Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 115 (2021)

On quantum versus classical query complexity

Scott Aaronson[*], Andris Ambainis[†], Andrej Bogdanov,
Tsun Ming Cheung, Krishnamoorthy Dinesh[‡]

Abstract

Aaronson and Ambainis (STOC 2015, SICOMP 2018) claimed that the acceptance probability of every quantum algorithm that makes q queries to an N-bit string can be estimated to within ϵ by a randomized classical algorithm of query complexity $O_q((N/\epsilon^2)^{1-1/2q})$. We describe a flaw in their argument but prove that the dependence on N in this upper bound is correct for one-query quantum algorithms ($q = 1$). **Update:** Bravyi, Gosset, and Grier had already obtained the improved bound $O(q\epsilon^{-1/q}N^{1-1/2q})$.

Given a quantum algorithm Q equipped with an oracle, let $c_Q(\epsilon)$ denote the minimum query complexity of a randomized classical algorithm that estimates the acceptance probability of Q within ϵ with probability at least 2/3. Aaronson and Ambainis [AA18] asked what is the value of $c_q(N, \epsilon) = \max_Q c_Q(\epsilon)$, where Q ranges over all quantum algorithms that make q queries to an N-bit oracle. They proved that $c_1(N, 1/3) = \Omega(\sqrt{N}/\log N)$[1] and conjectured that $c_q(N, 1/3) = \tilde{\Omega}_q(N^{1-1/2q})$. Moreover, Aaronson and Ambainis claimed an upper bound of $c_q(N, \epsilon) = O_q((N/\epsilon^2)^{1-1/2q})$.

In this note we describe a flaw in Aaronson's and Ambainis's proof of their upper bound on $c_q(N, \epsilon)$. Nevertheless, we show that the dependence on N is correct in the case $q = 1$.

Theorem 1. $c_1(N, \epsilon) = O(\sqrt{N}/\epsilon^2)$.

In Cheung's Master's thesis [Che21] it is proved more generally that $c_q(N, \epsilon) = O(N^{1-9/(2\cdot 9^q)}/\epsilon^2)$. This confirms the non-existence of a property that is testable with $O(1)$ quantum but not with $o(N)$ classical queries, a question that was raised by Buhrman et al. [BFNR08] which served as one of the motivations for Aaronson's and Ambainis's work.

The lower bound of Aaronson and Ambainis was generalized by Tal [Tal20], who showed that $c_q(N, 1/2 - 1/2^{2q}) = \tilde{\Omega}_q(N^{2/3-O(1/q)})$. Recently, Bansal and Sinha [BS21] and Sherstov, Storozhenko, and Wu [SSW21] independently improved it to $c_q(N, (1-\eta)/2) = \Omega_q((N/\log N)^{1-1/2q} \cdot \eta^2)$. It remains open whether the dependence on N is tight for $q \geq 2$.

Update: Bravyi, Gosset, and Grier [BGG21, Theorem 5] had already obtained the bound $c_q(N, \epsilon) = O(q\epsilon^{-1/q}N^{1-1/2q})$, settling the dependence on N for all q.

1 A revised analysis of the Aaronson-Ambainis estimator

To describe the flaw in [AA18] we specialize to the one-query case $q = 1$. First, [AA18] shows that the acceptance probability of a one-query quantum algorithm with oracle $x \in \{-1, 1\}^N$ can be written as $p(x, x)$ for some quadratic bilinear polynomial

$$p(x, y) = x^\top A y = \sum_{i,j=1}^{N} a_{ij} x_i y_j$$

[*] aaronson@cs.utexas.edu. The University of Texas at Austin
[†] andris.ambainis@lu.lv. University of Latvia
[‡] {andrejb, tmcheung}@cse.cuhk.edu.hk, krishnamoorthydinesh@cuhk.edu.hk. The Chinese University of Hong Kong

[1] Their proof applies only to non-adaptive classical algorithms.

