pages 513–519, 2002, arXiv:quant-ph/0112086.

[Špa08] Robert Špalek. The multiplicative quantum adversary. In *Proc. 23rd IEEE Complexity*, pages 237–248, 2008, arXiv:quant-ph/0703237.

[Sze04] Mario Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proc. 45th IEEE FOCS*, pages 32–41, 2004.

[Tsi87] B. S. Tsirelson. Quantum analogues of the Bell inequalities: The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987.

[Ung08] Falk Unger. *Noise in classical and quantum computation and non-locality.* PhD thesis, University of Amsterdam and CWI, 2008.

[YM11] David Yonge-Mallo. Adversary lower bounds in the Hamiltonian oracle model. 2011, arXiv:1108.2479 [quant-ph].

Centre for Quantum Technologies
*E-mail address:* troyjlee@gmail.com

Institute for Quantum Computing, University of Waterloo
*E-mail address:* r3mittal@uwaterloo.ca

Institute for Quantum Computing, University of Waterloo
*E-mail address:* breic@iqc.ca

Google, Inc.
*E-mail address:* spalek@google.com

Rutgers University
*E-mail address:* szegedy@cs.rutgers.edu

# A Properties of the filtered $\gamma_2$ norm

For reference, we list several useful properties of the filtered $\gamma_2$ norm. First, we give an alternative formulation that explains why $\gamma_2$ is also known as the Schur product operator norm:

**Lemma A.1.** *The $\gamma_2$ and filtered $\gamma_2$ norms can equivalently be expressed by*

$$\gamma_2(A) = \max_M \left\{ \|A \circ M\| : \|M\| \leq 1 \right\} \tag{A.1}$$

$$\gamma_2(A|Z) = \max_M \left\{ \|A \circ M\| : \max_j \|Z_j \circ M\| \leq 1 \right\} . \tag{A.2}$$

*Proof.* Both of these equations can be proven in the same way: Start with Eq. (2.1) or (3.1) for $\gamma_2(A)$ or $\gamma_2(A|Z)$, respectively, and take the dual. The semi-definite program given by Eq. (2.1) is always strictly feasible and that of Eq. (3.1) is strictly feasible provided that whenever $A_{x,y} \neq 0$ there is some $j$ with $(Z_j)_{x,y} \neq 0$, i.e., provided the maximum in (A.2) is finite. Thus by the duality principle [Lov03, Theorem 3.4] the primal and dual formulations are equal and the optimum is achieved. □

**Lemma A.2.** *The quantity $\gamma_2(\cdot|Z)$ is a norm when restricted to arguments supported on the union of the supports of the $Z_j$. For matrices $B, Y_1, \ldots, Y_n$ of the appropriate dimensions it satisfies:*

1. *If $A \neq 0$, then $\gamma_2(A|\{A\}) = 1$. For $J$ the all-ones matrix, $\gamma_2(A|\{J\}) = \gamma_2(A)$.*

2. *$\gamma_2(A|Z) = 0$ if and only if $A = 0$. $\gamma_2(A|Z) = \infty$, i.e., Eq. (3.1) is infeasible, if and only if there exists an entry $(x, y)$ such that $A_{x,y} \neq 0$ and $(Z_j)_{x,y} = 0$ for all $j$.*

3. *Positive scalability: $\gamma_2(sA|Z) = |s|\gamma_2(A|Z)$ and $\gamma_2(A|\{sZ_1, \ldots, sZ_n\}) = \frac{1}{|s|}\gamma_2(A|Z)$ for $s \neq 0$.*

4. *Triangle inequality: $\gamma_2(A + B|Z) \leq \gamma_2(A|Z) + \gamma_2(B|Z)$.*

5. *$\gamma_2(A|Z)$ is invariant under duplicating corresponding rows or columns of $A$ and all $Z_j$.*

6. *$\gamma_2(A|Y \cup Z) \leq \gamma_2(A|Z)$. This is an equality if each $Y_i$ is a restriction of some $Z_j$ to a rectangular submatrix.*

7. *Provided $\sum_j |p_j| = 1$, $\gamma_2(A|Z) = \gamma_2(A|Z \cup \{\sum_j p_j Z_j\})$. (Thus the second argument in $\gamma_2(A|Z)$ can be thought of as a convex set centered at the origin, where only the extremal points matter.)*

