comput. complex. **22** (2013), 429 – 462

1016-3328/13/020429-34
published online April 30, 2013
DOI 10.1007/s00037-013–0066-8

computational complexity

A STRONG DIRECT PRODUCT THEOREM FOR QUANTUM QUERY COMPLEXITY

Troy Lee and Jérémie Roland

Abstract. We show that quantum query complexity satisfies a strong direct product theorem. This means that computing k copies of a function with fewer than k times the quantum queries needed to compute one copy of the function implies that the overall success probability will be exponentially small in k. For a boolean function f, we also show an XOR lemma—computing the parity of k copies of f with fewer than k times the queries needed for one copy implies that the advantage over random guessing will be exponentially small. We do this by showing that the multiplicative adversary method, which inherently satisfies a strong direct product theorem, characterizes bounded-error quantum query complexity. In particular, we show that the multiplicative adversary bound is always at least as large as the additive adversary bound, which is known to characterize bounded-error quantum query complexity.

Keywords. Quantum query complexity, adversary method, strong direct product theorem, XOR lemma.

Subject classification. 68Q12, 68Q17, 81P68.

1. Introduction

A fundamental question in complexity theory is how the difficulty of computing k independent instances of a function scales with the difficulty of computing the function. Intuitively, if r resources are needed to compute a function f with error probability 1/3, we expect that even with $\alpha k r$ resources, for $\alpha < 1$, we can only succeed in computing k independent instances of f with probability exponentially small in k. Proving such a result is known as a strong direct product theorem. While intuitive, for some models of computation such a statement is simply false (Shaltiel 2003), and there are still relatively few computational models where strong direct product theorems have been shown. Notable examples of direct product-type results include Yao's XOR lemma (Goldreich *et al.* 2011) and Raz's parallel repetition theorem (Raz 1998). Closer to our setting, strong direct product theorems have been shown for bounded-round randomized communication complexity (Jain *et al.* 2012) and for randomized query complexity (Drucker 2011).

In this work, we show that quantum query complexity satisfies a strong direct product theorem. For boolean functions, we further show an XOR lemma. XOR lemmas are closely related to strong direct product theorems and state that computing the parity of k copies of a boolean function with fewer than k times the resources needed to compute one copy implies that the advantage over random guessing will be exponentially small. XOR lemmas can be shown quite generally to imply strong direct product theorems and even threshold direct product theorems (Unger 2009), which state that one cannot compute a μ fraction of the k copies with fewer than (a constant fraction of) μk times the resources with better than exponentially small (in μk) success probability. Thus, in the boolean case, we are also able to obtain a threshold direct product theorem.

For classical randomized query complexity, in addition to a strong direct product theorem, Drucker (2011) also showed an XOR lemma and a threshold direct product theorem. Thus, for both randomized and quantum query complexity, all the major open problems relating to direct product theorems have now been answered. The techniques used by Drucker are quite different from the ones used here.

Previous results for quantum query complexity. A result related to, but weaker than, a strong direct product theorem is a direct sum theorem. These state that the resources needed to compute k-copies of a function are at least k times the resources needed to compute the function—with the same error parameter. A direct sum theorem is known for quantum query complexity—it follows from results of Ambainis *et al.* (2010a) that the adversary method obeys a direct sum theorem and the fact that the adversary method characterizes quantum query complexity (Lee *et al.* 2011; Reichardt 2011).

Strong direct product theorems in quantum query complexity were previously known only for some special classes of functions and bounds shown by particular methods. In the first such result, Klauck *et al.* (2007) used the polynomial method (Beals *et al.* 1998) to show a strong direct product theorem for the quantum query complexity of the OR function. Via block sensitivity, this gives a polynomially tight strong direct product theorem for all functions—namely any algorithm using fewer than a constant fraction times $kQ_{1/3}(f)^{1/6}$ will have exponentially small success probability for computing k copies of f. Here and in the rest of the paper, we use $Q_\epsilon(f)$ to denote the ϵ-bounded-error quantum query complexity of f.

Sherstov (2011) recently showed how certain lower bound techniques based on looking at the distance of the function to a convex set inherently satisfy a strong direct product theorem. As an application, he was able to show that the polynomial method satisfies a strong direct product theorem *in general*. Thus, one obtains a strong direct product theorem for the quantum query complexity of any function where the polynomial method shows a tight lower bound. Super-linear gaps between the polynomial degree and quantum query complexity are known (Ambainis 2006), however, so this does not give a tight strong direct product theorem for all functions.

Direct product results have also been shown by the other main lower bound technique in quantum query complexity, the adversary method. The adversary method defines a potential function based on the state of the algorithm after t queries and bounds the change in this potential function from one query to the next. By developing a new kind of adversary method, Ambainis *et al.* (2006) showed a strong direct product theorem for all symmetric functions. Špalek (2008) formalized this technique into a generic method, coining it the multiplicative adversary method and showed that this method inherently satisfies a strong direct product theorem. The name multiplicative adversary contrasts with the additive adversary method, introduced earlier by Ambainis (2002) and later extended by Høyer *et al.* (2007). The additive adversary method bounds the difference of the potential function from one step to the next, while the multiplicative adversary method bounds the corresponding ratio.

Our results. There have recently been great strides in our understanding of the adversary methods. A series of works (Ambainis *et al.* 2010b; Childs *et al.* 2009; Farhi *et al.* 2008; Lee *et al.* 2011; Reichardt 2009, 2011; Reichardt & Špalek 2008) has culminated in showing that the additive adversary method characterizes the bounded-error quantum query complexity of any function whatsoever. Ambainis *et al.* (2011), answering an open question of Špalek (2008), showed that the multiplicative adversary is at least as large as the additive. Thus, the multiplicative adversary bound also characterizes bounded-error quantum query complexity.

This seems like it would close the question of a strong direct product theorem for quantum query complexity. The catch is the following. The multiplicative adversary method can be viewed as a family of methods parameterized by the bound c on the ratio of the potential function from one step to the next. The strong direct product theorem of Špalek (2008) holds for any value of c sufficiently bounded away from 1. The result of Ambainis *et al.* (2011), however, was shown in the limit $c \to 1$, which ends up degrading the resulting direct product theorem into a direct sum theorem. We show that the multiplicative adversary is at least as large as the additive adversary for a value of c bounded away from 1 (Claim 3.18). A similar result was independently proved by Belovs (2011). Together with the strong direct product theorem for the multiplicative adversary by Špalek (2008), this suffices to give a strong direct product theorem for quantum query complexity. Rather than use this "out of the box" strong direct product theorem, however, we prove the strong direct product theorem from scratch using a stronger output condition than those used previously (Ambainis *et al.* 2011; Špalek 2008). This results in better parameters and a better understanding of the multiplicative adversary method.

THEOREM 1.1 (Strong direct product theorem). *Let $f : \mathcal{D} \to E$ where $\mathcal{D} \subseteq D^n$ for finite sets D, E. For an integer $k > 0$ define $f^{(k)}(x^1, \ldots, x^k) = (f(x^1), \ldots, f(x^k))$. Then, for any $\delta \in [2/3, 1]$,*

$$Q_{1-\delta^{k/2}}\left(f^{(k)}\right) \geq \frac{k \ln(3\delta/2)}{8000} \cdot Q_{1/4}(f).$$

In the boolean case, we prove the following XOR lemma which also implies a threshold direct product theorem (Theorem 5.4).

LEMMA 1.2 (XOR Lemma). *Let $f : \mathcal{D} \to \{0,1\}$ where $\mathcal{D} \subseteq D^n$ for finite set D. For an integer $k > 0$ define $f^{\oplus k}(x_1, \ldots, x_k) = \sum_i f(x_i) \bmod 2$. For any $\delta \in [0,1]$,*

$$Q_{(1-\delta^{k/2})/2}\left(f^{\oplus k}\right) \geq \frac{k\delta}{8000} \cdot Q_{1/4}(f).$$

Proof technique. While the statement of our main theorems concern functions, a key to our proofs, especially for the XOR lemma, is to consider more general state generation problems, introduced by Ambainis *et al.* (2011). Instead of producing a classical value $f(x)$ on input x, the goal in state generation is to produce a specified target state $|\sigma_x\rangle$, again by making queries to the input x. We will refer to $\sigma(x, y) = \langle\sigma_x|\sigma_y\rangle$ as the target gram matrix. Evaluating a function f can be viewed as a special case of state generation where the target gram matrix is $F(x, y) = \delta_{f(x),f(y)}$, where $\delta_{a,b}$ denotes the Kronecker delta function.

Our most general result (Theorem 4.1) shows that for a restricted class of target gram matrices σ, to generate $\sigma^{\otimes k}$ with better than exponentially small success probability requires at least a constant fraction of k times the complexity of σ. The strong direct product theorem is obtained as a special case of this theorem by considering the gram matrix $F(x, y) = \delta_{f(x),f(y)}$. To obtain the XOR lemma, we apply this theorem with the state generation problem of computing f in the phase, that is to generate

$\sigma_f(x,y) = (-1)^{f(x)+f(y)}$. The advantage of considering this state is that $\sigma_f^{\otimes k}$ is the state generation problem corresponding to computing the parity of k copies of f in the phase. We then show that the complexities of f and the state generation problem of computing f in the phase are closely related.

Another key element of our proofs is a new characterization of the set of valid output gram matrices for an algorithm solving a state generation problem with success probability $1-\epsilon$ (Claim 2.9). A necessary condition for a matrix to be a valid output matrix is called an output condition. Usually, a lower bound uses an output condition that is necessary but not sufficient, and that is therefore a relaxation of the *true* output condition. In this case, the lower bound is shown against all gram matrices satisfying this relaxed output condition and thereby all valid output matrices as well. Examples of output conditions previously used with the adversary bound include being close to the target gram matrix in distance measured by the l_∞ or γ_2 (see Definition 2.6) norms. These output conditions, however, do not work for small success probabilities, which is critical to obtain the strong direct product theorem.

