Supplemental Material for "Provable and verifiable quantum advantage in sample complexity"

Marcello Benedetti,[1] Harry Buhrman,[1,2,3] and Jordi Weggemans[2,4]

[1] *Quantinuum, London, United Kingdom*
[2] *QuSoft, The Netherlands*
[3] *University of Amsterdam, Amsterdam, The Netherlands*
[4] *CWI, The Netherlands*

CONTENTS

I. COMPLEMENT SAMPLING USING QUANTUM SAMPLES

We consider the following sampling problem. For some $n \in \mathbb{N}$, let $S \subset \omega := \{0,1\}^n$ be a non-empty subset of cardinality K out of a set of $N := 2^n$ elements. Let $\bar{S} = \omega \setminus S$ be the complement set with cardinality $N - K$. Given access to a quantum sample $|S\rangle$ of the uniform distribution of the elements in S, the task of complement sampling is to return *any* element from $\bar{S}$.

In this section we discuss solutions to this problem where a quantum computer swaps the input state as

$$|S\rangle = \frac{1}{\sqrt{K}} \sum_{x \in S} |x\rangle \qquad \longrightarrow \qquad |\bar{S}\rangle = \frac{1}{\sqrt{N-K}} \sum_{x \notin S} |x\rangle \,.$$

Then, a measurement in the computational basis yields a uniformly random sample from $\bar{S}$.

Hereafter, we use Landau symbols $\mathcal{O}$ and Ω for asymptotic upper and lower bounds, respectively. For cases where polylogarithmic terms are omitted, we use $\tilde{\mathcal{O}}$ and $\tilde{\Omega}$ instead. For N fixed, we write $\mathcal{S}_K$ for the family of all subsets $S \subset \omega$ with cardinality K.

A. Aaronson, Atia and Susskind's construction

We will start by reviewing Aaronson, Atia and Susskind's construction [1], which formed our basis to arrive at our solution for the quantum sample variant to complement sampling. For most results in this work, except for Section III, we found alternative proofs and easier constructions without relying on this construction. However, we think that

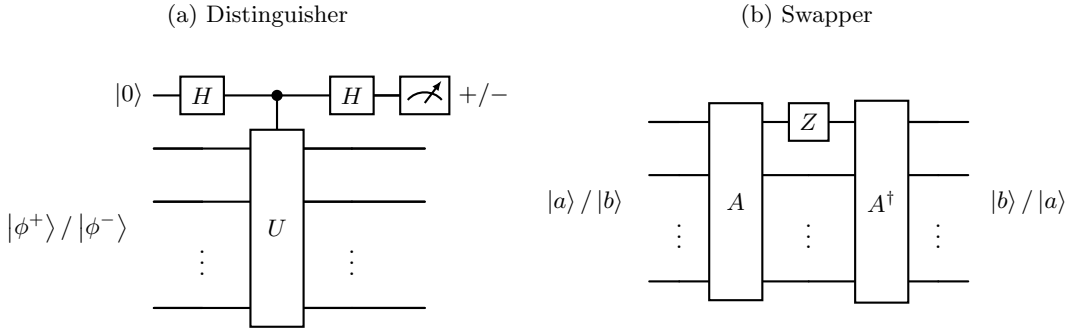

FIG. 1. Quantum circuits for Aaronson, Atia and Susskind's construction [1]. (a) Circuit distinguishing $|\phi^+\rangle$ from $|\phi^-\rangle$ using a unitary U that swaps $|a\rangle$ and $|b\rangle$. (b) Circuit swapping $|a\rangle$ and $|b\rangle$ using a unitary A that distinguishes $|\phi^+\rangle$ and $|\phi^-\rangle$.

poofs based on this construction provide the best intuition for why complement sampling can be perfectly solved on a quantum computer when $K = N/2$.

We start by recalling the four notions of complexities on quantum states from [1]: relative complexity, circuit complexity, swap complexity and distinguishability complexity.

Definition 1 (Relative complexity). *The relative complexity $\mathcal{C}_\epsilon(|a\rangle, |b\rangle)$ of two n-qubit pure quantum states $|a\rangle$,$|b\rangle$ is defined as the minimal number of gates in a circuit C such that*

$$|\langle b| \langle 0 \dots 0| C |a\rangle |0 \dots 0\rangle| \geq 1 - \epsilon.$$

Definition 2 (Circuit complexity). *The circuit complexity $\mathcal{C}_\epsilon(|a\rangle)$ of a n-qubit pure quantum state $|a\rangle$ is defined as the relative complexity with the n-qubit all zero state $|0 \dots 0\rangle$, i.e.*

$$\mathcal{C}_\epsilon(|a\rangle) = \mathcal{C}_\epsilon(|0 \dots 0\rangle, |a\rangle).$$

Definition 3 (Swap complexity). *The swap complexity $\mathcal{S}_\epsilon(|a\rangle, |b\rangle)$ of two n-qubit pure quantum states $|a\rangle$,$|b\rangle$ is defined as the minimal number of gates in a circuit C such that*

$$\frac{1}{2} \big| \langle a| \langle 0 \dots 0| C |b\rangle |0 \dots 0\rangle + \langle b| \langle 0 \dots 0| C |a\rangle |0 \dots 0\rangle \big| \geq 1 - \epsilon.$$

It can be verified that $\mathcal{S}_\epsilon \geq \mathcal{C}_\epsilon$ holds for any $0 \leq \epsilon \leq 1$ [1].

Definition 4 (Distinguishability complexity). *The distinguishability complexity $\mathcal{D}_\epsilon(|a\rangle, |b\rangle)$ of two n-qubit pure quantum states $|a\rangle$,$|b\rangle$ is defined as the minimal number of gates in a circuit C such that*

$$\Big| \|\Pi_1 C |a\rangle |0 \dots 0\rangle\|_2^2 - \|\Pi_1 C |b\rangle |0 \dots 0\rangle\|_2^2 \Big| \geq 1 - \epsilon,$$

where $\Pi_1 = |1\rangle\langle 1| \otimes \mathbb{I}$.

The key result from [1] is an equivalence (in terms of the order of circuit complexity) between a perfect swapper of two orthogonal states, and a perfect distinguisher for the corresponding conjugate states.

Lemma 1 (Adapted from [1], Theorem 2). *Let $0 \leq \epsilon < 1$ and $n \in \mathbb{N}$. Let $|a\rangle$, $|b\rangle$ be orthogonal n-qubit states and let $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$ and $|\phi^-\rangle = \frac{1}{\sqrt{2}}(|a\rangle - |b\rangle)$. The following holds:*

If $\mathcal{S}_\epsilon(|a\rangle, |b\rangle) = T_1$, then we have that $\mathcal{D}_\epsilon(|\phi^+\rangle, |\phi^-\rangle) = \mathcal{O}(T_1)$.

If $\mathcal{D}_\epsilon(|\phi^+\rangle, |\phi^-\rangle) = T_2$, then we have that $\mathcal{S}_\epsilon(|a\rangle, |b\rangle) = \mathcal{O}(T_2)$.

An intuition for this Lemma is given in Fig. 1 where the distinguisher and swapper circuits are defined in terms of each other. This indicates that the number of gates in these circuits should be of the same order.

The following claim is also given in [1, Corollary 1], but we include it to specify the values of ϵ, as this will be needed in Section III.