8. *If the supports of $Z_1$ and $Z_2$ are row- and column-disjoint, then $\gamma_2(A|Z) = \gamma_2(A|\{Z_1 + Z_2, Z_3, \ldots, Z_n\})$.*

9. *$\gamma_2(A \circ B|Z) \leq \gamma_2(A|Z)\gamma_2(B)$.*

10. *$\gamma_2(A \circ B|\{Z_j \circ B\}) \leq \gamma_2(A|Z) \leq \gamma_2(A|\{Z_j \circ B\})\gamma_2(B)$.*

11. *A composition property: $\gamma_2(A|Z) \leq \gamma_2(A|Y) \max_j \gamma_2(Y_j|Z)$.*

12. *A direct-sum property: $\gamma_2(A \oplus B|\{Y_j \oplus Z_j\}) = \max\{\gamma_2(A|Y), \gamma_2(B|Z)\}$.*

13. *A tensor-product property: $\gamma_2(A \otimes B|Y \otimes Z) = \gamma_2(A|Y)\gamma_2(B|Z)$, where $Y \otimes Z = \{Y_i \otimes Z_j\}$, all pairwise tensor products.*

*Proof.* By items (2), (3) and (4), $\gamma_2(\cdot|Z)$ is a norm on arguments restricted to the support of the $Z_j$. The proofs of the first three properties follow easily from the definition of filtered $\gamma_2$ norm, Eq. (3.1). Therefore we begin by showing the triangle inequality, property (4).

4. Given optimal vector solutions to Eq. (3.1) for $\gamma_2(A|Z)$ and for $\gamma_2(B|Z)$, simply concatenate corresponding vectors to obtain a solution for $\gamma_2(A + B|Z)$, with objective value at most $\gamma_2(A|Z) + \gamma_2(B|Z)$.

5. Invariance of $\gamma_2(A|Z)$ under copying rows follows by copying the associated solution vectors.

6. Any solution to Eq. (A.2) for $\gamma_2(A|Y \cup Z)$ also works for $\gamma_2(A|Z)$; hence $\gamma_2(A|Z) \geq \gamma_2(A|Y \cup Z)$. However, if $Y_i$ is a submatrix restriction of $Z_j$ then the constraint $\|Y_i \circ M\| \leq 1$ is redundant to $\|Z_j \circ M\| \leq 1$; hence adding $Y_i$ to $Z$ does not affect $\gamma_2(A|Z)$.

7. If $\max_j \|Z_j \circ M\| \leq 1$, then $\left\| \sum_j p_j Z_j \circ M \right\| \leq 1$; again, the new constraint is redundant.

8. Assuming without loss of generality that in a solution to $\gamma_2(A|Z)$ the vectors $|u_{xj}\rangle$ (respectively, $|v_{yj}\rangle$) are nonzero only on rows (columns) where $Z_j$ has nonzero entries, concatenating the vectors for $Z_1$ and for $Z_2$ gives a solution to $\gamma_2(A|\{Z_1 + Z_2, Z_3, \ldots, Z_n\})$. Thus $\gamma_2(A|\{Z_1 + Z_2, Z_3, \ldots, Z_n\}) \leq \gamma_2(A|Z)$. For the other direction, divide the vectors for $Z_1 + Z_2$ according to whether they correspond to a nontrivial row or column of $Z_1$, or of $Z_2$.

9. Begin with an optimal solution $\{|u_x\rangle, |v_y\rangle\}$ to Eq. (2.1) for $\gamma_2(B)$, and an optimal solution $\{|u_{xj}\rangle, |v_{yj}\rangle\}$ to Eq. (3.1) for $\gamma_2(A|Z)$. The tensor products $\{|u_{xj}\rangle \otimes |u_x\rangle, |v_{yj}\rangle \otimes |v_y\rangle\}$ give a solution for $\gamma_2(A \circ B|Z)$, with objective value at most $\gamma_2(A|Z)\gamma_2(B)$.

10. The second inequality in property (10) works in the same way as (9); the tensor product of vector solutions for $\gamma_2(B)$ and $\gamma_2(A|\{Z_j \circ B\})$ is a solution for $\gamma_2(A|Z)$. The first inequality follows since any vector solution for $\gamma_2(A|Z)$ also works for $\gamma_2(A \circ B|\{Z_j \circ B\})$.