ISSN 1433-8092

such that $|p(x,y)| \le 1$ for all $x, y \in \{-1,1\}^N$. The main ingredient is a randomized algorithm for estimating the value of $p(x,y)$ to within ϵ with high probability that reads only $O(\sqrt{N})$ bits of x and y (in expectation). The algorithm consists of two steps:

1. **Variable-splitting step** (Lemma 4.4 in [AA18]): Calculate the following regularity coefficients:

$$\Lambda_{\{1,2\}} := \sum_{ij} a_{ij}^2 \quad \Lambda_{\{1\}} := \sum_i (\sum_j a_{ij})^2 \quad \Lambda_{\{2\}} := \sum_j (\sum_i a_{ij})^2$$

   If any of $\Lambda_{\{1,2\}}$, $\Lambda_{\{1\}}$, and $\Lambda_{\{2\}}$ exceed $\delta = \epsilon^2/N$, replace some variable x_i by $(x_i' + x_i'')/2$ or some variable y_j by $(y_j' + y_j'')/2$, where x_i', x_i'', y_j', y_j'' are new variables. Then repeat variable-splitting.

   Otherwise, proceed to the estimation step.

2. **Estimation step** (Section 4.2 of [AA18]): Sample each input x_i, y_j independently with probability $q = 1/\sqrt{N}$ and output the value of the estimator $\mathbf{P} = (1/q^2) \sum_{\text{sampled } i,j} a_{ij} x_i y_j$.

The correctness of the algorithm is then argued through the following two claims:

Claim 2 (Corollary 4.5 in [AA18])**.** *There exists a choice of variable splittings for which the variable-splitting step terminates after at most $O(1/\delta) = O(N/\epsilon^2)$ iterations.*

Claim 3 (Section 4.3, 4.4 in [AA18])**.** $\mathbf{P}$ *is an unbiased estimator of $p(x,y)$ of variance $O(\delta/q^2) = O(\epsilon^2)$.*

By Claim 3 and Chebyshev's inequality it is concluded that the estimator is accurate with high probability:

Corollary 4. $\Pr[|\mathbf{P} - p(x,y)| = O(\epsilon)] \ge 2/3$.

By Claim 2, the number of variables N' in p after variable splitting is $N' = O(N/\epsilon^2)$. The expected query complexity of the algorithm is then $N'q = O(\sqrt{N}/\epsilon^2)$. This falls a little short of the $O(\sqrt{N}/\epsilon)$ bound claimed in [AA18]. The reason for this minor gap is that in [AA18] the sampling probability q is mistakenly set to $1/\sqrt{N'}$ instead of $1/\sqrt{N}$ (which would result in $O(\delta/q^2) = O(1)$ instead of $O(\epsilon^2)$ in Claim 3).

We now demonstrate that even with the more liberal choice of sampling probability $q = 1/\sqrt{N}$, Corollary 4 (and therefore Claim 3) is incorrect.

A counterexample to Corollary 4

Fix a sufficiently large absolute constant K. We assume $N > K^2/4$ and $\epsilon \le 1/K$. Let $\mathbf{b}_L$ denote the column vector consisting of $L/2$ 1s followed by $L/2$ -1s (assuming L is even). Consider the polynomial $p(x,y) = x^\top A y$ where $x \in \{-1,1\}^{K^2/4+N}, y \in \{-1,1\}^N$, and A is an $(K^2/4+N) \times N$ matrix

$$A = \begin{bmatrix} \frac{\epsilon}{KN} \mathbf{b}_{K^2/4} \cdot \mathbf{b}_N^\top \\ \frac{\epsilon}{2N^{3/2}} H \end{bmatrix}$$

where H denotes the $N \times N$ Walsh-Hadamard matrix.