Proposition 1. *Let $\mathcal{G}$ be a finite, self-inverse gate set. Let $|a\rangle, |b\rangle$ be two orthogonal n-qubit quantum states. If $\mathcal{D}_{4\epsilon}(|a\rangle, |b\rangle) = T$, then $\mathcal{C}_\epsilon(|a\rangle) \geq T$ and $\mathcal{C}_\epsilon(|b\rangle) \geq T$.*

Proof. We prove this by contradiction. Suppose, without loss of generality, that $\mathcal{C}_\epsilon(|a\rangle) < T$. Then, there exists a quantum circuit C with gate complexity strictly less than T such that

$$|\langle a| C |0\ldots0\rangle| \geq 1-\epsilon.$$

Defining $|o\rangle = C|0\ldots0\rangle$, we can write $|o\rangle$ as

$$|o\rangle = \sqrt{1-\alpha}e^{i\theta_1}|a\rangle + \sqrt{\alpha}e^{i\theta_2}|a^\perp\rangle$$

for some $\theta_1, \theta_2 \in [0, 2\pi]$ and $\alpha \in [0,1]$. The assumption $|\langle a|o\rangle| \geq 1-\epsilon$ implies that $\sqrt{1-\alpha} \geq 1-\epsilon$. Now, consider the circuit that receives an unknown state $|z\rangle$, where $z \in \{a, b\}$, applies $C^\dagger$, and measures in the computational basis. The probability of obtaining the all-zero outcome is then given by $|\langle z|o\rangle|^2$. For $|z\rangle = |a\rangle$, we have

$$|\langle a|o\rangle|^2 \geq (1-\epsilon)^2.$$

For $|z\rangle = |b\rangle$, orthogonality implies that

$$|\langle b|o\rangle|^2 = 1 - |\langle a|o\rangle|^2 \leq 1-(1-\epsilon)^2.$$

Thus, the bias of this measurement-based distinguisher is

$$\left||\langle a|o\rangle|^2 - |\langle b|o\rangle|^2\right| \geq (1-\epsilon)^2 - (1-(1-\epsilon)^2) = 1-4\epsilon+2\epsilon^2 \geq 1-4\epsilon.$$

Since $C^\dagger$ has the same gate complexity as C, this gives a distinguisher with gate complexity $< T$ and bias at least $1-4\epsilon$, contradicting the assumption that $\mathcal{D}_{4\epsilon}(|a\rangle, |b\rangle) = T$. The same argument applies if $\mathcal{C}_\epsilon(|b\rangle) < T$ or both $\mathcal{C}_\epsilon(|a\rangle)$ and $\mathcal{C}_\epsilon(|b\rangle)$ are smaller than T. Thus, we conclude that $\mathcal{C}_\epsilon(|a\rangle) \geq T$ and $\mathcal{C}_\epsilon(|b\rangle) \geq T$ must hold. □

B. A perfect swapper for subsets of cardinality $K = N/2$

Let us now show that the task of complement sampling, as stated in the beginning of Section I, can be perfectly solved using only a single quantum sample in the case where $K = N/2$.

Theorem 4 (Quantum complement swapper)**.** *Consider a subset $S \in \mathcal{S}_K$ with $K = N/2$. Then there exists a polynomial-time quantum algorithm which prepares the state $|\bar{S}\rangle$ from $|S\rangle$ and vice versa.*

Proof. Any function $f : \omega \to \{0,1\}$ induces a quantum phase state of the form

$$|y_f\rangle = \frac{1}{\sqrt{N}}\sum_{x\in\omega}(-1)^{f(x)}|x\rangle.$$

In particular, the constant function $f_{\text{con}}(x) = 0$ yields the state $|y_{f_{\text{con}}}\rangle = \frac{1}{\sqrt{N}}\sum_{x\in\{0,1\}^n}|x\rangle$. For every S and $\bar{S}$ with $K = N/2$ there exists a balanced function f_{bal} such that $f_{\text{bal}}(x) = 0$ if $x \in S$ and $f_{\text{bal}}(x) = 1$ otherwise. Then we can write

$$|S\rangle = \frac{1}{\sqrt{2}}\left(|y_{f_{\text{con}}}\rangle + |y_{f_{\text{bal}}}\rangle\right), \qquad |\bar{S}\rangle = \frac{1}{\sqrt{2}}\left(|y_{f_{\text{con}}}\rangle - |y_{f_{\text{bal}}}\rangle\right).$$

It is well known that the Deutsch–Jozsa algorithm [2] can perfectly distinguish the two phase states induced by constant and balanced functions. The Deutsch–Jozsa algorithm consists of a short quantum circuit V that applies $H^{\otimes n}$ followed by the two-outcome measurement operator $\Lambda = \{|0\ldots0\rangle\langle 0\ldots0|, I - |0\ldots0\rangle\langle 0\ldots0|\}$. We copy the measurement outcome to an ancilla qubit using a multi-controlled NOT gate, where we control on the main register being in the zero state. This circuit is illustrated in Fig. 2 (a). Then, by Lemma 1, there exists another quantum circuit V' which only uses two applications of V to swap $|S\rangle$ and $|\bar{S}\rangle$. This circuit is illustrated in Fig. 2 (b) where we ignore the ancilla qubit since, after circuit simplifications, it can be shown to idle in the $|0\rangle$ state throughout. Since V' consists of two applications of V, the multi-controlled NOT gate and two Z gates, it can be executed in time polynomial in n. □

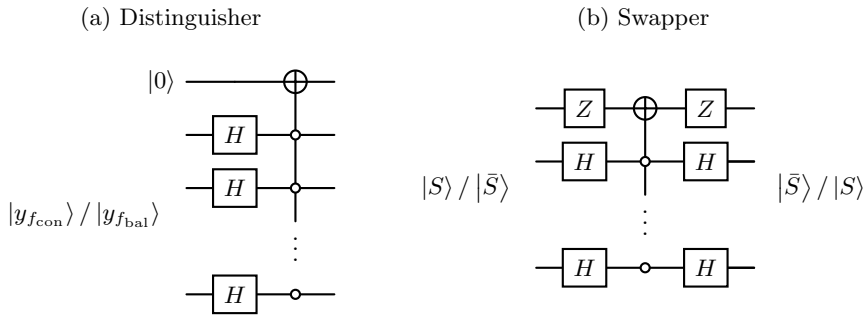

FIG. 2. Distinguisher and swapper for complement sampling. (a) Our distinguisher circuit can be thought of as the Deutsch-Jozsa algorithm followed by a copy of the measurement outcome into an ancilla qubit. (b) The swapper circuit is obtained by plugging our distinguisher circuit into Fig. 1 (b) and simplifying gates. The simplification leads to the ancilla qubit being idle, so we do not include it in this circuit diagram. Our swapper circuit corresponds to the Grover diffusion operator up to a global phase of -1.

Let us take a closer look at what the circuit used in the proof of Theorem 4 actually does, fixing the first qubit (so in the register which contains the qubit to which the outcome of Λ is 'copied') to be in $|0\rangle$ initially. Let the second n-qubit register be in an arbitrary state $|\psi\rangle$. A quick derivation shows that when the first qubit is fixed to be in $|0\rangle$ the circuit implements the mapping

$$|0\rangle|\psi\rangle \mapsto -|0\rangle U|\psi\rangle,$$

where $U = 2|+^n\rangle\langle+^n| - \mathbb{I}$ is the Grover diffusion operator [3]. Noting that $|+^n\rangle = \frac{1}{\sqrt{N}}\left(\sqrt{N-K}\,|\bar{S}\rangle + \sqrt{K}\,|S\rangle\right)$, it can be verified that

$$U|S\rangle = 2\sqrt{\frac{K}{N}\left(1-\frac{K}{N}\right)}\,|\bar{S}\rangle + \left(2\frac{K}{N}-1\right)|S\rangle.$$

Thus, the Grover diffusion operator is an imperfect swapper $U|S\rangle = \sqrt{1-\epsilon}\,|\bar{S}\rangle + \sqrt{\epsilon}\,|S\rangle$ with error $\epsilon = \left(2\frac{K}{N}-1\right)^2$. We see that zero error is achieved when S contains exactly half of the elements, $K = N/2$.

C. Impossibility of a perfect swapper for $K \neq N/2$

Consider again some fixed $N = 2^n$ with $n \in \mathbb{N}$. Let K be the cardinality, and let $\mathcal{S}_K = \{S : |S| = K\}$ be the family of all subsets $S \subset \omega$ that have cardinality K. We now prove that there exists no perfect swapper for any other case then $K = N/2$, given the condition that all auxiliary qubits must return to their initial state. In Section I E we will give a simpler proof. However, the proof given here uses the same construction used in the previous subsections, providing more intuition into the problem at hand.