11. Let $\{|u_{xj}\rangle, |v_{yj}\rangle\}$ be an optimal solution to Eq. (3.1) for $\gamma_2(A|Y)$ and for each $j$ let $\{|u^j_{xi}\rangle, |v^j_{yi}\rangle\}$ be an optimal solution for $\gamma_2(Y_j|Z)$. These vectors satisfy

$$\gamma_2(A|Y) \geq \max\Big\{\sum_j \||u_{xj}\rangle\|^2, \sum_j \||v_{yj}\rangle\|^2\Big\} \quad \gamma_2(Y_j|Z) \geq \max\Big\{\sum_i \||u^j_{xi}\rangle\|^2, \sum_i \||v^j_{yi}\rangle\|^2\Big\}$$
$$A_{x,y} = \sum_j (Y_j)_{x,y}\langle u_{xj}|v_{yj}\rangle \qquad (Y_j)_{x,y} = \sum_i (Z_i)_{x,y}\langle u^j_{xi}|v^j_{yi}\rangle \ .$$

Combining the last two equations gives $A_{x,y} = \sum_i (Z_i)_{x,y} \sum_j \langle u_{xj}|v_{yj}\rangle\langle u^j_{xi}|v^j_{yi}\rangle$. Thus the vectors $\oplus_j(|u_{xj}\rangle \otimes |u^j_{xi}\rangle)$ and $\oplus_j(|v_{yj}\rangle \otimes |v^j_{yi}\rangle)$ give a solution for $\gamma_2(A|Z)$, with objective value at most $\max\{\max_x \sum_{i,j} \||u_{xj}\rangle\|^2 \||u^j_{xi}\rangle\|^2, \max_y \sum_{i,j} \||v_{yj}\rangle\|^2 \||v^j_{yi}\rangle\|^2\} \leq \gamma_2(A|Y) \max_j \gamma_2(Y_j|Z)$.

12. A union of the vectors for $\gamma_2(A|Y)$ and $\gamma_2(B|Z)$ gives a vector solution for $\gamma_2(A \oplus B|Y_j \oplus Z_j)$.

13. The inequality $\gamma_2(A \otimes B|Y \otimes Z) \leq \gamma_2(A|Y)\gamma_2(B|Z)$ is straightforward; if $\{|u_{xi}\rangle, |v_{yi}\rangle\}$ form an optimal vector solution for $\gamma_2(A|Y)$ and $\{|\mu_{\alpha j}\rangle, |\nu_{\beta j}\rangle\}$ form an optimal vector solution for $\gamma_2(B|Z)$, then the vectors $|u_{(x,\alpha)(i,j)}\rangle = |u_{xi}\rangle \otimes |\mu_{\alpha j}\rangle$ and $|v_{(y,\beta)(i,j)}\rangle = |v_{yi}\rangle \otimes |\nu_{\beta j}\rangle$ satisfy

$$\sum_{i,j} (Y_i \otimes Z_j)_{(x,\alpha),(y,\beta)}\langle u_{(x,\alpha)(i,j)}|v_{(y,\beta)(i,j)}\rangle = \sum_i (Y_i)_{x,y}\langle u_{xi}|v_{yi}\rangle \sum_j (Z_j)_{\alpha,\beta}\langle \mu_{\alpha j}|\nu_{\beta j}\rangle = A_{x,y}B_{\alpha,\beta}$$

and therefore give a solution for $\gamma_2(A \otimes B|Y \otimes Z)$, with objective value at most $\gamma_2(A|Y)\gamma_2(B|Z)$.

For the other direction of the tensor-product inequality, let $M$ be an optimal solution to the dual SDP Eq. (A.2) for $\gamma_2(A|Y)$ and let $N$ be an optimal solution to the dual SDP for $\gamma_2(B|Z)$. We claim that $M \otimes N$ is a solution to the dual SDP for $\gamma_2(A \otimes B|Y \otimes Z)$. Indeed, for all $i, j$, $\|(Y_i \otimes Z_j) \circ (M \otimes N)\| = \|(Y_i \circ M) \otimes (Z_j \circ N)\| = \|Y_i \circ M\|\|Z_j \circ N\| \leq 1$. The objective value is $\|(A \otimes B) \circ (M \otimes N)\| = \gamma_2(A|Y)\gamma_2(B|Z)$. Thus $\gamma_2(A|Y)\gamma_2(B|Z) \leq \gamma_2(A \otimes B|Y \otimes Z)$. □