We give a new characterization of the true output condition in terms of fidelity. We then relax this condition by replacing the fidelity between quantum states by the fidelity between probability distributions arising from a measurement on those states. The key observation is that a witness for the adversary bound of the problem is a hermitian matrix, which can be interpreted as a physical observable that can be measured. Since the fidelity between two quantum states is upper bounded by the fidelity between the probability distributions arising from any measurement on those states, a relaxation of this output condition may be obtained by considering the measurement corresponding to an optimal witness for the adversary bound of the problem. A lower bound on the multiplicative bound under this relaxed output condition can be written as a linear program. By taking the dual of this linear program, we are able to lower bound the value on $\sigma^{\otimes k}$ in terms of the bound for σ by using a completely classical claim about product probability distributions (Corollary 3.13). This approach allows us to obtain a cleaner statement for the strong direct product theorem than what we would obtain from the output condition used by Špalek (2008) and Ambainis *et al.* (2011) and also clarifies the inner workings of the adversary method, which might be of independent interest.

2. Preliminaries

Let $\Re(z)$ denote the real part of a complex number z. Let $\delta_{a,b}$ denote the Kronecker delta function. We will refer throughout to a function $f : \mathcal{D} \to E$ where $\mathcal{D} \subseteq D^n$ for finite sets D, E. We let $f^{(k)} : \mathcal{D}^k \to E^k$ be the function computing k independent copies of f, namely $f^{(k)}(x^1, \ldots, x^k) = (f(x^1), \ldots, f(x^k))$. We let $f^{\oplus k}$ denote the parity function composed with $f^{(k)}$.

We also define some auxiliary matrices associated with f. Let $F(x, y) = \delta_{f(x),f(y)}$, and $\Delta_i(x, y) = \delta_{x_i,y_i}$ for $x, y \in \mathcal{D}$ and $i \in [n]$. For boolean functions, i.e., when $|E| = 2$, we also define the matrix $\sigma_f(x, y) = (-1)^{f(x)+f(y)}$ for $x, y \in \mathcal{D}$. Note that $\sigma_f = 2F - J$, where J is the all-1 matrix. We use $A \circ B$ for the entrywise product between two matrices A, B, also known as the Schur or Hadamard product.

For a probability distribution p, we use $\mathrm{E}_{A \leftarrow p}[g(A)]$ for the expected value of $g(A)$ when A is chosen according to p.

We use the notation $\rho \succeq 0$ to indicate that ρ is positive semidefinite, that is, hermitian with non-negative eigenvalues, and $\rho \succ 0$ if all eigenvalues are strictly positive. We use $\|v\|$ for the ℓ_2 norm of a vector v.

2.1. Distance measures. We will use several notions to measure the closeness of two quantum states or probability distributions. First we introduce the spectral norm and its dual the trace norm.

Definition 2.1. *Let M be a matrix.*

$$\|M\| = \max_{|u\rangle : \||u\rangle\| = 1} \mathrm{Tr}(M|u\rangle\langle u|).$$
$$\|M\|_{\mathrm{tr}} = \max_{P : \|P\| = 1} \mathrm{Tr}(MP).$$

We will also make extensive use of fidelity. A quantum state on a $|\mathcal{D}|$-dim Hilbert space is characterized by a $|\mathcal{D}| \times |\mathcal{D}|$ density matrix, that is, a positive semidefinite matrix ρ such that $\mathrm{Tr}(\rho) = 1$. Let ρ, σ be two such density matrices.

DEFINITION 2.2 (Fidelity). $\mathcal{F}(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_{\mathrm{tr}}$.

Note that $\mathcal{F}(\rho, \sigma) \in [0, 1]$ and the larger the fidelity the less distinguishable ρ and σ are, obtaining the extreme values $\mathcal{F}(\rho, \sigma) = 1$ if and only if $\rho = \sigma$ and $\mathcal{F}(\rho, \sigma) = 0$ if and only if $\rho\sigma = 0$. For pure states $\mathcal{F}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|$.

For classical probability distributions p, q we will abuse notation and simply write $\mathcal{F}(p, q)$ for $\mathcal{F}(\mathrm{diag}(p), \mathrm{diag}(q))$, where $\mathrm{diag}(p)$ is a diagonal matrix with the entries of p along the diagonal. Note that $\mathcal{F}(\mathrm{diag}(p), \mathrm{diag}(q)) = \sum_i p_i q_i$.

One property of the fidelity we will use is that it is jointly concave in its inputs (see Theorem 9.7 in Nielsen & Chuang 2000).

LEMMA 2.3. *Let p be a probability distribution on $[n]$ and ρ_i, σ_i be density matrices for $i \in [n]$. Then*

$$\mathcal{F}\big(\sum_i p_i \rho_i, \sum_i p_i \sigma_i\big) \geq \sum_i p_i \mathcal{F}(\rho_i, \sigma_i).$$

A positive operator valued measurement (POVM) is a set of positive semidefinite operators $\{E_i\}$ such that $\sum_i E_i = I$. We will make use of the following property of fidelity (see section 9.2.2 of (Nielsen & Chuang 2000)).

LEMMA 2.4. *Let ρ, σ be density matrices and $\{E_i\}$ a POVM. Then $\mathcal{F}(\rho, \sigma) \leq \mathcal{F}(p, q)$, where p, q are the probability distributions obtained from measuring $\{E_i\}$ on ρ, σ, i.e., $p(i) = \mathrm{Tr}(\rho E_i), q(i) = \mathrm{Tr}(\sigma E_i)$.*

The trace distance is related to the fidelity by the following lemma (see, for example, Eq. (9.110) in Nielsen & Chuang 2000).

LEMMA 2.5. *Let ρ, σ be density matrices. Then*

$$1 - \mathcal{F}(\rho, \sigma) \leq \frac{1}{2}\|\rho - \sigma\|_{\mathrm{tr}} \leq \sqrt{1 - \mathcal{F}(\rho, \sigma)^2}.$$

Finally, for a $|\mathcal{D}| \times |\mathcal{D}|$ matrix A, we will use the factorization norm $\gamma_2(A)$. This is also known as the Hadamard product operator norm (Bhatia 2007) and has recently found many applications in communication and query complexity (Lee *et al.* 2011, 2008; Linial *et al.* 2007; Linial & Shraibman 2009).

Definition 2.6 (Factorization norm).

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

By writing γ_2 as an optimization problem and taking the dual, one can obtain the following equivalent formulation as a maximization problem (see, for example, Lee *et al.* 2008).

Lemma 2.7.

$$\gamma_2(A) = \max_{\substack{|u\rangle, |v\rangle \\ \| |u\rangle \| = \| |v\rangle \| = 1}} \| A \circ |u\rangle\langle v| \|_{\mathrm{tr}}.$$

This maximization formulation makes a couple of facts apparent: first, that $\gamma_2(A) \geq \|A\|_{\mathrm{tr}}/|\mathcal{D}|$, and second that γ_2 obeys the triangle inequality $\gamma_2(A+B) \leq \gamma_2(A) + \gamma_2(B)$.

Finally, we will use the notion of relative entropy for a two-outcome event (Cover & Thomas 2006).

Definition 2.8 (Relative entropy). *For $0 \leq \lambda \leq 1$ and $0 < \mu < 1$, we denote by $D(\lambda || \mu)$ the relative entropy defined as*

$$D(\lambda || \mu) = \lambda \ln \frac{\lambda}{\mu} + (1 - \lambda) \ln \frac{1 - \lambda}{1 - \mu},$$

where $0 \ln 0 = 0$.

2.2. Quantum query complexity and state generation. The quantum query complexity of f, denoted $Q_\epsilon(f)$, is the minimum number of input queries needed to compute f with error probability at most ϵ. We refer to the survey by Buhrman & de Wolf (2002) for definitions and background on this model.

Although our main interest will be in the query complexity of functions, it will be useful to also talk about state generation problems, introduced by Ambainis *et al.* (2011). Instead of producing a classical value $f(x)$ on input x, the goal in state generation is to produce a specified target state $|\sigma_x\rangle$, again by making queries to the input x. As unitary transformations independent of the input can be made for free in the query model, a state generation problem is wholly determined by the gram matrix $\sigma(x, y) = \langle\sigma_x|\sigma_y\rangle$ of the target states $\{|\sigma_x\rangle\}_{x\in\mathcal{D}}$. We refer to σ as the target gram matrix.

State generation problems come in two variations, coherent and non-coherent. These variations differ in the *output condition*, the requirements placed on the final state of a successful algorithm.

Coherent state generation. An algorithm $\mathcal{P}$ solves the coherent quantum state generation problem σ with error at most ϵ if, for every $x \in \mathcal{D}$, it generates a state $|\mathcal{P}(x)\rangle \in \mathcal{H} \otimes \mathcal{H}'$ such that $\Re(\langle\mathcal{P}(x)|(|\sigma_x\rangle \otimes |\bar{0}\rangle)) \geq \sqrt{1-\epsilon}$, where $\mathcal{H}'$ denotes the workspace of the algorithm, and $|\bar{0}\rangle$ is a default state for $\mathcal{H}'$. The coherent quantum query complexity of σ, denoted $Q^c_\epsilon(\sigma)$, is the minimum number of queries needed to generate σ coherently with error at most ϵ.

We can equivalently rephrase the coherent output condition as $\Re(\langle\mathcal{P}(x)|V(|\sigma_x\rangle \otimes |\bar{0}\rangle)) \geq \sqrt{1-\epsilon}$ for some unitary V. This can be done as the unitary V can be appended to the algorithm at no extra cost. This formulation has the advantage that it only depends on the gram matrix σ of the vectors $\{|\sigma_x\rangle\}$ and the gram matrix $\sigma'(x, y) = \langle\mathcal{P}(x)|\mathcal{P}(y)\rangle$, rather than the vectors themselves.

In our strong direct product theorem, we will work directly with the coherent output condition, and this will be made much easier by the following claim that gives an equivalent reformulation in terms of fidelity.

Claim 2.9. *Let $\{|a_x\rangle\}, \{|b_x\rangle\}$ be two sets of unit vectors, and ρ, σ their corresponding gram matrices.*

$$\max_V \min_x \Re(\langle a_x|V|b_x\rangle) = \min_{|u\rangle:\||u\rangle\|=1} \mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma \circ |u\rangle\langle u|), \tag{2.10}$$

where the maximization is taken over all unitaries V.