Proposition 2. *Let $S \in \mathcal{S}_K$. Then for any $K \neq N/2$, there does not exist a quantum circuit which perfectly maps $|S\rangle$ to $|\bar{S}\rangle$ under the condition that all auxiliary qubits return to the $|0\rangle$ state.*

Proof. We will prove our claim using a reductio ad absurdum: suppose that there exists a circuit C which only depends on the cardinality of S (and of course N), and can swap any $|S\rangle$ to $|\bar{S}\rangle$ and vice versa (having $\epsilon = 0$ as per Definition 3). Again, define the states $|S\rangle$ and $|\bar{S}\rangle$ as

$$|S\rangle = \frac{1}{\sqrt{K}}\sum_{x\in S}|x\rangle, \qquad |\bar{S}\rangle = \frac{1}{\sqrt{N-K}}\sum_{x\in\bar{S}}|x\rangle,$$

where $\bar{S} = \omega \setminus S$. As before, define $|\phi^+\rangle := \frac{1}{\sqrt{2}}\left(|S\rangle + |\bar{S}\rangle\right)$ and $|\phi^-\rangle := \frac{1}{\sqrt{2}}\left(|S\rangle - |\bar{S}\rangle\right)$. Then by Lemma 1, there exists another circuit C' such that for all sets $S \in \mathcal{S}_K$

$$\left|\left\|\Pi_1 C'|\phi^+\rangle|0\ldots0\rangle\right\|_2^2 - \left\|\Pi_1 C'|\phi^-\rangle|0\ldots0\rangle\right\|_2^2\right| = 1,$$

which implies that either of the following holds:

$$\left\|\Pi_1 C' \left|\phi^+\right\rangle \left|0\ldots0\right\rangle\right\|_2^2 = 1 \text{ (resp. } 0) \qquad \text{and} \qquad \left\|\Pi_1 C' \left|\phi^-\right\rangle \left|0\ldots0\right\rangle\right\|_2^2 = 0 \text{ (resp. } 1). \tag{1}$$

In the following, we assume that the first case holds (the argument for the second case in parentheses is identical). Now consider two subsets $S_1, S_2 \in \mathcal{S}_K$ and let again $\left|\phi_i^+\right\rangle := \frac{1}{\sqrt{2}}\left(\left|S_i\right\rangle + \left|\bar{S}_i\right\rangle\right)$ and $\left|\phi^-\right\rangle := \frac{1}{\sqrt{2}}\left(\left|S_i\right\rangle - \left|\bar{S}_i\right\rangle\right)$ for $i \in \{1,2\}$. Eq. (1) tells us that

$$\left\|\Pi_1 C' \left|\phi_1^+\right\rangle \left|0\ldots0\right\rangle\right\|_2^2 = 1 \qquad \text{and} \qquad \left\|\Pi_1 C' \left|\phi_2^-\right\rangle \left|0\ldots0\right\rangle\right\|_2^2 = 0, \tag{2}$$

which implies that $\left|\phi_1^+\right\rangle$ and $\left|\phi_2^-\right\rangle$ can be perfectly discriminated. Writing out the states in full, we have

$$\left|\phi_1^+\right\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{K}}\sum_{x\in S_1}\left|x\right\rangle + \frac{1}{\sqrt{N-K}}\sum_{x\in \bar{S}_1}\left|x\right\rangle\right),$$

and

$$\left|\phi_2^-\right\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{K}}\sum_{x\in S_2}\left|x\right\rangle - \frac{1}{\sqrt{N-K}}\sum_{x\in \bar{S}_2}\left|x\right\rangle\right).$$

We have that the fidelity between $\left|\phi_1^+\right\rangle$ and $\left|\phi_2^-\right\rangle$ can be computed as

$$\begin{aligned}\left|\left\langle\phi_1^+\middle|\phi_2^-\right\rangle\right|^2 &= \frac{1}{4}\left|\frac{|S_1\cap S_2|}{K} - \frac{|S_1\cap \bar{S}_2|}{\sqrt{K(N-K)}} + \frac{|\bar{S}_1\cap S_2|}{\sqrt{K(N-K)}} - \frac{|\bar{S}_1\cap \bar{S}_2|}{N-K}\right|^2 \\ &= \frac{1}{4}\left|\frac{|A_1|}{K} - \frac{|A_2|}{\sqrt{K(N-K)}} + \frac{|A_3|}{\sqrt{K(N-K)}} - \frac{|A_4|}{N-K}\right|^2,\end{aligned}$$

where $A_1 := S_1\cap S_2, A_2 := S_1\cap \bar{S}_2, A_3 := \bar{S}_1\cap S_2$ and $A_4 := \bar{S}_1\cap\bar{S}_2$. A simple Venn diagram shows that $\bigcup_j A_j = \omega$ and $\bigcap_j A_j = \emptyset$. Now set $|A_1| = x$, where x should satisfy $\max\{2K-N,0\} \le x \le K$. We can then express the sizes of the other sets A_2, A_3 and A_4 as functions of N, K and x. We have $|A_2| = |S_1\cap\bar{S}_2| = |S_1\setminus(S_2\cap S_1)| = K-x$. A similar argument gives $|A_3| = K-x$. Using the principle of inclusion and exclusion, we have $|S_1\cup S_2| = |S_1| + |S_2| - A_1 = 2K - x$. Hence, $A_4 = N - |S_1\cup S_2| = N - 2K + x$. Our expression for the overlap thus becomes

$$\left|\left\langle\phi_1^+\middle|\phi_2^-\right\rangle\right|^2 = \frac{1}{4}\left|\frac{x}{K} - \frac{N-2K+x}{N-K}\right|^2 = \frac{1}{4}\left(\frac{(2K-N)(K-x)}{K(N-K)}\right)^2. \tag{3}$$

Note that Eq. (3) becomes 0 when $K = N/2$ (independent of x), as expected. For a fixed $1 \le K \le N-1$, we now want to maximize the overlap as a function of x, i.e. solve

$$F_{\max} = \max_{S_1,S_2}\left|\left\langle\phi_1^+\middle|\phi_2^-\right\rangle\right|^2 = \max_{\max\{0,2K-N\}\le x\le K}\frac{1}{4}\left(\frac{(2K-N)(K-x)}{K(N-K)}\right)^2.$$

The double derivative test gives

$$\frac{\partial^2}{\partial x^2}\frac{1}{4}\left(\frac{(2K-N)(K-x)}{K(N-K)}\right)^2 = \frac{(2K-N)^2}{2(K(N-K))^2} \ge 0,$$

which implies that the function is convex in x. Therefore, the maximum value is obtained at the extremal points of the interval. For simplicity, assume for now that $K \le N/2$ so that $\max\{0, 2K-N\} = 0$. For $x = K$, the overlap becomes 0 (this is because $S_1 = S_2$ and thus $\left|\phi_1^+\right\rangle$ is orthogonal to $\left|\phi_2^-\right\rangle$), but for $x = 0$ we find that the overlap becomes $\frac{1}{4}\left(\frac{N-2K}{K-N}\right)^2$. Note that this expression is valid (≤ 1) for $K \le 3N/4$, which is true by assumption. For $K > N/2$ the two extremal points are $x = K$, which again gives overlap 0, and $x = 2K-N$, which gives overlap

$\frac{1}{4}\left(2-\frac{N}{K}\right)^2$. The latter is valid ($\leq 1$) for $K \geq N/4$ which is again true by assumption. Combining both, we find that the maximum overlap as a function of N and K can be

$$\begin{cases} \frac{1}{4}\left(\frac{N-2K}{K-N}\right)^2 & \text{for } 1 \leq K \leq N/2, \\ \frac{1}{4}\left(2-\frac{N}{K}\right)^2 & \text{for } N/2 < K \leq N-1. \end{cases}$$

This is symmetric around $K = N/2$, as substituting $K = N - K'$ with $1 \leq K' \leq N-1$ in $\frac{1}{4}\left(2-\frac{N}{K}\right)^2$ gives back $\frac{1}{4}\left(\frac{N-2K'}{K'-N}\right)^2$, which means that our assumption that $K \leq N/2$ can be made without loss of generality. Therefore, for any $K \neq N/2$ we have that $\left|\phi_1^+\right\rangle$ and $\left|\phi_2^-\right\rangle$ will not be perfectly orthogonal. However, Eq. (2) tells us that we can discriminate these two non-orthogonal states perfectly, which is impossible by the Helstrom bound [4]. Hence, there cannot be a perfect swapper for any other case than $K = N/2$. □

The intuition of why a perfect swapper is only possible when $K = N/2$ follows from the fact that it can be used to construct a distinguisher for the conjugate states, which are generally non-orthogonal unless $K = N/2$.