For completeness, we also present the dual norm $\gamma_2^*(\,\cdot\,|Z)$. Let $\hat{A} = \left[\begin{smallmatrix} 0 & A \\ A^\dagger & 0 \end{smallmatrix}\right]$. Then

$$\begin{aligned} \gamma_2^*(A|Z) &= \max_{B:\gamma_2(B|Z)=1} \langle A, B\rangle \\ &= \max_{\{Y_j \succeq 0\}} \Big\{\frac{1}{2}\sum_j \langle Y_j, \hat{Z}_j \circ \hat{A}\rangle : \sum_j Y_j \circ \mathbf{1} = \mathbf{1}\Big\} \\ &= \min_{\Omega} \Big\{\tfrac{1}{2}\mathrm{Tr}\,\Omega : \Omega \circ \mathbf{1} = \Omega \text{ and } \forall j,\ \Omega - \hat{A} \circ \hat{Z}_j \succeq 0\Big\} \ . \end{aligned} \tag{A.3}$$

When $A$ and the $Z_j$ are Hermitian, then $\gamma_2^*(A|Z) = \min_\Omega\{\mathrm{Tr}\,\Omega : \Omega \circ \mathbf{1} = \Omega \text{ and } \forall j,\ \Omega \pm A \circ Z_j \succeq 0\}$, a slightly simpler form that we will use below in Appendix C. The dual norm $\gamma_2^*$ satisfies several similar properties to $\gamma_2$, such as $\gamma_2^*(A|Z) \leq \gamma_2^*(A|Y \cup Z)$, $\gamma_2^*(A|Z) = \gamma_2^*(A|Z \cup \{\sum_j p_j Z_j\})$ if $\sum_j |p_j| =$

1, and $\gamma_2^*(A \otimes B|Y \otimes Z) = \gamma_2^*(A|Y)\gamma_2^*(B|Z)$. It also satisfies $\gamma_2^*(A \circ B|Z) = \gamma_2^*(A|Z \circ B) \leq \gamma_2^*(A|Z)\gamma_2(B)$. We leave the proofs of these claims to the reader.

# B Application to continuous-time query complexity

The first step of the proof of Cleve et al. [CGM+09] is to show that the continuous-time model is equivalent to the fractional quantum query model, up to constant factors. For completeness, we now show that $\gamma_2(\sigma - \rho|\Delta)$ remains a lower bound on the fractional query complexity, up to a constant. Together with our upper bound, this gives Theorem 4.10. Yonge-Mallo [YM11] recently published a proof from 2007 which directly shows the general adversary bound is a lower bound on the continuous-time query model, and this was also independently observed by Landahl (unpublished).

Let us first describe the fractional query model. For simplicity, we restrict to the case of boolean input. Here the $\lambda$-fractional query operator $O_x(\lambda)$ behaves as $O_x(\lambda)|i\rangle|z\rangle = e^{i\lambda\pi x_i}|i\rangle|z\rangle$. Thus the usual query operator is obtained with $\lambda = 1$. The query cost is $\lambda$ times the number of applications of $O_x$.

As before the key step is to bound how much a single query can change the distance.

**Lemma B.1.** *Suppose that $\sigma$ can be reached from $\rho$ with one $\lambda$-fractional query. Then $\gamma_2(\rho - \sigma|\Delta) \leq \lambda\pi\sqrt{2}$.*

*Proof.* Let $\rho_x^i$ be the projection of $\rho_x$ onto the part of the query register holding $i$. Then

$$\begin{aligned}(\rho - \sigma)_{x,y} &= \sum_{j=1}^{n} \left(\langle\rho_x^j|\rho_y^j\rangle - \langle\rho_x^j|e^{-i\lambda\pi x_j}e^{i\lambda\pi y_j}|\rho_y^j\rangle\right) \\ &= \sum_{j:x_j\neq y_j} \langle\rho_x^j|\rho_y^j\rangle(1 - e^{i\lambda\pi(y_j - x_j)}) \\ &= \sum_{j:x_j\neq y_j} \langle\rho_x^j|\rho_y^j\rangle\left((1 - \cos(\lambda\pi)) + i(x_j - y_j)\sin(\lambda\pi)\right) \ .\end{aligned}$$