Proof. By writing the left hand side as a semidefinite program and taking the dual one can show that

$$\max_V \min_x \Re(\langle a_x|V|b_x\rangle) = \min_{|u\rangle:\||u\rangle\|=1} \max_V \Re(\mathrm{Tr}(V \sum_x |\langle x|u\rangle|^2 |a_x\rangle\langle b_x|)).$$

Letting $D(u)$ be a diagonal matrix with entries $\langle x|D(u)|x\rangle = \langle x|u\rangle$, we can rewrite the right hand side of this last expression as

$$\max_V \min_x \Re(\langle a_x|V|b_x\rangle) = \min_{|u\rangle:\||u\rangle\|=1} \|AD(u)(BD(u))^\dagger\|_{\mathrm{tr}},$$

where $A = \sum_x |a_x\rangle\langle x|$ and $B = \sum_x |b_x\rangle\langle x|$. Since $\rho = A^\dagger A$, $\sigma = B^\dagger B$ and $\rho \circ |u\rangle\langle u| = D(u)^\dagger \rho D(u)$, the claim follows using

$$\|XY^\dagger\|_{\mathrm{tr}} = \|(X^\dagger X)^{1/2}(Y^\dagger Y)^{1/2}\|_{\mathrm{tr}}$$

and the definition of the fidelity

$$\mathcal{F}(X^\dagger X, Y^\dagger Y) = \|(X^\dagger X)^{1/2}(Y^\dagger Y)^{1/2}\|_{\mathrm{tr}}. \qquad \square$$

Non-coherent state generation. An algorithm $\mathcal{P}$ solves the non-coherent state generation problem σ with error at most ϵ if there exists a set of states $|\phi_x\rangle \in \mathcal{H}'$ such that $\Re(\langle \mathcal{P}(x)|(|\sigma_x\rangle \otimes |\phi_x\rangle)) \geq \sqrt{1-\epsilon}$ for all $x \in \mathcal{D}$. We denote by $Q_\epsilon(\sigma)$ the non-coherent query complexity of generating σ with error ϵ.

Evaluating a function f can be seen as a special case of non-coherent state generation where the target gram matrix is $F(x,y) = \delta_{f(x),f(y)}$. In other words, $Q_\epsilon(f) = Q_\epsilon(F)$, justifying our abuse of notation.

Coherent and non-coherent complexities for functions. Clearly $Q_\epsilon(\sigma) \leq Q^c_\epsilon(\sigma)$. In general, it is easier to lower bound the coherent quantum query complexity, but more interesting to lower bound the non-coherent complexity. Luckily, for state generation problems corresponding to functions, the coherent and non-coherent complexities are closely related as shown in the next two claims.

Claim 2.11. *Let f be a function. Then*

$$Q_\epsilon(F) \leq Q^c_\epsilon(F) \leq 2Q_{1-\sqrt{1-\epsilon}}(F).$$

Proof. The lower bound holds for a general target gram matrix σ, as the success condition in the coherent case implies the non-coherent one.

For the upper bound, let A_x be an algorithm computing $f(x)$ with success probability $1-\eta$. Let $p = |E|$ be the size of the output set, which we assume to be $E = \{0, \ldots, p-1\}$ for simplicity. In what follows, $+$ will denote addition modulo p when applied on elements of E. Thus, the algorithm applied on $|0\rangle|\bar{0}\rangle$, where the first register is the output register and the second register corresponds to some workspace initialized in a default state, prepares a state

$$A_x|0\rangle|\bar{0}\rangle = \sum_{j \in E} \alpha_j |j + f(x)\rangle|\psi_j\rangle,$$

where by assumption $|\alpha_0| \geq \sqrt{1-\eta}$, and the states $|\psi_j\rangle$ describe the final state of the workspace register. Let us now copy the output register into an additional register initialized in the state $|0\rangle$ using an addition gate G, and finally uncompute the original output register together with the workspace by using the algorithm A_x in reverse.

We analyze the overlap of $A_x^{-1} G A_x |0\rangle|\bar{0}\rangle|0\rangle$ with $|0\rangle|\bar{0}\rangle|f(x)\rangle$. After applying G on $A_x|0\rangle|\bar{0}\rangle|0\rangle$, we have the state

$$|v\rangle = \sum_{j \in E} \alpha_j |j + f(x)\rangle|\psi_j\rangle|j + f(x)\rangle.$$

Now we look at the overlap of $|0\rangle|\bar{0}\rangle|f(x)\rangle$ with $A_x^{-1}|v\rangle$ or, equivalently, the overlap of $A_x|0\rangle|\bar{0}\rangle|f(x)\rangle$ with $|v\rangle$. Since

$$A_x|0\rangle|\bar{0}\rangle|f(x)\rangle = \sum_{j \in E} \alpha_j |j + f(x)\rangle|\psi_j\rangle|f(x)\rangle,$$

we have

$$\langle 0|\langle \bar{0}|\langle f(x)|A_x^{-1}|v\rangle = \sum_{j \in E} |\alpha_j|^2 \langle f(x)|j + f(x)\rangle \geq 1 - \eta.$$

Therefore, this algorithm coherently computes $f(x)$ with success probability $1-\epsilon \geq (1-\eta)^2$. Inverting this relation, we obtain $\eta \geq 1-\sqrt{1-\epsilon}$. $\square$

We will also consider another type of state generation problem associated with a function, the problem of computing the function in the phase. For a boolean function $f : \mathcal{D} \to \{0,1\}$ let $\sigma_f(x,y) = (-1)^{f(x)+f(y)}$. While the non-coherent complexity of σ_f is trivial, the coherent complexity of σ_f is closely related to that of F.

Claim 2.12.

$$Q^c_{(1-\sqrt{1-\epsilon})/2+\epsilon/4}(F) \leq Q^c_\epsilon(\sigma_f) \leq 2Q_{(1-\sqrt{1-\epsilon})/2}(F).$$

Proof. For the lower bound, we turn an algorithm for σ_f into an algorithm for $F = (J+\sigma_f)/2$ by the following standard technique: We introduce an ancilla qubit prepared in the state $(|0\rangle + |1\rangle)/\sqrt{2}$, apply the original algorithm conditionally on this ancilla being in state $|1\rangle$ and then apply the Hadamard operator H on the ancilla qubit.

Say that the output gram matrix of the original algorithm is ρ. By Claim 2.9, and using joint concavity of fidelity (Lemma 2.3), we can upper bound the error of the new algorithm as

$$\begin{aligned}
&\min_{\substack{|u\rangle\\ \||u\rangle\|=1}} \mathcal{F}\left(\tfrac{J+\rho}{2} \circ |u\rangle\langle u|, \tfrac{J+\sigma_f}{2} \circ |u\rangle\langle u|\right) \\
&\quad\geq \frac{1}{2}\Big(\min_{\substack{|u\rangle\\ \||u\rangle\|=1}} \mathcal{F}(|u\rangle\langle u|, |u\rangle\langle u|) + \min_{\substack{|u\rangle\\ \||u\rangle\|=1}} \mathcal{F}\left(\rho \circ |u\rangle\langle u|, \sigma_f \circ |u\rangle\langle u|\right)\Big) \\
&\quad\geq \frac{1}{2}\Big(1 + \min_{|u\rangle : \||u\rangle\|=1} \mathcal{F}\left(\rho \circ |u\rangle\langle u|, \sigma_f \circ |u\rangle\langle u|\right)\Big) \\
&\quad\geq \frac{1}{2} + \frac{\sqrt{1-\epsilon}}{2}.
\end{aligned}$$

For the upper bound, let us consider an algorithm A_x computing $f(x)$ (in a register) with success probability $1-\eta$. Thus, the algorithm applied on $|0\rangle|\bar{0}\rangle$, where the first register is the output register and the second register corresponds to some workspace initialized in a default state, prepares a state

$$A_x|0\rangle|\bar{0}\rangle = \sum_{j\in\{0,1\}} \alpha_j |j + f(x)\rangle|\psi_j\rangle,$$

where by assumption $|\alpha_0| \geq \sqrt{1-\eta}$, and the states $|\psi_j\rangle$ describe the final state of the workspace register. Let Φ be a phase gate acting on the output register as $|b\rangle \mapsto (-1)^{f(x)}|b\rangle$. We can turn an algorithm A_x computing in a register into an algorithm computing in the phase by first applying A_x to compute the output, then applying the phase gate Φ, and finally applying A_x^{-1} to uncompute the output.

After applying Φ on $A_x|0\rangle|\bar{0}\rangle$, we have the state $\Phi A_x|0\rangle|\bar{0}\rangle = \sum_{j\in\{0,1\}}(-1)^{j+f(x)}\alpha_j|j + f(x)\rangle|\psi_j\rangle$. Now we look at the overlap of $(-1)^{f(x)}|0\rangle|\bar{0}\rangle$ with $A_x^{-1}\Phi A_x|0\rangle|\bar{0}\rangle$ or, equivalently, the overlap of $(-1)^{f(x)}A_x|0\rangle|\bar{0}\rangle$ with $\Phi A_x|0\rangle|\bar{0}\rangle$. We have

$$(-1)^{f(x)}\langle 0|\langle\bar{0}|A_x^{-1}\Phi A_x|0\rangle|\bar{0}\rangle = \sum_{j\in\{0,1\}} (-1)^j |\alpha_j|^2 \geq 1 - 2\eta.$$

Therefore, we obtain a success probability $1-\epsilon \geq (1-2\eta)^2$. Inverting this relation, we obtain $\eta \geq (1-\sqrt{1-\epsilon})/2$. □

3. Adversary methods

In this section, we introduce both the additive and multiplicative adversary lower bound methods. Even when one is only interested in the functional case, it is useful to view these methods as lower bounds on quantum state generation as this allows the separation of the method into two distinct parts. The first part is a lower bound on exact coherent quantum state generation. This is where the two methods differ. The second part is the output condition, a minimization of the bound for exact coherent quantum state generation over all valid output gram matrices. The set of valid output gram matrices is determined by the target gram matrix σ, the error parameter ϵ, and whether one is considering coherent or non-coherent state generation. This second step is common to both the additive and multiplicative methods. Finally, we show that the multiplicative bound is at least as large as the additive bound.

3.1. Additive method. We first review the derivation of the additive adversary method to compare it with the multiplicative method in the next section. We will actually present a generalization of the additive adversary method due to Lee *et al.* (2011).