D. Probabilistic zero-error algorithm for any subset

Now that we know a perfect swapper is not possible for $K \neq N/2$, we want to create the "next best thing": a swapper that sometimes fails, but only knowingly so. In the following, we show that one can trade success probability to achieve this zero-error requirement. An algorithm satisfying this property is known as a *Las Vegas algorithm*. The idea is to rebalance the Grover diffusion operator $U = 2\left|+^n\right\rangle\!\left\langle+^n\right| - \mathbb{I}$ by adding a term proportional to $\mathbb{I}$. To this end, we use one auxiliary qubit and the parameterised gate

$$W(q) = e^{i \arccos\left(\sqrt{q}\right) Y} = \begin{pmatrix} \sqrt{q} & -\sqrt{1-q} \\ \sqrt{1-q} & \sqrt{q} \end{pmatrix}, \tag{4}$$

to construct the following circuit:

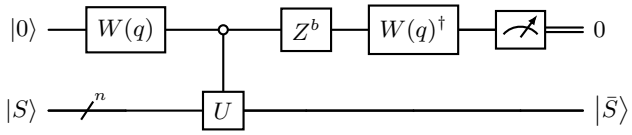

FIG. 3. Probabilistic zero-error algorithm to swap $|S\rangle$ of cardinality K to $\left|\bar{S}\right\rangle$ of cardinality $N-K$.

Here we have that U is controlled on the zero state of the auxiliary qubit, denoted by C_U, and Z is the Z-gate with some power $b \in \{0,1\}$. The role of b is to allow for both positive and negative deviations from the optimal case of $K = N/2$.

Theorem 5 (Zero-error swapper)**.** *Let $S \in \mathcal{S}_K$. Then for any $1 \leq K \leq N-1$, there exists a zero-error (Las Vegas) quantum algorithm that maps $|S\rangle$ to $\left|\bar{S}\right\rangle$, with probability*

$$\frac{\min\{K, N-K\}}{\max\{K, N-K\}}.$$

Proof. We will show that for suitable choices of q and b, the quantum circuit in Fig. 3 achieves the desired transformation and corresponding success probability. The state just before the measurement of the auxiliary qubit can be written as

$$W(q)^\dagger Z^b C_U W(q) \left|0\right\rangle \left|S\right\rangle = \left|0\right\rangle \left(qU + (-1)^b(1-q)\mathbb{I}\right)\left|S\right\rangle + \left|1\right\rangle \left|g\right\rangle$$

with $|g\rangle$ some state we do not care about. Using that $U = 2\left|+^n\right\rangle\!\left\langle+^n\right| - \mathbb{I}$ gives

$$qU + (-1)^b(1-q)\mathbb{I} = 2q\left|+^n\right\rangle\!\left\langle+^n\right| + \left((-1)^b(1-q) - q\right)\mathbb{I}.$$

We will use

$$|+^n\rangle\langle+^n|\,|S\rangle = \frac{K}{N}\,|S\rangle + \sqrt{\frac{K(N-K)}{N^2}}\,|\bar{S}\rangle\,,$$

to obtain

$$\left(qU + (-1)^b(1-q)\mathbb{I}\right)|S\rangle = \left(2q\frac{K}{N} + (-1)^b(1-q) - q\right)|S\rangle + 2q\sqrt{\frac{K(N-K)}{N^2}}\,|\bar{S}\rangle\,. \tag{5}$$

To ensure that $\left(qU + (-1)^b(1-q)\mathbb{I}\right)|S\rangle$ is proportional to $|\bar{S}\rangle$, we set the coefficient of $|S\rangle$ to zero:

$$2q\frac{K}{N} + (-1)^b(1-q) - q = 0.$$

This gives two valid solutions:

- If $b = 0$, the equation becomes $2q\frac{K}{N} + 1 - 2q = 0$, which gives

$$q = \frac{1}{2(1-K/N)}.$$

- If $b = 1$, the equation becomes $2q\frac{K}{N} - 1 = 0$, which gives

$$q = \frac{N}{2K}.$$

It remains to check when these choices yield a valid success probability[1], i.e., when

$$\left\|\left(qU + (-1)^b(1-q)\mathbb{I}\right)|S\rangle\right\|^2 \leq 1$$

Using that in Eq. (5) we made the amplitude on $|S\rangle$ zero and the $|\bar{S}\rangle$-part only depends implicitly on b via q, we obtain

$$\left\|\left(qU + (-1)^b(1-q)\mathbb{I}\right)|S\rangle\right\|^2 = 4q^2\frac{K(N-K)}{N^2}.$$

We find that for the value of q corresponding to $b = 0$, the norm squared is $\frac{K}{N-K}$, which is at most 1 when $K \leq N/2$. When $b = 1$, we obtain $\frac{N-K}{K}$, which is at most 1 when $K \geq N/2$. To summarise, we set

$$\begin{cases} b = 0,\ q = \frac{1}{2(1-K/N)} & \text{if } K < \frac{N}{2}, \\ b = 1,\ q = \frac{N}{2K} & \text{if } K \geq \frac{N}{2}, \end{cases}$$

in Fig. 3. The overall success probability is given by the postselection success probabilities in both cases (the norms squared), which can be captured in a single expression as

$$\frac{\min\{K, N-K\}}{\max\{K, N-K\}}.$$

□

[1] Or equivalently, for what values of q the operator $W(q)$ is indeed unitary.

E. Zero error: optimality, auxiliary qubits and multiple samples

In this section we will show that the algorithm from Section I D is in fact optimal when considering its restricted setting: we assume that there is some designated flag qubit, which when measured to be in the $|0\rangle$ state indicates that the state has been successfully created to ensure zero error (putting the flag being in $|0\rangle$ is without loss of generality). Except for this flag qubit, we require that the circuit uses no other extra auxiliary qubits. We prove the following proposition.

Proposition 3. *For all $1 \leq K \leq N-1$, under the zero-error condition and the use of only one extra ancilla (used as a flag to indicate that the operation was successful), Theorem 5 is optimal in terms of achievable success probability for swapping $|S\rangle$ to $|\bar{S}\rangle$.*

Proof. For cardinality K, let C_K be an arbitrary circuit which probabilistically swaps $|S\rangle$ for subsets $S \in \mathcal{S}_K$ to $|\bar{S}\rangle$. For full generality, we allow the preparation of $|\bar{S}\rangle$ up to an arbitrary global phase factor $e^{i\theta}$, $\theta \in [0, 2\pi]$. Let $0 \leq \epsilon \leq 1$. To achieve success probability $1-\epsilon$, C_K should be able to perform a mapping of the form

$$C_K |0\rangle |S_j\rangle = e^{i\theta_j}\sqrt{1-\epsilon}\,|0\rangle |\bar{S}_j\rangle + \sqrt{\epsilon}\,|1\rangle |g_j\rangle \tag{6}$$

for some arbitrary subset $S_j \in \mathcal{S}_K$, phase $\theta_j \in [0, 2\pi]$, and some garbage state $|g_j\rangle$ (which can absorb its own phase factor). Now take two subsets S_1 and S_2 defined as $S_1 = [K]$ and $S_2 = \omega \setminus [N-K]$ where, with a slight abuse of notation, we write $[K]$ to indicate the K first strings in ω under lexicographical ordering. We distinguish two cases:

- $K \geq N/2$. In this case, we have that the inner product between input states $|S_1\rangle$ and $|S_2\rangle$ is given by (recall that each $|S_j\rangle$ only has positive amplitudes on its computational basis states)

  $$\langle S_1|S_2\rangle = \frac{2K-N}{K}.$$

  The overlap between output states is given by

  $$\left|\langle 0|\langle S_1| C_K^\dagger C_K |0\rangle |S_2\rangle\right| = \left|(1-\epsilon)e^{i(\theta_2-\theta_1)}\langle \bar{S}_1|\bar{S}_2\rangle + \epsilon\langle g_1|g_2\rangle\right|.$$