Now we define positive semi-definite matrices $\{P_j\}_{j\in[n]}$ satisfying $\rho - \sigma = \sum_j P_j \circ \Delta_j$. For this let us define a couple of auxiliary matrices. Let $M_j(x,y) = \langle\rho_x^j|\rho_y^j\rangle$ and let $E_j(x,y) = \langle e_{x_j}|e_{y_j}\rangle$, where $e_b = b + i(1-b)$ for $b \in \{0,1\}$. From this definition it is clear that $E_j$ is positive semi-definite, and note that

$$E_j(x,y) = \begin{cases} i(x_j - y_j) & \text{if } x_j \neq y_j \\ 1 & \text{otherwise} \ . \end{cases}$$

Finally, we can define $P_j = (1 - \cos(\lambda\pi))M_j + \sin(\lambda\pi)M_j \circ E_j$. Then $P_j$ is positive semi-definite, and satisfies $\rho - \sigma = \sum_j P_j \circ \Delta_j$. As $\sum_{j\in[n]} M_j(x,x) = 1$ for all $x$ we can upper bound the cost $\max_x \sum_{j\in[n]} P_j(x,x)$ by $p(\lambda) = (1 - \cos(\lambda\pi)) + \sin(\lambda\pi)$. Note that $p(0) = 0$ and the maximum value of the derivative of $p(\lambda)$ is $\pi\sqrt{2}$. Thus for $\lambda \geq 0$ we have $p(\lambda) \leq \lambda\pi\sqrt{2}$. □

# C Function composition

In this section we prove the composition lemmas, Lemmas 5.1 and 5.2.

We begin with some notation. Let $g : \mathcal{C} \to D$ where $\mathcal{C} \subseteq C^m$ and $f : D^n \to E$ for finite sets $C$, $D$ and $E$. Let $G = \{\delta_{g(x),g(y)}\}_{x,y}$, and $F = \{\delta_{f(x),f(y)}\}_{x,y}$. For a string $x \in \mathcal{C}^n$ we write $x = (x^1, \ldots, x^n)$ where each $x^i \in \mathcal{C}$, and we let $\tilde{x} = g(x^1) \cdots g(x^n) \in D^n$. For a $|D|^n$-by-$|D|^n$ matrix $A$, define a $|\mathcal{C}|^n$-by-$|\mathcal{C}|^n$ matrix $\tilde{A}$ by $\tilde{A}_{x,y} = A_{\tilde{x},\tilde{y}}$. With this notation, $\tilde{\Delta}_p = J^{\otimes(p-1)} \otimes (J - G) \otimes J^{\otimes(n-p)}$, and the filtering matrices for the composed function $f \circ g^n$ are $\Delta_{(p,q)} = J^{\otimes(p-1)} \otimes \Delta_q \otimes J^{\otimes(n-p)}$. To shorten expressions like these, we will use the notation $(B)_p \otimes \bigotimes_{i \neq p} A_i = A^{\otimes p-1} \otimes B \otimes A^{\otimes n-p}$.

*Proof of Lemma 5.1.* The lemma is a consequence of the composition property (11) from Lemma A.2, together with several other properties of the filtered $\gamma_2$ norm. Let $\bar{F} = J - F$, so $\mathrm{Adv}^{\pm}(f \circ g^n) = \gamma_2(\tilde{\bar{F}} | \{\Delta_{(p,q)} \circ \tilde{\bar{F}}\})$. We have

$$\begin{aligned}
\gamma_2(\tilde{\bar{F}} | \{\Delta_{(p,q)} \circ \tilde{\bar{F}}\}) &\leq \gamma_2(\tilde{\bar{F}} | \{\tilde{\Delta}_p \circ \tilde{\bar{F}}\}) \max_{\rho} \gamma_2(\tilde{\Delta}_\rho \circ \tilde{\bar{F}} | \{\Delta_{(p,q)} \circ \tilde{\bar{F}}\}) && \text{property (11)} \\
&\leq \gamma_2(\bar{F} | \{\Delta_p \circ \bar{F}\}) \max_{\rho} \gamma_2(\tilde{\Delta}_\rho | \{\Delta_{(\rho,q)} : q \in [m]\}) && (5, 10, 6) \\
&= \mathrm{Adv}^{\pm}(f) \gamma_2(J - G | \Delta) \ . && (1, 13)
\end{aligned}$$