Consider an algorithm that exactly and coherently generates the target state σ_x by making T queries to the input x, for all $x \in \mathcal{D}$. Let $|\psi_x^t\rangle$ be the state of this algorithm on input x after t queries, and $\rho^t(x,y) = \langle\psi_x^t|\psi_y^t\rangle$ be the corresponding gram matrix. Note that $\rho^0 = J$, the all ones matrix, and, by assumption, $\rho^T = \sigma$.

Now let Γ be a matrix, $|v\rangle$ a unit vector, and consider the potential function $\Phi(t) = \mathrm{Tr}((\Gamma \circ \rho^t)|v\rangle\langle v|)$. The additive change in this potential function from the beginning to the end of the protocol is

$$\begin{aligned}\mathrm{Tr}((\Gamma \circ (J-\sigma))|v\rangle\langle v|) &= \sum_{t=0}^{T-1} \mathrm{Tr}((\Gamma \circ (\rho^t - \rho^{t+1}))|v\rangle\langle v|)\\ &\le T \max_t \mathrm{Tr}((\Gamma \circ (\rho^t - \rho^{t+1}))|v\rangle\langle v|).\end{aligned}$$

A standard argument (see, for example, Høyer *et al.* 2007) then goes that if we impose the condition on Γ that

$$I \pm \Gamma \circ (J - \Delta_i) \succeq 0 \quad \text{for all } i \in [n],$$

then $\mathrm{Tr}((\Gamma \circ (\rho^t - \rho^{t+1}))|v\rangle\langle v|) \le 2$, for all t and unit vectors $|v\rangle$.

As this argument holds for any Γ and v, we can maximize over them. The maximization over v gives rise to the spectral norm as $\|M\| = \max_{|v\rangle:\||v\rangle\|=1} \mathrm{Tr}(M|v\rangle\langle v|)$. This leads to the following definition (Lee *et al.* 2011).

Definition 3.1 (Additive adversary method).

$$\begin{aligned}\mathrm{Adv}^*(\sigma) = \underset{\Gamma}{\text{maximize}} \quad & \|\Gamma \circ (J-\sigma)\|\\ \text{subject to} \quad & I \pm \Gamma \circ (J - \Delta_i) \succeq 0 \text{ for all } i \in [n],\end{aligned}$$

where the maximization is over $|\mathcal{D}| \times |\mathcal{D}|$ *hermitian matrices* Γ.

The preceding argument shows the following.

THEOREM 3.2 (Lee *et al.* 2011). *For any target gram matrix* σ,

$$Q_0^c(\sigma) \geq \frac{\mathrm{Adv}^*(\sigma)}{2}.$$

Lee *et al.* (2011) have also shown that this lower bound is tight for the bounded-error query complexity of functions.

THEOREM 3.3 (Lee *et al.* 2011). *For any function* f,

$$Q_{1/4}(f) \leq 1000 \cdot \mathrm{Adv}^*(F).$$

Up to the constant factor, this upper bound holds more generally for *well-behaved* state generation problems. A state generation problem is well-behaved if the query complexity $Q_\epsilon(\sigma)$ does not depend dramatically on the error ϵ, that is if $Q_{1/4}(\sigma) = \Theta(Q_\epsilon(\sigma))$ for any small constant ϵ. This property holds for the query complexity of any function, but does not hold in general for state generation problems.

REMARK 3.4. *The adversary bound* $\mathrm{Adv}^{\pm}$ *from Høyer et al. (2007) was originally defined in the functional case, that is, for target gram matrices* F *of the form* $F(x, y) = \delta_{f(x),f(y)}$ *for a function* f. *This definition had an additional constraint that* $\Gamma \circ F = 0$. *This constraint only affects the bound up to a multiplicative factor of two (Lee et al. 2011).*

$$\mathrm{Adv}^{\pm}(F) \leq \mathrm{Adv}^*(F) \leq 2\mathrm{Adv}^{\pm}(F). \tag{3.5}$$

The constraint $\Gamma \circ F = 0$ *allows one to show that* $\mathrm{Adv}^{\pm}(F)/2$ *is a lower bound even on the non-coherent complexity of generating* F. *One can see that* $\mathrm{Adv}^*(F)/4$ *is a lower bound on the non-coherent complexity of generating* F *either by Eq. (3.5) or by Claim 2.11 showing that the coherent and non-coherent state generation complexities of functions are related by a factor of two.*

3.2. Multiplicative adversary method. The multiplicative bound is derived by considering the same potential function $\Phi(t)$, but looks at the ratio of this function at the beginning and end of the protocol, rather than the difference. Equivalently, one can consider the logarithmic potential function $\ln(\Phi(t))$ and again look at the additive change over the course of the protocol. To ensure that the argument to the logarithm is positive, we now restrict the maximization to positive definite matrices $\Gamma \succ 0$.

DEFINITION 3.6 (Multiplicative adversary method). $\mathrm{Madv}(\sigma) = \sup_{c>1} \mathrm{Madv}^{(c)}(\sigma)$, *where*

$$\begin{aligned} \mathrm{Madv}^{(c)}(\sigma) = \ \frac{1}{\ln(c)} \ & \underset{\Gamma \succ 0, |v\rangle}{\text{maximize}} \ln\left(\mathrm{Tr}((\Gamma \circ \sigma)|v\rangle\langle v|)\right) \\ & \text{subject to } \mathrm{Tr}(\Gamma|v\rangle\langle v|) = 1 \quad \text{and} \\ & c^{-1}\Gamma \preceq \Gamma \circ \Delta_i \preceq c\, \Gamma \quad \text{for all } i \in [n], \end{aligned}$$

and the maximization is over $|\mathcal{D}| \times |\mathcal{D}|$ positive definite matrices Γ and unit vectors $|v\rangle$. We will refer to a matrix $\Gamma \succ 0$ satisfying $c^{-1}\Gamma \preceq \Gamma \circ \Delta_i \preceq c\,\Gamma$ for all i as a multiplicative witness.

THEOREM 3.7 (Ambainis *et al.* 2011; Špalek 2008). *For any state generation problem σ,*

$$Q_0^c(\sigma) \geq \frac{\mathrm{Madv}(\sigma)}{2}.$$

PROOF. Consider an algorithm that coherently generates σ by making T queries. Let us denote by ρ^t the gram matrix of the states after the t-th query, that is, $\rho^t(x, y) = \langle \psi_x^t | \psi_y^t \rangle$, where $|\psi_x^t\rangle$ is the state of the algorithm on input x after t queries. We define a potential function $\Phi(t) = \mathrm{Tr}((\Gamma \circ \rho^t)|v\rangle\langle v|)$, where $\Gamma \succ 0$. Then

$$\begin{aligned} \frac{\Phi(T)}{\Phi(0)} = \frac{\mathrm{Tr}((\Gamma \circ \sigma)|v\rangle\langle v|)}{\mathrm{Tr}((\Gamma \circ J)|v\rangle\langle v|)} &= \prod_{t=0}^{T-1} \frac{\mathrm{Tr}((\Gamma \circ \rho^{t+1})|v\rangle\langle v|)}{\mathrm{Tr}((\Gamma \circ \rho^t)|v\rangle\langle v|)} \\ &\leq \left(\max_t \frac{\mathrm{Tr}((\Gamma \circ \rho^{t+1})|v\rangle\langle v|)}{\mathrm{Tr}((\Gamma \circ \rho^t)|v\rangle\langle v|)} \right)^T. \end{aligned}$$

Analogously to the additive bound, we now show that the constraint $c^{-1}\Gamma \preceq \Gamma \circ \Delta_i \preceq c\,\Gamma$ for all $i \in [n]$ implies

$$\max_t \frac{\mathrm{Tr}((\Gamma \circ \rho^{t+1})|v\rangle\langle v|)}{\mathrm{Tr}((\Gamma \circ \rho^{t})|v\rangle\langle v|)} \leq c.$$

This argument is very similar to proofs in Ambainis *et al.* (2011) and Špalek (2008), so we only sketch the idea here. Recall from Ambainis *et al.* (2011) that we can assume that there are only two types of queries, called computing and uncomputing queries (this restriction can only increase the query complexity by a factor at most 2, hence the factor 1/2 in the final lower bound). Let us first consider a computing query, in which case the state right before the query can be written as $|\psi_x^{t-1}\rangle = \sum_i |\psi_{x,i}^{t-1}\rangle|i\rangle|0\rangle$, while after the query it reads $|\psi_x^t\rangle = \sum_i |\psi_{x,i}^{t-1}\rangle|i\rangle|x_i\rangle$ (the situation for uncomputing queries is similar except that the roles of $|\psi_x^{t-1}\rangle$ and $|\psi_x^t\rangle$ are interchanged). Setting $\rho_i^t(x,y) = \langle\psi_{x,i}^t|\psi_{y,i}^t\rangle$, we can decompose the gram matrix before the t-th query as $\rho^{t-1} = \sum_i \rho_i^{t-1}$ and the gram matrix after the query as $\rho^t = \sum_i \rho_i^{t-1} \circ \Delta_i$.

The condition $\Gamma \circ \Delta_i \preceq c\,\Gamma$ then immediately implies that

$$\mathrm{Tr}((\Gamma \circ \rho^t)|v\rangle\langle v|) \leq c\ \mathrm{Tr}((\Gamma \circ \rho^{t-1})|v\rangle\langle v|).$$

For uncomputing queries, the roles of ρ^{t-1} and ρ^t are interchanged, and we obtain the same conclusion from the constraint $\Gamma \preceq c\,\Gamma \circ \Delta_i$. $\square$

REMARK 3.8. *The constraints on Γ given here are expressed differently from Ambainis et al. (2011) and Špalek (2008), the latter using the constraint $\|\Gamma^{1/2}(\Gamma \circ \Delta_i)^{-1/2}\|^2 \leq c$ and $\|(\Gamma\circ\Delta_i)^{1/2}\Gamma^{-1/2}\|^2 \leq c$. It is straightforward to show, however, that these conditions are equivalent to $c^{-1}\Gamma \preceq \Gamma \circ \Delta_i \preceq c\,\Gamma$.*

When the value of c is fixed, the multiplicative bound becomes a semidefinite program. Indeed, setting $W = \Gamma \circ |v\rangle\langle v|$, we have:

$$\begin{aligned} \mathrm{Madv}^{(c)}(\sigma) = \ & \frac{1}{\ln(c)} \ \underset{W \succ 0}{\mathrm{maximize}} \ln\left(\mathrm{Tr}(W\sigma)\right) \\ & \qquad \text{subject to } \mathrm{Tr}(WJ) = 1 \quad \text{and} \\ & \qquad c^{-1}W \preceq W \circ \Delta_i \preceq c\,W \quad \text{for all } i \in [n]. \end{aligned}$$

Thus, we can view the multiplicative adversary bound as a maximization over semidefinite programs.