  By our construction of the subsets we have that $\langle \bar{S}_1|\bar{S}_2\rangle = 0$. Since the overlap must be conserved under unitary transformations, we must have

  $$\frac{2K-N}{K} = \epsilon|\langle g_1|g_2\rangle|.$$

  To find an upper bound on $1-\epsilon$, we can solve the following optimization problem:

  $$\begin{aligned} \max_{\epsilon, |\langle g_1|g_2\rangle|} \quad & 1-\epsilon \\ \text{s.t.} \quad & \frac{2K-N}{K} = \epsilon|\langle g_1|g_2\rangle|, \\ & \epsilon \in [0,1], \\ & |\langle g_1|g_2\rangle| \in [0,1]. \end{aligned}$$

  This gives us

  $$1-\epsilon \leq 1 - \frac{2K-N}{K} = \frac{N-K}{K}.$$

- $K \leq N/2$. Let us first make the observation that $C_K^\dagger$ gives us an approximate swapper to go from $|\bar{S}\rangle$ to $|S\rangle$, where the cardinality of $\bar{S}$ is $K' = N-K$ with $K' \geq N/2$. In particular

  $$\left|\langle 0|\langle S_j| C_K^\dagger |0\rangle |\bar{S}_j\rangle\right| = \left|\langle 0|\langle \bar{S}_j| C_K |0\rangle |S_j\rangle\right| = \sqrt{1-\epsilon},$$

  by Eq. (6). This means we can write the action of $C_K^\dagger$ on $|0\rangle |\bar{S}_j\rangle$ to be of the form

  $$C_K^\dagger |0\rangle |\bar{S}_j\rangle = e^{i\bar{\theta}_j}\sqrt{1-\epsilon}\,|0\rangle |S_j\rangle + \sqrt{\epsilon}\,|1\rangle |\bar{g}_j\rangle,$$

again for some garbage state $|\bar{g}_j\rangle$ and some phase $\bar{\theta}_j \in [0, 2\pi]$. We can now follow the same argument as the previous case, but looking at the inner product between the complement states. By our construction of the subsets we have

$$\left\langle \bar{S}_1 \middle| \bar{S}_2 \right\rangle = \frac{2K' - N}{K'} = \frac{N - 2K}{N - K}.$$

But this must be equal to

$$\left| \langle 0| \left\langle \bar{S}_1 \right| C_K C_K^\dagger |0\rangle \left| \bar{S}_2 \right\rangle \right| = \left| (1-\epsilon) e^{i(\bar{\theta}_2 - \bar{\theta}_1)} \langle S_1 | S_2 \rangle + \epsilon \langle \bar{g}_1 | \bar{g}_2 \rangle \right|,$$

where $\langle S_1 | S_2 \rangle = 0$. Maximizing $1 - \epsilon$ as done in the previous case, but using K' instead of K, we find

$$1 - \epsilon \leq \frac{N}{K'} - 1 = \frac{K}{N - K}.$$

The optimal probability of the two cases can be summarized as $\frac{\min\{K, N-K\}}{\max\{K, N-K\}}$, showing that under these conditions the algorithm from Section I D is optimal. □

Note that the above proof is an arguably easier proof of the result in Section I C, as it only relies on the conservation of the absolute value of the inner product under unitary transformations. There are several more observations we can make in relation to the proof of Proposition 3.

a. Adding extra auxiliary qubits. The proof of Proposition 3 shows that, under the condition of having only one additional ancilla to work as flag qubit, the algorithm of Section I D achieves the optimal success probability. What happens when extra auxiliary qubits are allowed (initialized in $|0 \dots 0\rangle$), but we still require a single flag qubit to make the algorithm zero-error (yet still probabilistic)?

In the $K \geq N/2$ case, it is easy to show that the above proof still holds. For simplicity, we remove the relaxation that the state can be prepared up to an arbitrary global phase, as it does not change the argument. We can write the action of the swapping circuit in this case as

$$C_K |0\rangle |S_j\rangle |0 \dots 0\rangle = \sqrt{1-\epsilon}\, |0\rangle \left|\bar{S}_j\right\rangle |h_j\rangle + \sqrt{\epsilon}\, |1\rangle |G_j\rangle,$$

for some garbage states $|h_j\rangle$ and $|G_j\rangle$, so we still have that the inner product of the output state after applying C_K to two input subset states $|S_1\rangle$ and $|S_2\rangle$ is given by

$$(1-\epsilon) \left\langle \bar{S}_1 \middle| \bar{S}_2 \right\rangle \langle h_1 | h_2 \rangle + \epsilon \langle G_1 | G_2 \rangle = \epsilon \langle G_1 | G_2 \rangle,$$

again using that $\left\langle \bar{S}_1 \middle| \bar{S}_2 \right\rangle = 0$ when $K \geq N/2$.

However, the argument for $K \leq N/2$ no longer works since $|h_1\rangle$ and $|h_2\rangle$ might be two completely different states. This means that to swap backwards using $C_K^\dagger$ one might have to use a different initial state for each different input $\bar{S}$.

An easy example that contradicts the bound in the case when extra auxiliary qubits can be used is the case of $K = 1$ (where the bound gives a maximum success probability very close to 0). An algorithm that only depends on K and perfectly makes $\left|\bar{S}\right\rangle$ is easily given: (i) read out the only element $x \in S$, (ii) create the uniform superposition, (iii) use an additional flag qubit to mark the state $|S\rangle$, which can be done by applying a marking operation that marks all basis states conditioned on not being equal to $|x\rangle$, and (iv) use exact Grover's algorithm to create $\left|\bar{S}\right\rangle$ [5]. This algorithm clearly needs additional auxiliary qubits, as it cannot both have a uniform superposition on n qubits and a flag controlled on $|x\rangle$ (which also needs to be stored in a n-qubit register). Note that in this case the problem is also classically trivial, because after observing a single sample x one has perfectly learned $\bar{S} = \omega \setminus \{x\}$.

b. Multiple copies and no additional auxiliary qubits. The above argument also shows that a joint measurement on multiple copies does not help if one does not have additional auxiliary qubits. That is, suppose one again requires a probabilistic zero-error algorithm that is able to perform the mapping

$$C_K |0\rangle |S_j\rangle^{\otimes k} = \sqrt{1-\epsilon}\, |0\rangle \left|\bar{S}_j\right\rangle |h_j\rangle + \sqrt{\epsilon}\, |1\rangle |G_j\rangle,$$

using k copies of the input state. Considering $K \geq N/2$, the same argument as before gives us a maximum success probability of

$$1 - \epsilon \leq 1 - \left(\frac{2K - N}{K}\right)^k,$$

which is the same success probability one would achieve by repeating our algorithm k times, resetting the flag qubit to $|0\rangle$ whenever $|1\rangle$ is measured and the protocol failed. We currently do not know whether a joint measurement on multiple copies and the use of extra auxiliary qubits would yield an algorithm with higher success probability.

II. CLASSICAL SAMPLES: LOWER AND UPPER BOUNDS

A. Index query model

We will start by adopting a slightly more powerful classical setting, as it allows for easier proofs and it will be useful later when we construct S from pseudorandom permutations, which naturally assumes a query complexity setting. Instead of assuming that the algorithm is given samples from S at random, we assume that it has access to an oracle for which it can query elements of S (assuming arbitrary unknown ordering on each of the possible input sets S). That is, it has access to an oracle O_{index} such that $O_{\text{index}}(i) = y$ with y the ith element from S. Clearly, this is a stronger access model, as the player can generate the random samples by querying O_{index} for randomly generated indices.