The last step uses $\gamma_2((J - G)_\rho \otimes \bigotimes_{i \neq \rho} J_i | \{(\Delta_q)_\rho \otimes \bigotimes_{i \neq \rho} J_i\}_q) = \gamma_2(J - G | \Delta) \gamma_2(J | J)^{n-1}$. □

*Proof of Lemma 5.2.* For the lower bound, we will use the dual formulation of the adversary bound. Either by writing Eq. (3.3) as an SDP and taking the dual or by noting that $\mathrm{Adv}^{\pm}(g) = \max_W \{\langle J - G, W \rangle : \gamma_2^*(W | \Delta \circ (J - G)) \leq 1\}$ and using Eq. (A.3), we find

$$\begin{aligned}
\mathrm{Adv}^{\pm}(g) = \underset{\Omega, W}{\text{maximize}} \quad & \langle J, W \rangle \\
\text{subject to} \quad & \Omega \circ \mathbf{1} = \Omega \\
& \mathrm{Tr}(\Omega) = 1 \\
& W \circ G = 0 \\
& \Omega \pm W \circ \Delta_j \succeq 0 \ .
\end{aligned} \tag{C.1}$$

We first note some basic properties of an optimal dual solution $\Omega, W$.

**Claim C.1.** *Let $g : \mathcal{C} \to D$, where $\mathcal{C} \subseteq C^m$. Then there is an optimal solution $\Omega, W$ to Eq. (C.1) that satisfies $\mathrm{Adv}^{\pm}(g)\Omega \pm W \succeq 0$. If $D = \{0, 1\}$ we may also assume $\sum_{x:g(x)=1} \Omega_{x,x} = \sum_{x:g(x)=0} \Omega_{x,x} = \frac{1}{2}$.*

*Proof.* Let $\Omega, W$ be an optimal solution to Eq. (C.1) and let $d_g = \mathrm{Adv}^{\pm}(g) = \langle W, J \rangle$. Note that $d_g\Omega + W \succeq 0$ if and only if $d_g\Omega - W \succeq 0$ since $\Omega$ is diagonal and $W = W \circ (J - G)$ is bipartite. Suppose that $d_g\Omega - W \not\succeq 0$. Then there exists $\phi \succeq 0$, such that $\langle \phi, W \rangle > d_g \langle \phi, \Omega \rangle$. By normalizing $\phi$, we may assume that $\langle \phi, \Omega \rangle = 1$. This shows $\phi \circ \Omega, \phi \circ W$ is a feasible solution for $g$ with objective value greater than $d_g$, a contradiction.

Now for the second part. We may reorder the rows and columns of $\Omega, W$ so that all elements $x$ with $g(x) = 0$ come first, then all elements $y$ with $g(y) = 1$. Then the matrices $\Omega \pm W \circ \Delta_i$ have the form

$$\begin{bmatrix} \Omega_0 & 0 \\ 0 & \Omega_1 \end{bmatrix} \pm \begin{bmatrix} 0 & X \\ X^\dagger & 0 \end{bmatrix} \circ \Delta_i \ ,$$

where $W = \left[\begin{smallmatrix} 0 & X \\ X^\dagger & 0 \end{smallmatrix}\right]$. Thus for any $c > 0$,

$$\begin{bmatrix} c\Omega_0 & 0 \\ 0 & \frac{1}{c}\Omega_1 \end{bmatrix} \pm \begin{bmatrix} 0 & X \\ X^\dagger & 0 \end{bmatrix} \circ \Delta_i \succeq 0 \ .$$

If we did not originally have $\mathrm{Tr}(\Omega_0) = \mathrm{Tr}(\Omega_1)$ then choosing $c = \sqrt{\frac{\mathrm{Tr}(\Omega_1)}{\mathrm{Tr}(\Omega_0)}}$ to balance them will result in a solution with smaller trace, a contradiction to the optimality of $\Omega, W$. □

Notice that because of the second item we have that

$$\sum_{\substack{x,y \\ g(x)=a, g(y)=b}} \mathrm{Adv}^{\pm}(g)\Omega_{x,y} + W_{x,y} = \mathrm{Adv}^{\pm}(g)/2 \tag{C.2}$$

for any $a, b \in \{0,1\}$. This is the main property of boolean functions we use.