3.3. Output condition. Thus far, we have seen lower bounds on the problem of *exact coherent* state generation. To obtain a lower bound in the bounded-error setting—coherent or non-coherent—one can minimize the exact coherent bound over the set of valid final gram matrices of a successful algorithm.

We will restrict our discussion to the coherent output condition. As our main results are for functions, by showing lower bounds on the coherent state generation problems F and σ_f associated with a function f, we obtain lower bounds on the query complexity of f by Claims 2.11 and 2.12.

Recall that a successful coherent ϵ-error algorithm $\mathcal{P}$ for the set of target vectors $\{\sigma_x\}$ must satisfy $\Re(\langle \mathcal{P}(x)|V(|\sigma_x\rangle \otimes |\bar{0}\rangle)) \geq \sqrt{1-\epsilon}$ for some unitary V. The set of σ' satisfying this condition can be hard to deal with, so previous works have typically relaxed this condition and used an output condition that defines a larger, simpler set. For example, the original Ambainis output condition minimized over σ' satisfying $\ell_\infty(\sigma - \sigma') \leq 2\sqrt{\epsilon}$ for error parameter ϵ. A stronger output condition based on the γ_2 norm that $\gamma_2(\sigma - \sigma') \leq 2\sqrt{\epsilon}$ was introduced by Høyer *et al.* (2007). As $\gamma_2(v) \geq \ell_\infty(v)$, this output condition defines a smaller set. The γ_2 output condition was later shown to be approximately tight in the sense that if $\gamma_2(\sigma - \sigma') \leq \epsilon$, then there is a unitary V such that $\langle \sigma_x | V | \sigma'_x \rangle \geq 1 - 2\sqrt{\epsilon}$ for all x (Lee *et al.* 2011). While approximately tight in the bounded-error setting, this condition is not strong enough for proving strong direct product theorems, where we need to obtain non-trivial bounds for exponentially small success probabilities.

In this paper, we will work directly with the coherent output condition, or more precisely its reformulation in terms of fidelity from Claim 2.9. The following quantities then give lower bounds for ϵ-error coherent quantum state generation:

Definition 3.9 (Additive and multiplicative bounds).

$$\mathrm{Adv}_\epsilon(\sigma) = \min_\rho \mathrm{Adv}^*(\rho),$$

$$\mathrm{Madv}_\epsilon(\sigma) = \min_\rho \mathrm{Madv}(\rho),$$

where both minimizations are over gram matrices ρ such that

$$\min_{|u\rangle:\||u\rangle\|=1} \mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma \circ |u\rangle\langle u|) \geq \sqrt{1-\epsilon}.$$

In light of Claim 2.9, we can slightly improve one of the bounds in (Lee *et al.* 2011, Lemma 4.8), which compares the tight output condition based on the fidelity to the output condition based on the factorization norm γ_2.

CLAIM 3.10. *Let $\{|a_x\rangle\}, \{|b_x\rangle\}$ be two sets of vectors, and ρ, σ their corresponding gram matrices. Say that*

$$\sqrt{1-\epsilon} = \max_V \min_x \Re(\langle a_x|V|b_x\rangle),$$

where the maximization is taken over all unitary matrices V. Then

$$1 - \sqrt{1-\epsilon} \leq \frac{1}{2}\gamma_2(\rho - \sigma) \leq \sqrt{\epsilon}.$$

PROOF. This directly follows from Claim 2.9 and the relation between the trace distance and fidelity (Lemma 2.5).

$$\begin{aligned} 1 - \mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma \circ |u\rangle\langle u|) &\leq \frac{1}{2}\|(\rho - \sigma) \circ |u\rangle\langle u|\|_{\mathrm{tr}} \\ &\leq \sqrt{1 - \mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma \circ |u\rangle\langle u|)^2}. \quad \square \end{aligned}$$

Note that a multiplicative witness Γ yields a good zero-error multiplicative adversary bound if $\mathrm{Tr}(\Gamma(\sigma \circ |v\rangle\langle v|))$ is large. To obtain a bound for ϵ-error algorithms, we need to show that $\mathrm{Tr}(\Gamma(\rho \circ |v\rangle\langle v|))$ remains large for any gram matrix ρ such that $\mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma \circ |u\rangle\langle u|) \geq \sqrt{1-\epsilon}$ for all unit vectors $|u\rangle$. The following lemma will be useful.

LEMMA 3.11. *Let p, q be two distributions for a discrete random variable A taking values in a finite subset of $\mathbb{R}_{>0}$ (where $\mathbb{R}_{>0}$ denotes the set of positive reals). If $\mathcal{F}(p, q) \geq \sqrt{\delta}$, then*

$$\mathrm{E}_{A \leftarrow q}[A] \geq \delta \left(\mathrm{E}_{A \leftarrow p}[A^{-1}]\right)^{-1}.$$

PROOF. Let $\{a_1, \ldots, a_N\}$ be the support of A. Let $p_i = \Pr_{A \leftarrow p}[A = a_i]$ and $q_i = \Pr_{A \leftarrow q}[A = a_i]$. We need to lower bound the value of the following optimization program:

$$\underset{q_i \geq 0:\ \sum_i q_i = 1}{\text{minimize}} \sum_i q_i a_i \text{ subject to } \mathcal{F}(p, q) \geq \sqrt{\delta}.$$

Introducing vectors $|u\rangle = \sum_i \sqrt{p_i}|i\rangle$ and $|v\rangle = \sum_i \sqrt{q_i}|i\rangle$, and letting $D(A)$ be a diagonal matrix whose (i, i) entry is a_i, this can be rewritten as

$$\begin{aligned}
&\underset{|v\rangle : \|v\| = 1}{\text{minimize}} \langle v|D(A)|v\rangle \text{ subject to } |\langle u|v\rangle|^2 \geq \delta \\
&\quad = \underset{\rho \succeq 0 : \mathrm{Tr}\rho = 1}{\text{minimize}} \mathrm{Tr}[D(A)\rho] \text{ subject to } \mathrm{Tr}[|u\rangle\langle u|\rho] \geq \delta.
\end{aligned}$$

This is a semidefinite program, whose dual can be written as

$$\underset{\lambda \geq 0, \mu}{\text{maximize}}\ \lambda\delta + \mu \text{ subject to } D(A) \succeq \lambda|u\rangle\langle u| + \mu I.$$

Setting $\mu = 0$, this is at least

$$\delta \cdot \underset{\lambda \geq 0}{\text{maximize}}\ \lambda \text{ subject to } D(A) \succeq \lambda|u\rangle\langle u|.$$

Let $|w\rangle = \sum_i \sqrt{p_i/a_i}|i\rangle$. The constraint is equivalent to $I \succeq \lambda|w\rangle\langle w|$, which in turn is equivalent to $\lambda\||w\rangle\langle w|\| = \lambda\|w\|^2 \leq 1$. The lemma then follows from $\|w\|^2 = \sum_i p_i a_i^{-1}$. $\square$

To apply this lemma, we need an upper bound on $\mathrm{E}_{A \leftarrow p}[A^{-1}]$. In our applications, we usually do not know explicitly the distribution p, but we do know its expectation and the extremal values in its support. The next claim allows us to upper bound $\mathrm{E}_{A \leftarrow p}[A^{-1}]$ in terms of these quantities.

CLAIM 3.12. *Let $0 < a_0 \leq \bar{a} \leq a_1$, and A be a random variable with finite support taking values in a bounded set $S \subseteq [a_0, a_1]$. If $\mathrm{E}_{A \leftarrow p}[A] = \bar{a}$, then $\mathrm{E}_{A \leftarrow p}[A^{-1}] \leq \frac{a_0 + a_1 - \bar{a}}{a_0 a_1}$.*

PROOF. $\mathrm{E}_{A \leftarrow p}[A^{-1}]$ is at most the value of the following linear program:

$$\underset{p_a \geq 0}{\text{maximize}} \sum_{a \in S} p_a a^{-1} \text{ subject to } \sum_{a \in S} p_a a = \bar{a}, \quad \sum_{a \in S} p_a = 1.$$

The dual program can be written as

$$\underset{\lambda,\mu}{\text{minimize}}\ \lambda - \bar{a}\mu \text{ subject to } \ \mu a^2 - \lambda a + 1 \leq 0 \ \forall a \in S.$$

Since $a_0 \leq a \leq a_1$, the constraint is satisfied for $\lambda = \frac{a_0+a_1}{a_0 a_1}$ and $\mu = \frac{1}{a_0 a_1}$, which leads to $\mathrm{E}_{A\leftarrow p}[A^{-1}] \leq \frac{a_0+a_1-\bar{a}}{a_0 a_1}$. □

Putting the last two claims together, we get the following corollary which is key to our strong direct product theorem.

COROLLARY 3.13. *Let $a_1 \geq a_0 > 0$ and p be a distribution for a random variable A with finite support taking values in $[a_0, a_1]$. If $\mathrm{E}_{A\leftarrow p}[A] = \bar{a}$ and q is a distribution over $(\mathbb{R}_{>0})^k$ such that $\mathcal{F}(p^{\otimes k}, q) \geq \sqrt{\delta^k}$, then*

$$\mathrm{E}_{(A_1,\ldots,A_k)\leftarrow q}\left(\Pi_{l=1}^k A_l\right) \geq \left(\frac{\delta a_0 a_1}{a_0 + a_1 - \bar{a}}\right)^k.$$

3.4. Comparison of the adversary bounds. We first give a variation of the result by Ambainis *et al.* (2011) that the multiplicative adversary bound is stronger than the additive bound. The main difference with Ambainis *et al.* (2011) is that this claim relies on the bound $\mathrm{Adv}^*(\sigma)$ which is potentially stronger for general quantum state generation problems.