Intuitively, one would immediately expect that complement sampling with classical samples is intractable: when there is no structure to the set S (and therefore no structure to $\bar{S}$), the best strategy seems to be to just gather (or in this case, query) as many distinct samples from S and just output anything which is not in the set of collected samples. We will now formalise this intuition and show that it is indeed the correct way to look at the problem. We will use Yao's minimax principle [6] to prove our lower bound. For a fixed input size N and cardinality K, let $F(S, y)$ be a function describing the relation for a subset S from a subset family $\mathcal{S}_K$ and candidate element y such that

$$F(S,y) = \begin{cases} 1 \text{ if } y \in \bar{S} \\ 0 \text{ otherwise.} \end{cases}$$

Let A be the set of all possible deterministic algorithms allowed to make T queries to the index oracle for S. Write $P(a, S) = F(S, a(S))$, which means that $P = 1$ if and only if algorithm $a \in A$ on input S outputs any $y \in \bar{S}$. One can view $P(a, S)$ as a $|A| \times |\mathcal{S}_K|$- matrix where the rows label the different choices of deterministic query algorithms, the columns label the different instances of S, and the corresponding entry is 1 only if the algorithm provides an output satisfying the relation. By the minimax theorem, we have

$$\min_{\mu} \max_{a} e_a^{\intercal} P \mu = \max_{\rho} \min_{S} \rho^{\intercal} P e_S,$$

where e_a (resp. e_S) is a $|A|$-dimensional (resp. $|\mathcal{S}_K|$-dimensional) unit vector, and ρ (resp. μ) is a $|A|$-dimensional (resp. $|\mathcal{S}_K|$-dimensional) vector of non-negative reals that sum to one. Since μ is a probability distribution, we have that the left-hand side describes the probability that the best deterministic T-query algorithm is correct on the hardest distribution over inputs μ. On the right-hand side, $\rho^{\intercal} P e_S$ is the success probability on input S achieved by the randomised algorithm given by probability distribution ρ over deterministic algorithms. Hence, the right-hand side gives the highest worst-case success probability achievable by randomised T-query algorithms. Since this holds for all T, we have that

$$R_\epsilon(F) = \max_{\mu} D_\epsilon^{\mu}(F),$$

where R_ϵ is the worst-case randomized query complexity, achieving success probability $1-\epsilon$ on all inputs, and D_ϵ^{μ} the deterministic query complexity on succeeding on a $(1-\epsilon)$-fraction over all inputs weighted by μ. Note that we now have a maximization over input distributions μ, since we have moved from expressing the smallest success probability given a fixed maximum number of queries T to the largest minimum number of queries T given a fixed fraction of inputs on which we have to be correct. For any fixed candidate distribution μ, we always have

$$R_\epsilon(F) \geq D_\epsilon^{\mu}(F).$$

We first prove the following lemma, showing that when starting with the uniform distribution over all families of subsets $\mathcal{S}_K$ (and thus also a uniform distribution over complementary subsets $\bar{S}$), learning entries of S will leave the resulting conditional distribution over the family of complementary subsets that do not include these elements still uniform.

Lemma 2. *Let $\mu_{\mathcal{S}_K}$ be the uniform distribution over all $S \in \mathcal{S}_K$. Let $X \in \mathcal{S}_K$ be sampled according to $\mu_{\mathcal{S}_K}$, and define $\bar{X} = \omega \setminus X$. Fix a set $Q = \{x_1, \ldots, x_q\}$ containing $q \leq K$ distinct elements from ω. Then, conditioned on $Q \subseteq X$, the distribution of $\bar{X}$ is uniform over $\mathcal{S}_{N-K}^{Q}$, the family of all subsets of size $N-K$ that do not contain any element from Q.*

Proof. By Bayes' rule,

$$\Pr\big[\bar{X} = \bar{S} \mid x_1, \ldots, x_q \in X\big] = \frac{\Pr\big[\bar{X} = \bar{S} \wedge x_1, \ldots, x_q \in X\big]}{\Pr[x_1, \ldots, x_q \in X]}.$$

Using that

$$\Pr[x_1,\ldots,x_q \in X] = \frac{\binom{N-q}{K-q}}{\binom{N}{K}},$$

and the fact that

$$\Pr\big[\bar{X} = \bar{S} \wedge x_1,\ldots,x_q \in X\big] = \begin{cases} \binom{N}{N-K}^{-1} & \text{if } x_1,\ldots,x_q \notin \bar{S}, \\ 0 & \text{otherwise,} \end{cases}$$

we obtain

$$\Pr\big[\bar{X} = \bar{S} \mid x_1,\ldots,x_q \in X\big] = \begin{cases} \binom{N-q}{N-K}^{-1} & \text{if } x_1,\ldots,x_q \notin \bar{S}, \\ 0 & \text{otherwise,} \end{cases}$$

which is uniform over all $\bar{S} \in \mathcal{S}^Q_{N-K}$. □

We can now use Yao's principle to derive our lower bound.

Theorem 6. *Let $\delta \in [0, \frac{1}{2}]$ and $1 \leq K \leq N-1$. Then we have that any randomized algorithm that produces a sample from $\bar{S}$ with probability $\geq 1/2 + \delta$ on all inputs S needs to make at least*

$$N - \frac{2(N-K)}{2\delta+1}.$$

queries to the index oracle for S.

Proof. Consider $\mu_{\mathcal{S}_K}$ to be the uniform distribution over all possible inputs S coming from the family $\mathcal{S}_K$. Let $X \in \mathcal{S}_K$ be sampled according to $\mu_{\mathcal{S}_K}$. Let A be a deterministic algorithm with access to the index oracle O_{index}. We can assume that no element of X is queried twice by A, as there would be no benefit in doing so. Then after q queries, the algorithm has seen q elements of X. Let $Q = \{x_1,\ldots,x_q\}$ be the set of all strings $x_j \in \{0,1\}^n$ that were observed, and write $\mathcal{S}^Q_{N-K}$ for the family of all sets $\bar{S}$ that do not contain the strings $x_j \in Q$. By Lemma 2, we have that the conditional distribution $\mu_{\mathcal{S}^Q_{N-K}}$ over $\mathcal{S}^Q_{N-K}$ is still uniform over these elements. The cardinality of the set $\mathcal{S}^Q_{N-K}$ is then given by

$$\left|\mathcal{S}^Q_{N-K}\right| = \binom{N-q}{N-K}.$$

So for any output y, the probability that it is part of the $\bar{S}'$ for some $\bar{S}' \in \mathcal{S}^Q_{N-K}$ is then given by

$$\frac{\binom{N-q-1}{N-K-1}}{\binom{N-q}{N-K}} = \frac{N-K}{N-q}, \tag{7}$$

which holds independently of the strings that are in Q (except for how many of them there are). Hence, A can only be correct on any such fraction of the inputs. By Yao's principle, this is also the highest success probability a q-query randomized algorithm can achieve on the worst-case input. Then, to be correct with probability at least $\frac{1}{2} + \delta$ we require $(N-K)/(N-q) \geq \frac{1}{2} + \delta$, which is satisfied when

$$q \geq N - \frac{2(N-K)}{2\delta+1}.$$

□

We proceed by giving a matching upper bound, showing that Theorem 6 is tight.

Proposition 4. *For any $\delta \in [0, \frac{1}{2}]$, for any S there exists a randomized algorithm which makes $N - \frac{2(N-K)}{2\delta+1}$ queries to O_{index} and outputs an element $y \in \bar{S}$ with probability $\frac{1}{2} + \delta$.*

Proof. Again, let $Q = \{x_1, \ldots, x_q\}$ be the set of strings $x_j \in S$ that are observed after q queries. The algorithm will now simply sample a uniformly random y from the set $\omega \setminus Q$. Given any S, the probability that this y is from $\bar{S}$ is then given by

$$\Pr\left[y \in \bar{S}\right] = \frac{N-K}{N-q} = \frac{N-K}{N-\left(N-\frac{2(N-K)}{2\delta+1}\right)} = \frac{1}{2}+\delta.$$

□

Note that if $K = \Omega(N)$ and $\delta = \Omega(1)$, then we have $q \geq \Omega(N)$.

B. Exact sample complexity bounds

In the above, we used an index query model because it greatly simplifies the analysis and provides lower bounds for the sample complexity setting. However, it is possible to translate these bounds to the sample complexity setting by computing the probability of observing q unique samples after drawing d samples. Since the results of Section II A hold in terms of *unique* observations of elements from S, we can exactly compute matching upper and lower bounds on the sample complexity.