We claim that $|p(x,y)| \le 1$ on all $\{-1,1\}$ inputs: For $x = (x_0, x_1)$ where $x_0 \in \{-1,1\}^{K^2/4}$, the contribution of the top $K^2/4$ rows is $\frac{\epsilon}{KN} {x_0}^\top \mathbf{b}_{K^2/4} \cdot \mathbf{b}_N^\top y \le K\epsilon/4 \le 1/4$. The contribution of the bottom N rows is at most $\epsilon/2$ (as the Hadamard matrix has spectral norm $\sqrt{N}$), thus $|p(x,y)| \le 1$ for all relevant inputs.

We can calculate $\Lambda_{\{1,2\}} = (\epsilon/KN)^2 \cdot (K^2/4) \cdot N + (\epsilon^2/4N^3) \cdot N^2 \le \epsilon^2/2N < \epsilon^2/(K^2/4+N)$. As for $\Lambda_{\{1\}}$ and $\Lambda_{\{2\}}$, the top $K^2/4$ rows do not affect their value at all as the corresponding entries cancel out, so both of them are determined by H and can be checked to evaluate to $\epsilon^2/2N < \epsilon^2/(K^2/4+N)$. Thus the algorithm does not find it necessary to perform variable splitting.

Now consider what happens when the input is $x = (\mathbf{b}_{K^2/4}, \mathbf{1}_N)$ and $y = \mathbf{b}_N$. The value of the polynomial on this input is

$$p(x,y) = \frac{\epsilon}{KN} \mathbf{b}_{K^2/4}^\top \cdot \mathbf{b}_{K^2/4} \cdot \mathbf{b}_N^\top \cdot \mathbf{b}_N + \frac{\epsilon}{2N^{3/2}} \mathbf{1}_N^\top H y = \frac{K\epsilon}{4} \pm \frac{\epsilon}{2}$$

However, the estimator $\mathbf{P}$ misses the first $K^2/4$ rows with probability $1 - O(K^2/\sqrt{N})$. Conditioned on this, the value of its estimate would have been to within a constant factor as if it was run on the polynomial $(\epsilon/N^{3/2})x^T Hy$ with input $x = \mathbf{1}_N, y = \mathbf{b}_N$, which should produce an estimate of magnitude $O(\epsilon)$ with high probability. For K sufficiently large, the estimated and true value of $p(x, y)$ are likely to be more than $O(\epsilon)$ apart.

In fact, it can be checked that the variance of the estimator $\mathbf{P}$ on the given example is $\Omega(\epsilon^2\sqrt{N})$, which is larger than the $O(\epsilon^2)$ bound from Claim 3.

2 Proof of Theorem 1

To prove Theorem 1, we dispense of the variable splitting step and show that there exists a possibly non-uniform choice of probabilities that makes the estimation step work. The choice of sampling probabilities is derived from the factorial (dual) form of Grothendieck's inequality [Pis12, Page 239] (see also [AAI+16, Lemma A.6]). Let $\|A\|_\square = \max_{x,y\in\{-1,1\}^N} x^T Ay$ denote the cut norm of A and $\|A\| = \max_{\|x\|_2=1}\|Ax\|_2$ denote its spectral norm.

Proposition 5 (Factorial Grothendieck's Inequality)**.** *There exists an absolute constant K_G (Grothendieck's real constant) such that for every $A \in \mathbb{R}^{n\times n}$, there exists $\alpha, \beta \in \mathbb{R}^n_{\geq 0}$ such that $\|\alpha\|_2 = \|\beta\|_2 = 1$ such that for $\widetilde{A} = [\widetilde{a}_{ij}]$ satisfying $a_{ij} = \alpha_i \widetilde{a}_{ij}\beta_j$, $\|\widetilde{A}\| \leq K_G \cdot \|A\|_\square$.*