CLAIM 3.14 (Ambainis *et al.* 2011). *For any state generation problem σ*

$$\mathrm{Madv}(\sigma) \geq \mathrm{Adv}^*(\sigma).$$

PROOF. Let Γ be an optimal witness for $\mathrm{Adv}^*(\sigma) = b$, and $|v\rangle$ be the principal eigenvector of $\Gamma \circ (J - \sigma)$. Note that we may assume without loss of generality that $|v\rangle$ corresponds to a positive eigenvalue of $\Gamma \circ (J - \sigma)$. Let $\Gamma' = \Gamma - \mathrm{Tr}((\Gamma \circ \sigma)|v\rangle\langle v|)I$, and notice that Γ' is also a witness for $\mathrm{Adv}^*(\sigma) = b$, satisfying $\mathrm{Tr}(\Gamma'|v\rangle\langle v|) = b$ and

$$\begin{aligned}\mathrm{Tr}((\Gamma' \circ \sigma)|v\rangle\langle v|) &= \mathrm{Tr}((\Gamma \circ \sigma)|v\rangle\langle v|) \\ &\quad - \mathrm{Tr}((\Gamma \circ \sigma)|v\rangle\langle v|) \cdot \mathrm{Tr}((I \circ \sigma)|v\rangle\langle v|) = 0,\end{aligned}$$

as $\mathrm{Tr}((I \circ \sigma)|v\rangle\langle v|) = \mathrm{Tr}(|v\rangle\langle v|) = 1$.

Let $d = \|\Gamma'\|$ and note that $d \geq b$. Finally, for $\kappa > 0$ a small constant to be chosen later, define $\Gamma_m = (I + \kappa(dI - \Gamma'))/(1 + \kappa(d-b))$. Therefore, we have $\mathrm{Tr}(\Gamma_m |v\rangle\langle v|) = 1$ and $\mathrm{Tr}((\Gamma_m \circ \sigma)|v\rangle\langle v|) = (1+\kappa d)/(1+\kappa(d-b))$.

We now show that the condition $c^{-1}\Gamma_m \preceq \Gamma_m \circ \Delta_i \preceq c\Gamma_m$ is satisfied for $c = 1 + \kappa$. We show $(1+\kappa(d-b))(\Gamma_m \circ (c\Delta_i - J)) \succeq 0$ which implies $\Gamma_m \circ (c\Delta_i - J) \succeq 0$ as $1 + \kappa(d-b) > 0$.

$$
\begin{aligned}
&(1+\kappa(d-b))(\Gamma_m \circ (c\Delta_i - J)) \\
&\quad = \big((1+\kappa d)I - \kappa\Gamma'\big) \circ \big((\Delta_i - J) + \kappa\Delta_i\big) \\
&\quad = \kappa(I + \Gamma' \circ (J - \Delta_i)) + \kappa^2(dI - \Gamma') \circ \Delta_i.
\end{aligned}
$$

From the constraint of the additive bound, we know that $I + \Gamma' \circ (J - \Delta_i) \succeq 0$ for all $i \in [n]$. Also as $dI - \Gamma' \succeq 0$, taking the Hadamard product with $\Delta_i \succeq 0$ gives $(dI - \Gamma') \circ \Delta_i \succeq 0$. Therefore, we have $\Gamma_m \circ (c\Delta_i - J) \succeq 0$. One can show $\Gamma_m \circ (cJ - \Delta_i) \succeq 0$ in a similar fashion. This implies that Γ_m is a witness for

$$
\mathrm{Madv}(\sigma) \geq \frac{\ln\left(\frac{1+\kappa d}{1+\kappa(d-b)}\right)}{\ln(1+\kappa)}.
$$

As the above argument holds for any $\kappa > 0$, the claim follows as

$$
\lim_{\kappa \to 0^+} \frac{\ln\left(\frac{1+\kappa d}{1+\kappa(d-b)}\right)}{\ln(1+\kappa)} = b. \qquad \square
$$

Adapting results from Špalek (2008) and Ambainis *et al.* (2011), this implies a strong direct product theorem for $\mathrm{Madv}(\sigma)$ as long as the bound is obtained for $c = 1 + \Omega(1/\mathrm{Adv}^*(\sigma))$. Unfortunately, showing that we can take c bounded away from 1 requires bounding $d = \|\Gamma'\|$, which we do not know how to do for a general state generation problem σ. In general, we can only use this statement in the limit $c \to 1$, in which case the direct product theorem degrades into a direct sum theorem. This is why Ambainis *et al.* (2011) were not able to conclude a strong direct product theorem. We observe that for interesting cases such as F or σ_f, we *can* bound the norm of the witness Γ'. Note that every entry of $J - F$ is either 0 or 1, and similarly, every entry of $J - \sigma_f$ is either 0 or 2. For state generation problems with this property, we can show the following theorem.

Claim 3.15. *Suppose that* $\mathrm{Adv}^*(\sigma) = b$ *and that every entry of* $J - \sigma$ *is either* 0 *or* λ*, for some positive real number* λ*. Then there is a matrix* Γ' *witnessing* $\mathrm{Adv}^*(\sigma) \geq \frac{\lambda b}{\gamma_2(J-\sigma)}$ *such that* $\|\Gamma'\| = \frac{b}{\gamma_2(J-\sigma)}$ *and* $\Gamma' \circ (J - \sigma) = \lambda\Gamma'$*.*

Proof. Let Γ be an optimal witness for $\mathrm{Adv}^*(\sigma)$. Define $\Gamma' = (\gamma_2(J - \sigma))^{-1}(\Gamma \circ (J - \sigma))$. All entries of $J - \sigma$ being either 0 or λ gives the property $(J - \sigma) \circ (J - \sigma) = \lambda(J - \sigma)$. Thus $\Gamma' \circ (J - \sigma) = \lambda\Gamma'$. This implies that Γ' is a feasible witness as

$$\|\Gamma' \circ (J - \Delta_i)\| \leq \frac{\gamma_2(J - \sigma)}{\gamma_2(J - \sigma)}\|\Gamma \circ (J - \Delta_i)\| \leq 1,$$

since $\|A \circ B\| \leq \gamma_2(A) \cdot \|B\|$ for any A, B of the same size. Furthermore, $\|\Gamma'\| = b/\gamma_2(J - \sigma)$ and Γ' witnesses a bound of $\lambda\|\Gamma'\| = \lambda b/\gamma_2(J - \sigma)$. □

For certain state generation problems including F and σ_f, we are thus able to obtain a quantitative version of Claim 3.14.

Claim 3.16. *Suppose that every entry of* $J - \sigma$ *is either* 0 *or* $\lambda \in \mathbb{R}_{>0}$*, and let* $d = \frac{\mathrm{Adv}^*(\sigma)}{\gamma_2(J-\sigma)}$*. Then, for any* $\kappa > 0$*, there is a multiplicative witness* Γ_m *and a vector* $|v\rangle$ *such that*

$$\begin{aligned}
\mathrm{Tr}(\Gamma_m |v\rangle\langle v|) &= 1,\\
\mathrm{Tr}(\Gamma_m(\sigma \circ |v\rangle\langle v|)) &= 1 + \lambda\kappa d,\\
I \preceq \Gamma_m &\preceq (1 + 2\kappa d)I,\\
c^{-1}\Gamma_m \preceq \Gamma_m \circ \Delta_i &\preceq c\, \Gamma_m \text{ for all } i,
\end{aligned}$$

where $c = 1 + \kappa$*. Therefore,* Γ_m *satisfies the constraints of Definition 3.6 and witnesses that*

$$\mathrm{Madv}(\sigma) \geq \frac{\ln(1 + \lambda\kappa d)}{\ln(1 + \kappa)}. \tag{3.17}$$

Proof. From Claim 3.15, there exists Γ witnessing $\mathrm{Adv}^*(\sigma) \geq \lambda d$ such that $\|\Gamma\| = d$. Let $|v\rangle$ be the principal eigenvector of Γ, and $\Gamma_m = I + \kappa(dI - \Gamma)$. Note that we may assume without loss of generality that $|v\rangle$ corresponds to a positive eigenvalue of Γ.

Therefore, we have $\Gamma_m \succeq I$ and $\mathrm{Tr}(\Gamma_m |v\rangle\langle v|) = 1$. As $\Gamma \circ (J - \sigma) = \lambda\Gamma$, it follows that $|v\rangle$ is also a principal eigenvector of $\Gamma \circ (J - \sigma)$, and the objective value achieved by Γ is $\mathrm{Tr}(\Gamma((J-\sigma)\circ|v\rangle\langle v|)) = \lambda d$. Thus, $\mathrm{Tr}(\Gamma(\sigma \circ |v\rangle\langle v|)) = (1-\lambda)d$ and $\mathrm{Tr}(\Gamma_m(\sigma \circ |v\rangle\langle v|)) = 1 + \lambda\kappa d$. The third condition follows from $-dI \preceq \Gamma \preceq dI$.

The fact that the condition $c^{-1}\Gamma_m \preceq \Gamma_m \circ \Delta_i \preceq c\Gamma_m$ is satisfied for $c = 1 + \kappa$ follows by the same argument as in the proof of Claim 3.14. □

We can now show that the bound for $\mathrm{Madv}(\sigma)$ can be obtained with $c = 1 + \Omega(1/\mathrm{Adv}^*(\sigma))$.

CLAIM 3.18. *Suppose that every entry of $J - \sigma$ is either 0 or $\lambda \in \mathbb{R}_{>0}$. Then, there exists $c \geq 1 + \frac{1}{\mathrm{Adv}^*(\sigma)}$ such that*

$$\mathrm{Madv}^{(c)}(\sigma) \geq \frac{\lambda \ln(2)}{2} \mathrm{Adv}^*(\sigma).$$

PROOF. Note that if $J = \sigma$, then $\mathrm{Adv}^*(\sigma) = 0$ and there is nothing to prove. Therefore, we may assume that $J \neq \sigma$, in which case there must exist an entry of $J - \sigma$ equal to $\lambda > 0$. This implies that $\gamma_2(J - \sigma) \geq \lambda$. By the triangle inequality, we also have $\gamma_2(J - \sigma) \leq \gamma_2(J) + \gamma_2(\sigma) \leq 2$ (the fact that $\gamma_2(\sigma) \leq 1$ follows from the factorization $\sigma_{x,y} = \langle \sigma_x | \sigma_y \rangle$). The claim then follows from Claim 3.16 with $\kappa = 1/(\lambda d) \geq 1/\mathrm{Adv}^*(\sigma)$. Specifically, the numerator of (3.17) becomes $\ln(2)$, and the denominator $\ln(1 + 1/(\lambda d)) \leq 1/(\lambda d) \leq 2/(\lambda \mathrm{Adv}^*(\sigma))$ as $1 + x \leq e^x$ and $d \geq \mathrm{Adv}^*(\sigma)/2$. □

4. Strong direct product theorem

We first prove the following theorem, which will lead to both the strong direct product theorem and the XOR lemma in the boolean case.