Sampling from a subset of size K, the probability of getting $q \leq d$ unique samples from d draws (with replacement, each with equal probability) is given by (see for example [7] for a derivation)

$$\Pr[X = q] = \frac{K\left\{{d \atop q}\right\}}{(K-q)!K^d},$$

where $\left\{{d \atop q}\right\}$ is the Stirling number of the second kind, representing the number of ways to partition d objects into q non-empty subsets. Combining this with Eq. (7), which gives the probability of outputting a successful sample conditioned on observing q unique elements, gives a lower bound on the overall success probability of

$$\sum_{q=0}^{q=d} \frac{N\left\{{d \atop q}\right\}}{(K-q)!K^d}\frac{N-K}{N-q}.$$

This in principle gives the exact expression and has a matching upper bound by the same argument as the previous subsection. However, since the expression is cumbersome to work with and we are interested in proving asymptotic lower bounds, we will continue with the query model in the following sections to consider more practically relevant notions of hardness of the problem.

C. Average-case lower bounds

A consequence of Theorem 6 is that the problem is hard on average (with respect to the uniform distribution), as a worst-to-average case reduction with respect to the uniform distribution can easily be constructed. We will use that applying a random permutation on N elements as an operation to any fixed subset of K out of the N elements will result in a uniformly random subset of K elements.

Lemma 3. *Suppose σ is drawn uniformly at random from the set of all permutations on $\{0,1\}^n$. Then, for any fixed subset $S \in \mathcal{S}_K$, the associated subset*

$$S' = \{x \in \{0,1\}^n : \sigma^{-1}(x) \in S\} = \sigma(S).$$

is uniformly drawn from the family $\mathcal{S}_K$.

Proof. We have that S' is the image of S under the permutation σ. We want to understand the distribution of S' when σ is chosen uniformly at random. Fix any subset $T \in \mathcal{S}_K$. We count how many permutations σ satisfy $\sigma(S) = T$. To construct such a permutation, we must:

- define a bijection from S to T, of which there are $K!$ choices;
- define a bijection from $\overline{S}$ to $\overline{T}$, of which there are $(N-K)!$ choices.

So the total number of permutations P such that $\sigma(S) = T$ is

$$|\{\sigma : \sigma(S) = T\}| = K! \cdot (N-K)!,$$

which is the same for all $T \in \mathcal{S}_K$. Since σ is chosen uniformly at random, this means that $S' = \sigma(S)$ is uniformly distributed over $\mathcal{S}_K$. In particular,

$$\Pr[S' = T] = \frac{K! \cdot (N-K)!}{N!} = \frac{1}{\binom{N}{K}} = \frac{1}{|\mathcal{S}_K|},$$

as required. □

We now show that the existence of a q-query randomized algorithm A that achieves a certain guaranteed expected success probability averaged over its internal randomness and the input distribution, implies the existence of a randomized q-query algorithm B that achieves the same guaranteed expected success probability on every input. This yields the desired conclusion that the complement sampling problem is average-case just as hard as worst-case.

Corollary 1. *Let $K = N/2$, $N = 2^n$ for $n \in \mathbb{N}$, and δ be any function such that $\delta(n) \in [0, \frac{1}{2}]$ for all $n \in \mathbb{N}$. Suppose we sample some $S \in \mathcal{S}_K$ uniformly at random. Then any classical randomized algorithm A which makes q queries to O_{index} and satisfies*

$$\Pr\left[A \text{ outputs } y \in \bar{S}\right] \geq \frac{1}{2} + \delta \tag{8}$$

has

$$q \geq N\left(1 - \frac{1}{2\delta + 1}\right),$$

where the probability in Eq. (8) is taken over S and the randomness of A.

Proof. Let σ be a permutation on bit strings of length n chosen uniformly at random, and write σ^{-1} for its inverse. By Lemma 3, starting from any $S \in \mathcal{S}_{N/2}$, $\bar{S} = \omega \setminus S$, we have that the subset S' and its complement $\bar{S}' = \omega \setminus S'$ given by

$$S' = \{x' \mid x' = \sigma(x), x \in S\}, \qquad \bar{S}' = \{y' \mid y' = \sigma(y), y \in \bar{S}\},$$

are uniformly at random from all possible $S \in \mathcal{S}_{N/2}$. Since we are only interested in query (or sample) complexity, we do not care that the time complexity of implementing this permutation σ, nor its inverse σ^{-1}, generally scales exponentially in n (recall $n = \log N$). Now suppose there exists a q-query classical algorithm A that succeeds with expected probability at least $\frac{1}{2} + \delta$ over this distribution of inputs. We can then use A to create a q-query algorithm B that works with a certain success probability on all inputs in the following way:

1. Pick a permutation σ uniformly at random and let σ^{-1} be its inverse.
2. Let $O'_{\text{index}}(i) = \sigma(O_{\text{index}}(i))$. Run algorithm A with $O'_{\text{index}}(i)$, resulting in some y'.
3. Return $y = \sigma^{-1}(y')$.

We then have that B succeeds on any input S with probability at least $\frac{1}{2} + \delta$ as well, since

$$\Pr\left[B \text{ outputs } y \in \bar{S}\right] = \mathbb{E}_{S'}\left[\Pr\left[A \text{ outputs } y' \in \bar{S}'\right]\right] \geq \frac{1}{2} + \delta,$$

using that the expectation of a probability is again a probability. By Theorem 6, we then need at least

$$q \geq N\left(1 - \frac{1}{2\delta + 1}\right)$$

queries to O_{index}, completing the proof. □

D. Hardness for strong pseudorandom subsets

We will now prove our strongest hardness result: we will show that under the existence of one-way functions, there should be families of subsets that are easy to generate and verify, but still hard to perform complement sampling on classically. To achieve this, we will use strong pseudorandom permutations. Following the convention in cryptography, we say that a function f is negligible ($f(n) = \text{negl}(n)$), if for every constant c, we have $f(n) = o\left(\frac{1}{n^c}\right)$.

Definition 5 (Strong pseudorandom permutations)**.** *For some $n \in \mathbb{N}$, let $P : \{0,1\}^n \times \{0,1\}^\lambda \to \{0,1\}^n$ be an efficient keyed permutation. We say that P is a strong pseudorandom permutation if for all probabilistic polynomial-time distinguishers D there exists a negligible function $\text{negl}(n)$ such that*

$$\left|\Pr\left[D^{P_k(\cdot),P_k^{-1}(\cdot)}(1^n)\right] - \Pr\left[D^{\sigma(\cdot),\sigma^{-1}(\cdot)}(1^n)\right]\right| \leq \text{negl}(\lambda),$$

where $k \leftarrow \{0,1\}^\lambda$ is chosen uniformly at random and σ is chosen uniformly at random from the set of permutations on n-bit strings.

It is known that strong pseudorandom permutations can be constructed from pseudorandom permutations [8], which exist if and only if one-way functions exist [9].

We will show the hardness result for cardinality $K = N/2$, as it forms the basis of our proposed advantage experiment. However, the argument can be easily extended to any K superpolynomial in N.

Theorem 7. *For some $n \in \mathbb{N}$, let $P : \{0,1\}^n \times \{0,1\}^\lambda \to \{0,1\}^n$ be a strong pseudorandom permutation with security parameter $\lambda = n$, and P^{-1} its inverse. Given a key $k \in \{0,1\}^\lambda$, let S be the image of $\{0ik : i \in \{0,1\}^{n-1}\}$ under P, and $\bar{S}$ be the image of $\{1ik : i \in \{0,1\}^{n-1}\}$ under P. Let $O_{\text{index}} : [2^{n-1}] \to \{0,1\}^n$ be the oracle that on input i returns the i-th element in S. Picking a key uniformly at random, for all polynomial-time algorithms A that make a polynomial number of queries to O_{index} it must hold that*

$$\Pr\left[A^{O_{\text{index}}} \text{ outputs a } y \in \bar{S}\right] \leq \frac{1}{2} + \text{negl}(n).$$

Proof. We will use a proof by contradiction. Suppose that there exists a polynomial-time algorithm A with query access to some O_{index}, which, when picking a key uniformly at random to construct S, outputs a $y \in \bar{S}$ with probability $\geq \frac{1}{2} + 1/\text{poly}(n)$. We show that if such an A existed it would imply a distinguisher between a strong pseudorandom permutation and truly random one, violating the assumption that S was generated by a strong pseudorandom permutation. First, note that defining S to be the set of all outputs of a uniformly random permutation for which the preimage has '0' as the first bit ensures that S is chosen uniformly at random from $\mathcal{S}_{N/2}$ (Lemma 3). Second, note that any given P already acts as O_{index} when applied to strings from $\{0ik : i \in \{0,1\}^{n-1}\}$. Let F be a samplable family of either (i) permutations or (ii) strong pseudorandom permutations. Then, we can design a distinguisher using $A^{O_{\text{index}}}$ as a subroutine in the following way:

1. We sample a permutation $f \in F$, which is either strong pseudorandom or truly random, and construct O_{index} from f.
2. We run $A^{O_{\text{index}}}$ to obtain a candidate string $\hat{y}$, supposedly from the complement.
3. We check whether $f^{-1}(\hat{y})$ has the first bit equal to '1'. If this is the case, we return ✓. If not, we return ✗.