Now let $d_f = \mathrm{Adv}^{\pm}(f)$, $d_g = \mathrm{Adv}^{\pm}(g)$, and let $\Lambda, V$ and $\Omega, W$ be optimal solutions to Eq. (C.1) for $f$ and $g$, respectively, satisfying the conditions of Claim C.1 as appropriate. Our proposed solution to Eq. (C.1) for the composed function $f \circ g^n$ is the diagonal matrix $d_g^{n-1}\tilde{\Lambda} \circ \Omega^{\otimes n}$ and weight matrix $\tilde{V} \circ (d_g\Omega + W)^{\otimes n}$. Notice that the weight matrix satisfies the constraint $\tilde{F} \circ \left(\tilde{V} \circ (d_g\Omega + W)^{\otimes n}\right) = 0$ as $F \circ V = 0$.

Let us check the objective value.

$$\begin{aligned}
\left\langle J, (\tilde{V} \circ (d_g\Omega + W)^{\otimes n})\right\rangle &= \sum_{\substack{a,b\in\{0,1\}^n \\ f(a)\neq f(b)}} V_{a,b} \sum_{\substack{x,y \\ \tilde{x}=a, \tilde{y}=b}} \prod_i \left(d_g\Omega_{x^i,y^i} + W_{x^i,y^i}\right) \\
&= \sum_{\substack{a,b\in\{0,1\}^n \\ f(a)\neq f(b)}} V_{a,b} \prod_i \sum_{\substack{x^i,y^i \\ g(x^i)=a_i, g(y^i)=b_i}} \left(d_g\Omega_{x^i,y^i} + W_{x^i,y^i}\right) \\
&= d_f\left(\frac{d_g}{2}\right)^n .
\end{aligned}$$

The last line follows by Eq. (C.2).

It remains to show that $d_g^{n-1}\tilde{\Lambda} \circ \Omega^{\otimes n} \pm \tilde{V} \circ (d_g\Omega + W)^{\otimes n} \circ \Delta_{(p,q)} \succeq 0$ for all $(p,q)$. As $\mathrm{Tr}(d_g^{n-1}\tilde{\Lambda} \circ \Omega^{\otimes n}) = d_g^{n-1}/2^n$ by Claim C.1, this will complete the proof.

We know that $d_g\Omega + W \succeq 0$, $\Omega + W \circ \Delta_q \succeq 0$. Also $\tilde{\Lambda} \pm \tilde{V} \circ \tilde{\Delta}_p \succeq 0$ follows from $\Lambda \pm V \circ \Delta_p \succeq 0$ as they are equal up to repetition of some rows and columns. Thus

$$\begin{aligned}
0 &\preceq (\tilde{\Lambda} \pm \tilde{V} \circ \tilde{\Delta}_p) \circ \left((\Omega + W \circ \Delta_q)_p \otimes \bigotimes_{i\neq p} (d_g\Omega_i + W_i)\right) \\
&= d_g^{n-1}\tilde{\Lambda} \circ \Omega^{\otimes n} \pm \tilde{V} \circ \tilde{\Delta}_p \circ \left((\Omega + W \circ \Delta_q)_p \otimes \bigotimes_{i\neq p} (d_g\Omega_i + W_i)\right) .
\end{aligned}$$

This equality follows as $\tilde{\Lambda}_{x,y} = 0$ unless $\tilde{x} = \tilde{y}$, meaning that $g(x^i) = g(y^i)$ for all $i$. On the other hand, $W_{x^i,y^i} = 0$ if $g(x^i) = g(y^i)$, which kills all terms involving $\tilde{\Lambda}$ and $W$.

Now substitute $\tilde{\Delta}_p = (J - G)_p \otimes \bigotimes_{i\neq p} J_i$ and simplify $(J - G) \circ (\Omega + W \circ \Delta_q) = (d_g\Omega + W) \circ \Delta_q$ since $(J - G) \circ \Omega = \Delta_q \circ \Omega = G \circ W = 0$. Since $\Delta_{(p,q)} = (\Delta_q)_p \otimes \bigotimes_{i\neq p} J_i$, this gives $d_g^{n-1}\tilde{\Lambda} \circ \Omega^{\otimes n} \pm \tilde{V} \circ (d_g\Omega + W)^{\otimes n} \circ \Delta_{(p,q)} \succeq 0$, as desired. □