THEOREM 4.1. *Let σ be a gram matrix for a state generation problem such that all entries of $J - \sigma$ are either 0 or λ, and let $d = (\gamma_2(J - \sigma))^{-1}\mathrm{Adv}^*(\sigma)$. Then for any $\kappa > 0$ and any $\delta \in (0, 1]$*

$$Q^c_{1-\delta^k}\left(\sigma^{\otimes k}\right) \geq \frac{k \ln\left(\delta \frac{1+2\kappa d}{1+\kappa d(2-\lambda)}\right)}{2\ln(1+\kappa)}.$$

PROOF. Let $|v\rangle, \Gamma_m$ satisfy the conditions in Claim 3.16. As a witness for $\sigma^{\otimes k}$ we take $\Gamma_m^{\otimes k}$. Let us first see that this matrix satisfies the multiplicative constraint with the same value $c = 1+\kappa$.

We label the constraint matrices $\Delta_{p,q}$ for $\sigma^{\otimes k}$ by $p \in [k]$ and $q \in [n]$. These are $|\mathcal{D}|^k$-by-$|\mathcal{D}|^k$ matrices where

$$\Delta_{p,q}((x^1,\dots,x^k),(y^1,\dots,y^k)) = \delta_{x^p_q, y^p_q}.$$

In other words, $\Delta_{p,q} = J^{\otimes p-1} \otimes \Delta_q \otimes J^{\otimes k-p}$. Thus, $\Gamma^{\otimes k} \circ \Delta_{p,q} = \Gamma_m^{\otimes p-1} \otimes \Gamma_m \circ \Delta_q \otimes \Gamma_m^{\otimes k-p}$. Since $c^{-1}\Gamma_m \preceq \Gamma_m \circ \Delta_q \preceq c\,\Gamma_m$ for all $q \in [n]$, and $\Gamma_m \succeq 0$, we immediately have

$$c^{-1}\Gamma_m^{\otimes k} \preceq \Gamma_m^{\otimes k} \circ \Delta_{p,q} \preceq c\,\Gamma_m^{\otimes k}$$

for any $p \in [k], q \in [n]$.

To lower bound the objective value, we must lower bound

$$\mathrm{Madv}_{1-\delta^k}(\sigma^{\otimes k}) \geq \frac{1}{\ln(c)} \min_\rho \ln \mathrm{Tr}(\Gamma_m^{\otimes k}(\rho \circ (|v\rangle\langle v|)^{\otimes k})),$$

where the minimum is taken over all positive semidefinite matrices ρ such that $\rho \circ I = I$ and

$$\min_{|u\rangle : \||u\rangle\|=1} \mathcal{F}(\rho \circ |u\rangle\langle u|, \sigma^{\otimes k} \circ |u\rangle\langle u|) \geq \delta^{k/2}. \tag{4.2}$$

Let ρ be any gram matrix satisfying these conditions. Setting $|u\rangle = |v\rangle^{\otimes k}$, condition (4.2) then implies $\mathcal{F}(\rho \circ (|v\rangle\langle v|)^{\otimes k}, (\sigma \circ |v\rangle\langle v|)^{\otimes k}) \geq \delta^{k/2}$ and we can apply Corollary 3.13 with p being the distribution arising from measuring Γ_m on $\sigma \circ |v\rangle\langle v|$, and q the distribution arising from measuring $\Gamma_m^{\otimes k}$ on $\rho \circ (|v\rangle\langle v|)^{\otimes k}$. Note that both $\sigma \circ |v\rangle\langle v|$ and $\rho \circ (|v\rangle\langle v|)^{\otimes k}$ are density matrices, that is positive semidefinite with trace one, so this gives rise to a valid probability distribution. More explicitly, write Γ_m in terms of its eigenvalue decomposition as $\Gamma_m = \sum_i \alpha_i |\xi_i\rangle\langle\xi_i|$. Then define the distribution p over the eigenvalues $\{\alpha_i\}$ of Γ_m as $p(\alpha_i) = \mathrm{Tr}(|\xi_i\rangle\langle\xi_i|\sigma \circ |v\rangle\langle v|)$. Similarly, define q as a distribution over k-tuples of eigenvalues $(\alpha_{i_1},\dots,\alpha_{i_k})$ of Γ

as $q(\alpha_{i_1},\ldots,\alpha_{i_k}) = \mathrm{Tr}(|\xi_{i_1}\rangle\langle\xi_{i_1}| \otimes \cdots \otimes |\xi_{i_k}\rangle\langle\xi_{i_k}|\rho \circ (|v\rangle\langle v|)^{\otimes k})$. By Lemma 2.4, as $\mathcal{F}(\rho \circ (|v\rangle\langle v|)^{\otimes k}, (\sigma \circ |v\rangle\langle v|)^{\otimes k}) \geq \delta^{k/2}$, we also have $\mathcal{F}(p^{\otimes k}, q) \geq \delta^{k/2}$. The properties of Γ_m given in Claim 3.16 give that the extreme values of the support of p are $a_0 = 1, a_1 = 1+2\kappa d$, and the expected value is $\bar{a} = 1 + \lambda\kappa d$. Putting these parameters into Corollary 3.13 gives

$$\mathrm{Tr}(\Gamma_m^{\otimes k}(\rho \circ (|v\rangle\langle v|)^{\otimes k})) \geq \delta^k \left(\frac{1+2\kappa d}{1+\kappa d(2-\lambda)}\right)^k,$$

and in turn

$$\mathrm{Madv}_{1-\delta^k}(\sigma^{\otimes k}) \geq \frac{k\ln(\delta\frac{1+2\kappa d}{1+\kappa d(2-\lambda)})}{\ln(1+\kappa)}. \qquad \square$$

We then obtain the following strong direct product theorem for the quantum query complexity of any function (boolean or not).

THEOREM 1.1. *For any function f, any $\delta \in [2/3, 1]$, and any integer $k > 0$, we have*

$$Q_{1-\delta^{k/2}}(f^{(k)}) \geq \frac{k\ln(3\delta/2)}{8}\mathrm{Adv}^*(F) \geq \frac{k\ln(3\delta/2)}{8000}Q_{1/4}(F).$$

PROOF. Recall that $F(x,y) = \langle f(x)|f(y)\rangle$. Thus, all entries of $J - F$ are either 0 or 1, and $J - F$ satisfies the condition of Theorem 4.1 with $\lambda = 1$. This factorization of F also shows that $\gamma_2(F) \leq 1$, and so $\gamma_2(J-F) \leq \gamma_2(J) + \gamma_2(F) \leq 2$. Applying Theorem 4.1 with $\lambda = 1$ and $\kappa = 1/d$, we obtain

$$Q^c_{1-\delta^k}(F^{\otimes k}) \geq \frac{k\ln(3\delta/2)}{4}\mathrm{Adv}^*(F).$$

This lower bound is for computing $f^{(k)}$ coherently, and we obtain the lower bound for $f^{(k)}$ using Claim 2.11. The second inequality follows from Theorem 3.3. $\square$

5. Boolean functions

5.1. XOR Lemma. We now focus on boolean functions. Before proving the XOR lemma, we prove a strong direct product theorem for the problem of computing a function in the phase. Let $\sigma_f = 2F - J$ be the gram matrix corresponding to computing a boolean function f in the phase.

CLAIM 5.1. *Let $d = \mathrm{Adv}^*(F)$. For any δ, κ,*

$$Q^c_{1-\delta^k}(\sigma_f^{\otimes k}) \geq \frac{k \ln(\delta(1+2\kappa d))}{2\ln(1+\kappa)}.$$

PROOF. Notice that $J - \sigma_f = 2(J-F)$, therefore $(J-\sigma_f)\circ(J-\sigma_f) = 2(J-\sigma_f)$, $\gamma_2(J-\sigma_f) = 2$ and $\mathrm{Adv}^*(\sigma_f) = 2\mathrm{Adv}^*(F)$. The claim then follows from Theorem 4.1 with $\lambda = 2$. □

Setting $\kappa = 1/(\delta d)$, we immediately obtain the strong direct product theorem for σ_f.

COROLLARY 5.2. *For any δ,*

$$Q^c_{1-\delta^k}(\sigma_f^{\otimes k}) \geq \frac{k\delta}{4}\mathrm{Adv}^*(F).$$

Let $f^{\oplus k}$ be the function computing the parity of k independent copies of f. Since computing $f^{\oplus k}$ in the phase is the same as generating the state $\sigma_f^{\otimes k}$, we obtain the XOR lemma from the strong direct product theorem for σ_f and Claim 2.12, plus the fact (Theorem 3.3) that $\mathrm{Adv}^*(F)$ characterizes $Q_{1/4}(F)$.

LEMMA 1.2 (XOR Lemma). *For any boolean function f, any $\delta \in [0,1]$ and any integer $k > 0$,*

$$Q_{(1-\delta^{k/2})/2}(f^{\oplus k}) \geq \frac{k\delta}{8}\mathrm{Adv}^*(F) \geq \frac{k\delta}{8000}Q_{1/4}(F).$$

5.2. Threshold and strong direct product theorems. Finally, we prove a threshold direct product theorem. This will follow from Claim 5.1 together with the following threshold lemma (Unger 2009, Lemma 2).