The estimator $\mathbf{P}(\epsilon)$ outputs the empirical average of $O(1/\epsilon^2)$ samples of the following estimator $\mathbf{P}$: Independently sample variable x_i with probability α_i, variable y_j with probability β_j, and output

$$\mathbf{P} = \sum_{ij} \frac{a_{ij}}{\alpha_i\beta_j} x_i y_j \mathbf{X}_i \mathbf{Y}_j,$$

where $\mathbf{X}_i$ and $\mathbf{Y}_j$ are indicator random variables for the events that x_i and y_j were sampled, respectively, so that $\Pr[\mathbf{X}_i = 1] = \alpha_i$ and $\Pr[\mathbf{Y}_j = 1] = \beta_j$. Then $\mathbf{P}$, and therefore also $\mathbf{P}(\epsilon)$, is an unbiased estimator of $p(x, y)$:

$$\mathsf{E}[\mathbf{P}] = \sum_{ij} a_{ij} x_i y_j = p(x, y).$$

The expected number of queries made by $\mathbf{P}$ is

$$\mathsf{E}\Big[\sum X_i + \sum Y_j\Big] = \sum_i \alpha_i + \sum_j \beta_j \leq \sqrt{N}\sqrt{\sum\nolimits_i \alpha_i^2} + \sqrt{N}\sqrt{\sum\nolimits_j \beta_j^2} = 2\sqrt{N},$$

so $\mathbf{P}(\epsilon)$ makes at most $O(\sqrt{N}/\epsilon^2)$ queries as desired. We now prove

Proposition 6. *Assuming $\|A\|_\square = 1$, $\mathsf{Var}[\mathbf{P}] \leq 3K_G^2$.*

From here, $\mathsf{Var}[\mathbf{P}(\epsilon)] \leq 1/3\epsilon^2$ and so by Chebyshev's inequality $\mathbf{P}(\epsilon)$ is within ϵ of $\mathsf{E}[\mathbf{P}(\epsilon)] = p(x, y)$ with probability at least 2/3, proving Theorem 1.

Proof of Proposition 6. Let $i \neq i'$ and $j \neq j'$. Then

$$\begin{aligned}
\mathsf{Var}(\mathbf{X}_i\mathbf{Y}_j) &= \alpha_i\beta_j(1-\alpha_i\beta_j) \leq \alpha_i\beta_j,\\
\mathsf{Cov}(\mathbf{X}_i\mathbf{Y}_{j'}, \mathbf{X}_i\mathbf{Y}_j) &= \alpha_i\beta_j\beta_{j'}(1-\alpha_i) \leq \alpha_i\beta_j\beta_{j'},\\
\mathsf{Cov}(\mathbf{X}_i\mathbf{Y}_j, \mathbf{X}_{i'}\mathbf{Y}_j) &= \alpha_i\alpha_{i'}\beta_j(1-\beta_j) \leq \alpha_i\alpha_{i'}\beta_j,\\
\mathsf{Cov}(\mathbf{X}_i\mathbf{Y}_j, \mathbf{X}_{i'}\mathbf{Y}_{j'}) &= 0.
\end{aligned}$$

By the sum-of-variances formula,

$$\begin{aligned}\mathsf{Var}[\mathbf{P}] &= \sum_{i,j}\frac{a_{ij}^2}{\alpha_i^2\beta_j^2}\mathsf{Var}(\mathbf{X}_i\mathbf{Y}_j)+\sum_{i,j\neq j'}\frac{a_{ij}c_{ij'}}{\alpha_i^2\beta_j\beta_{j'}}y_jy_{j'}\,\mathsf{Cov}(\mathbf{X}_i\mathbf{Y}_j,\mathbf{X}_i\mathbf{Y}_{j'})+\sum_{i\neq i',j}\frac{a_{ij}c_{i'j}}{\alpha_i\alpha_{i'}\beta_j^2}x_ix_{i'}\,\mathsf{Cov}(\mathbf{X}_{i'}\mathbf{Y}_j,\mathbf{X}_i\mathbf{Y}_j)\\ &\le \sum_{i,j}\frac{a_{ij}^2}{\alpha_i\beta_j}+\sum_i\frac{1}{\alpha_i}(\sum_j a_{ij}y_j)^2+\sum_j\frac{1}{\beta_j}(\sum_i a_{ij}x_i)^2.\end{aligned}$$

Let $\Lambda=\sum_{ij}\frac{a_{ij}^2}{\alpha_i\beta_j}$, $\Lambda_1(y)=\sum_i\frac{1}{\alpha_i}(\sum_j a_{ij}y_j)^2$, and $\Lambda_2(x)=\sum_j\frac{1}{\beta_j}(\sum_i a_{ij}x_i)^2$. We show that each of them at most K_G^2 for all $x,y\in\{-1,1\}^N$.

$$\begin{aligned}\Lambda &= \sum_{i,j}\alpha_i\widetilde{a_{ij}}^2\beta_j\\ &\le \sqrt{\sum_{i,j}\alpha_i^2\widetilde{a_{ij}}^2}\sqrt{\sum_{i,j}\beta_i^2\widetilde{a_{ij}}^2} && \text{[by Cauchy-Schwarz]}\\ &= \sqrt{\sum_i\alpha_i^2\|\widetilde{a}_i\|_2^2}\sqrt{\sum_i\beta_i^2\|\widetilde{a}_i\|_2^2} && [\widetilde{a}_i \text{ is the } i\text{-th row of } \widetilde{A}]\\ &\le \|\widetilde{A}\|^2 && [\|\widetilde{a}_i\|_2\le\|\widetilde{A}\|]\\ &\le K_G^2 && \text{[by Proposition 5]}\\ \Lambda_1(y) &= \sum_i\frac{1}{\alpha_i}(\sum_j\alpha_j\widetilde{a_{ij}}\beta_jy_j)^2\\ &= \sum_i\alpha_i(\sum_j\widetilde{a_{ij}}\beta_jy_j)^2\\ &= \sum_i\alpha_i|(\widetilde{A}(\beta\cdot y))_i|^2 && [(\beta\cdot y)_j=\beta_jy_j]\\ &\le \|\widetilde{A}\|^2 && [\alpha_i\le 1 \text{ and } \|\beta\cdot y\|_2=\|\beta\|_2=1]\\ &\le K_G^2. && \text{[by Proposition 5]}\end{aligned}$$

By symmetry we also get that $\Lambda_2(x)\le K_G^2$ for all $x\in\{-1,1\}^N$, so $\mathsf{Var}[P]\le 3K_G^2$. ☐

Cheung's Master's thesis [Che21] shows that the estimation algorithm can be implemented in time polynomial in N and $1/\epsilon$.

References

[AA18] Scott Aaronson and Andris Ambainis. Forrelation: A problem that optimally separates quantum from classical computing. *SIAM J. Comput.*, 47(3):982–1038, 2018.

[AAI+16] Scott Aaronson, Andris Ambainis, Janis Iraids, Martins Kokainis, and Juris Smotrovs. Polynomials, quantum query complexity, and Grothendieck's inequality. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 25:1–25:19, 2016.

[BFNR08] H. Buhrman, L. Fortnow, I. Newman, and H. Röhrig. Quantum property testing. *SIAM J. Comput.*, 37(5):1387–1400, 2008.

[BGG21] Sergey Bravyi, David Gosset, and Daniel Grier. Classical algorithms for Forrelation, 2021. arXiv 2102.06963.

[BS21] Nikhil Bansal and Makrand Sinha. k-forrelation optimally separates quantum and classical query complexity. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1303–1316. ACM, 2021.

[Che21] Tsun Ming Cheung. Classical simulation of quantum algorithms via Grothendieck's theorem. Master's thesis, The Chinese University of Hong Kong, July 2021.

[Pis12] Gilles Pisier. Grothendieck's theorem, past and present. *Bulletin of the American Mathematical Society*, 49(2):237–323, May 2012.

[SSW21] Alexander A. Sherstov, Andrey A. Storozhenko, and Pei Wu. An optimal separation of randomized and quantum query complexity. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1289–1302. ACM, 2021.

[Tal20] Avishay Tal. Towards optimal separations between quantum and randomized query complexities. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 228–239, 2020.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il