Lemma 5.3 (Unger 2009). *Let $Y_1, \ldots, Y_k \in \{-1, +1\}$ be random variables, $-1 \le \beta \le 1$ and $C > 0$ be such that*

$$\mathrm{E}\left(\prod_{i \in S} Y_i\right) \le C\beta^{|S|}$$

for all $S \subseteq [k]$. Let λ be such that $\beta \le \lambda \le 1$. Then

$$\Pr\left[\sum_{i=1}^{k} Y_i \ge \lambda k\right] \le Ce^{-kD(1/2+\lambda/2||1/2+\beta/2)}.$$

Theorem 5.4. *For any function f, any $\delta \in [0, 1)$, any μ such that $\frac{1+\sqrt{\delta}}{2} \le \mu \le 1$ and any integers $k, L > 0$, let $\mathcal{P}_i(x_1, \ldots, x_k) \in \{-1, 1\}$ be the i-th output of a T-query algorithm for $f^{(k)}$, where*

$$T \le \frac{k\delta}{L(1-\delta)} \mathrm{Adv}^*(F),$$

and let $X = \{i \in [k] : \mathcal{P}_i(x_1, \ldots, x_k) = f(x_i)\}$. Then,

$$\Pr\left[|X| \ge \mu k\right] \le e^{\frac{k}{L} - kD\left(\mu || \frac{1+\sqrt{\delta}}{2}\right)}.$$

Proof. Let $d = \mathrm{Adv}^*(F)$ and, for any $i \in [k]$ and any set $S \subseteq [k]$, let us consider the random variables $Y_i = \mathcal{P}_i(x_1, \ldots, x_k) \cdot f(x_i) \in \{-1, 1\}$ and the expectations $\beta_S = \mathrm{E}(\prod_{i \in S} Y_i)$. By definition, we have

$$Q_{(1-\beta_S)/2}(f^{\oplus|S|}) \le T.$$

Moreover, we also have from Claims 2.12 and 5.1 that

$$Q_{(1-\beta_S)/2}(f^{\oplus|S|}) \ge \frac{1}{2} Q^c_{1-\beta_S^2}(\sigma_f^{\otimes|S|}) \ge \frac{\ln(\beta_S^2(1+2\kappa d)^{|S|})}{4\ln(1+\kappa)}$$

for any $\kappa > 0$, which together with the previous inequality leads to

$$\beta_S \le (1+\kappa)^{2T}(1+2\kappa d)^{-|S|/2}.$$

For $\kappa = (1-\delta)/(2\delta d)$, this implies $\beta_S \le e^{k/L}\delta^{|S|/2}$. Using Lemma 5.3 with $\beta = \sqrt{\delta}$, $C = e^{k/L}$ and $\lambda = 2\mu - 1$, we then obtain

$$\Pr\left[\sum_{i=1}^{k} Y_i \geq \lambda k\right] \leq e^{\frac{k}{L} - kD\left(\frac{1+\lambda}{2} \| \frac{1+\sqrt{\delta}}{2}\right)}.$$

The theorem then follows from $|X| = (k + \sum_{i=1}^{k} Y_i)/2$. □

In the special case $\mu = 1$, we obtain the following strong direct product theorem for boolean functions (for some values of δ, L this can lead to better parameters than Theorem 1.1, which is nevertheless more general as it also holds for non-boolean functions).

COROLLARY 5.5. *For any boolean function f, any $\delta \in [0, 1)$ and any integers $k, L > 0$,*

$$Q_{1-(e^{1/L}(1+\sqrt{\delta})/2)^k}(f^{(k)}) \geq \frac{k\delta}{L(1-\delta)} \mathrm{Adv}^*(F).$$

Acknowledgements

JR acknowledges support from ARO/NSA under grant W911NF-09-1-0569, and from the action *Mandats de Retour* of the *Politique Scientifique Fédérale Belge*. TL would like to thank Ben Reichardt for many insightful conversations on these topics. The authors also thank Oded Regev for interesting comments and in particular for suggesting to prove the XOR lemma. After completion of this work, the authors learned that a result similar to Claim 3.18, which was the key missing element to prove the strong direct product theorem, was independently proved by Belovs (2011). An extended abstract of this paper has appeared in the Proceedings of the 27th IEEE Conference on Computational Complexity (Lee & Roland 2012).

References

ANDRIS AMBAINIS (2002). Quantum Lower Bounds by Quantum Arguments. Journal of Computer and System Sciences **64**(4), 750–767.

ANDRIS AMBAINIS (2006). Polynomial degree vs. quantum query complexity. Journal of Computer and System Sciences **72**(2), 220–238.

ANDRIS AMBAINIS, ANDREW M. CHILDS, FRANÇOIS LE GALL & SEIICHIRO TANI (2010a). The quantum query complexity of certification. Quantum Information and Computation **10**, 181–188.

ANDRIS AMBAINIS, ANDREW M. CHILDS, BEN W. REICHARDT, ROBERT ŠPALEK & SHENGYU ZHANG (2010b). Any AND-OR Formula of Size N Can Be Evaluated in Time $N^{1/2+o(1)}$ on a Quantum Computer. SIAM Journal on Computing **39**(6), 2513.

ANDRIS AMBAINIS, LOÏCK MAGNIN, MARTIN ROETTELER & JÉRÉMIE ROLAND (2011). Symmetry-assisted adversaries for quantum state generation. In Proceedings of the 26th Annual IEEE Conference on Computational Complexity, 167–177. IEEE Computer Society.

ANDRIS AMBAINIS, ROBERT ŠPALEK & RONALD DE WOLF (2006). A new quantum lower bound method with applications to direct product theorems and time-space tradeoffs. In Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 618–633. ACM. ISBN 1-59593-134-1.

ROBERT BEALS, HARRY BUHRMAN, RICHARD CLEVE, MICHELE MOSCA & RONALD DE WOLF (1998). Quantum Lower Bounds by Polynomials. In Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science, 352. IEEE Computer Society. ISBN 0-8186-9172-7.

ALEKSANDRS BELOVS (2011). Personal communication.

RAJENDRA BHATIA (2007). Positive definite matrices. Princeton University Press.

HARRY BUHRMAN & RONALD DE WOLF (2002). Complexity measures and decision tree complexity: A survey. Theoretical Computer Science **288**(1), 21–43.

ANDREW M. CHILDS, RICHARD CLEVE, STEPHEN P. JORDAN & DAVID YEUNG (2009). Discrete-query quantum algorithm for NAND trees. Theory of Computing **5**, 119–123.

THOMAS M. COVER & JOY A. THOMAS (2006). Elements of Information Theory. Wiley.

ANDREW DRUCKER (2011). Improved Direct Product Theorems for Randomized Query Complexity. In Proceedings of the 26th Annual IEEE Conference on Computational Complexity, 1–11. IEEE Computer Society.

EDWARD FARHI, JEFFREY GOLDSTONE & SAM GUTMANN (2008). A Quantum Algorithm for the Hamiltonian NAND Tree. Theory of Computing **4**, 169–190.

ODED GOLDREICH, NOAM NISAN & AVI WIGDERSON (2011). On Yao's XOR-Lemma. In Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation, volume 6650 of Lecture Notes in Computer Science, 273–301. Springer.

PETER HØYER, TROY LEE & ROBERT ŠPALEK (2007). Negative weights make adversaries stronger. In Proceedings of the 39th Annual ACM Symposium on Theory of Computing, 526–535. ACM. ISBN 978-1-59593-631-8.

RAHUL JAIN, ATTILA PERESZLENYI & PENGHUI YAO (2012). A direct product theorem for bounded-round public coin randomized communication complexity. In Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science. IEEE Computer Society.

HARTMUT KLAUCK, ROBERT ŠPALEK & RONALD DE WOLF (2007). Quantum and Classical Strong Direct Product Theorems and Optimal Time-Space Tradeoffs. SIAM Journal on Computing **36**(5), 1472–1493.

TROY LEE, RAJAT MITTAL, BEN W. REICHARDT, ROBERT ŠPALEK & MARIO SZEGEDY (2011). Quantum query complexity of state conversion. In Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science, 344–353. IEEE Computer Society.

TROY LEE & JÉRÉMIE ROLAND (2012). A strong direct product theorem for quantum query complexity. In Proceedings of the 27th IEEE Conference on Computational Complexity, 234–246. IEEE Computer Society.

TROY LEE, ADI SHRAIBMAN & ROBERT ŠPALEK (2008). A Direct Product Theorem for Discrepancy. In Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, 71–80. IEEE Computer Society.

Nati Linial, Shahar Mendelson, Gideon Schechtman & Adi Shraibman (2007). Complexity measures of sign matrices. Combinatorica **27**, 439–463.

Nati Linial & Adi Shraibman (2009). Lower bounds in communication complexity based on factorization norms. Random Structures and Algorithms **34**(3), 368–394.

Michael A. Nielsen & Isaac L. Chuang (2000). Quantum Computation and Quantum Information. Cambridge University Press.

Ran Raz (1998). A Parallel Repetition Theorem. SIAM Journal on Computing **27**(3), 763–803.

Ben W. Reichardt (2009). Span Programs and Quantum Query Complexity: The General Adversary Bound Is Nearly Tight for Every Boolean Function. In Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science, 544–551. IEEE Computer Society. ISSN 0272-5428.

Ben W. Reichardt (2011). Reflections for quantum query algorithms. In Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms, 560–569. SIAM.

Ben W. Reichardt & Robert Špalek (2008). Span-program-based quantum algorithm for evaluating formulas. In Proceedings of the 40th Annual ACM Symposium on Theory of Computing, 103–112. ACM. ISBN 978-1-60558-047-0.

Ronen Shaltiel (2003). Towards proving strong direct product theorems. Computational Complexity **12**(1-2), 1–22.

Alexander A. Sherstov (2011). Strong direct product theorems for quantum communication and query complexity. In Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, 41–50. ACM.

Robert Špalek (2008). The Multiplicative Quantum Adversary. In Proceedings of the 23rd Annual IEEE Conference on Computational Complexity, 237–248. IEEE Computer Society. ISBN 978-0-7695-3169-4.

Falk Unger (2009). A Probabilistic Inequality with Applications to Threshold Direct-Product Theorems. In Proceedings of the 50th Annual

IEEE Symposium on Foundations of Computer Science, 221–229. IEEE Computer Society.

Manuscript received 30 April 2012

Troy Lee
Centre for Quantum Technologies,
National University of Singapore,
3 Science Drive 2 Blk S-15,
Singapore 117543, Singapore.
troyjlee@gmail.com

Jérémie Roland
NEC Laboratories America,
Princeton, NJ 08540, USA.

Current address of Jérémie Roland:
QuIC - ULB,
Ecole Polytechnique de Bruxelles,
Université Libre de Bruxelles,
1050 Brussels, Belgium.
jroland@ulb.ac.be