Let $\delta = 1/\sqrt{N}$, $N = 2^n$, such that $\delta \in \text{negl}(n)$. Then, Corollary 1 readily implies that whenever $q = \text{poly}(n)$ and S is uniformly random (i.e. f comes from a family of truly random permutation), then any q-query classical algorithm (even without the requirement that it runs in polynomial time) succeeds with probability smaller than $\frac{1}{2} + \delta = \frac{1}{2} + \text{negl}(n)$. However, if f comes from a family of strong pseudorandom permutations then by assumption it must hold that we observe a ✓with probability $\frac{1}{2} + 1/\text{poly}(n)$. Hence, under this assumption, we can distinguish strong pseudorandom permutations from truly random ones, by repeating the above distinguisher a polynomial number of times to detect the small polynomial bias towards returning ✓in the pseudorandom case. Since this is not possible by the definition of strong pseudorandomness, such an A cannot exist. Thus the statement must be true. □

The above argument does not extend to pseudorandom permutations that are not strong, as we need to use P^{-1} to check whether the generated string is indeed from $\bar{S}$.

E. Circuit lower bounds for uniform samplers

Finally, we conclude our classical hardness results by showing that we can even prove an exponential *circuit lower bound* on any sampler (quantum or classical) that uses only a polynomial number of classical samples and produces a uniformly random sample from $\bar{S}$ with probability 1. Unlike the previous lower bounds, which hold even when the uniformly random criterion is removed, this bound relies crucially on the requirement that every element from $\bar{S}$ has a non-zero probability to be produced by the sampler. The key idea is to use Kolmogorov complexity of binary strings to arrive at a circuit lower bound.

Definition 6. *The Kolmogorov complexity $\mathcal{K}(x)$ of any binary string $x \in \{0,1\}^n$ is the length of the shortest computer program x^* that can produce this string on the Universal Turing Machine and then halts.*

It is easy to show that there must exist incompressible strings (so with $\mathcal{K}(x) \geq |x|$) for every input length n: there are 2^n binary strings of length n but only $2^n - 1$ binary strings of length strictly less than n to describe the program.

We will formulate our bound in terms of the circuit complexity of a *Boolean circuit*. A Boolean circuit is a directed acyclic graph consisting of logic gates such as AND, OR, and NOT, wired together to compute a Boolean function. The size s of the circuit is defined as the total number of gates (the total number of vertices in the graph). Any Boolean circuit of size s can be described using at most $\mathcal{O}(s \log s)$ bits: each gate can be encoded by its type and the indices of its input wires, each requiring $\mathcal{O}(\log s)$ bits. Consequently, every such circuit can be converted into a binary string x^* of length $|x^*| = \mathcal{O}(s \log s)$, which serves as a program for the universal Turing machine that simulates the circuit. Hence, if a string x has Kolmogorov complexity $\mathcal{K}(x)$ as defined in Definition 6, then any circuit computing x must have size at least $s = \Omega\left(\mathcal{K}(x)/\log \mathcal{K}(x)\right)$.

We will use this to show a circuit lower bound on any classical sampler that uses only a polynomial number of samples and is able to sample uniformly at random from the complementary subset $\bar{S}$.

Theorem 8 (Circuit complexity lower bound with classical samples)**.** *There exists an $S \in \mathcal{S}_K$, with $K = N/2$, such that any sampler, which given as an input $l = \mathrm{poly}(n)$ samples from S produces samples y from $\bar{S}$ uniformly at random, has circuit complexity $\tilde{\Omega}(N)$.*

Proof. We will simply assume that all l samples are distinct, since this not being true only increases the lower bound that follows from our argument. Take a subset $\bar{S}$ with Kolmogorov complexity $\mathcal{K}(\bar{S}) = N/2$. Such a set can be constructed by taking a Kolmogorov random string $z \in \{0,1\}^{N/2}$ of length $N/2$, with each element z_i indexed by a string $i \in \{0,1\}^{n-1}$. We define $y_i = (z_i, i)$ as the concatenation of bit z_i and the string i. For each $i \in \{0,1\}^{n-1}$, we add the string y_i to the set $\bar{S}$. This way, one can fully reconstruct z if one knows all the strings in $\bar{S}$. Then, any sampler M that produces samples from $\bar{S}$ uniformly at random, given the description of any l elements from S, is a description of the set $\bar{S}$ (and thus of the string z), since one can run the sampler until all $N/2$ distinct labels from $\bar{S}$ are observed. Since l elements from S can be described with $l \log N$ bits, the description length of the program for M given l such elements has to be at least $N/2 - l \log N - \mathcal{O}(1)$. Since any circuit description (of the elementary gates) of size s can represent a program for M in $\mathcal{O}(s \log s)$ bits, we must have that the circuit complexity of any such circuit which implements M is lower bounded by

$$\Omega\left(\frac{N/2 - l \log N - \mathcal{O}(1)}{\log\left(N/2 - l \log N - \mathcal{O}(1)\right)}\right) = \tilde{\Omega}(N),$$

when $l = \mathrm{poly}(n)$. □

III. CIRCUIT COMPLEXITY AND DISTINGUISHABILITY OF SUBSET STATES

In the final section we will show that even though we have proven in the above sections that a quantum algorithm can perfectly swap between any $|S\rangle$ and $|\bar{S}\rangle$, there exist pairs of $|S\rangle$, $|\bar{S}\rangle$ so that it is computationally intractable to find out which one was given as an input. As an immediate corollary, this result implies that generally states of the form of $|S\rangle$ must have exponential circuit complexity. This means that in order to turn complement sampling into a suitable quantum advantage experiment, the hardness result for strong PRPs is indeed necessary. We will first prove the easier case of the exact setting and then move on the general case. We will utilize the complexity definitions given in Section I A. Again, we fix the subset cardinality to $K = N/2$ where $N = 2^n$ and n is the number of qubits.

A. Warm-up: exact case

As a warm-up, we first consider the simple example where we care about being able to distinguish $|S\rangle$ and $|\bar{S}\rangle$ perfectly. Consider again the phase states

$$|y_{f_{\text{con}}}\rangle = \frac{1}{\sqrt{N}} \sum_{x\in\omega} |x\rangle\,, \qquad |y_{f_{\text{bal}}}\rangle = \frac{1}{\sqrt{N}} \sum_{x\in\omega} (-1)^{f_{\text{bal}}(x)} |x\rangle\,, \tag{9}$$

for some Boolean function $f_{\text{bal}} : \omega \to \{0,1\}$. Recall that

$$|S\rangle = \frac{1}{\sqrt{2}}(|y_{f_{\text{con}}}\rangle + |y_{f_{\text{bal}}}\rangle), \qquad |\bar{S}\rangle = \frac{1}{\sqrt{2}}(|y_{f_{\text{con}}}\rangle - |y_{f_{\text{bal}}}\rangle),$$

where

$$S = \{x : f_{\text{bal}}(x) = 0\}, \qquad \bar{S} = \{x : f_{\text{bal}}(x) = 1\}.$$

The total number of balanced functions is lower bounded by

$$\binom{N}{N/2} \geq 2^{\Omega(N)}.$$

For every choice of f_{bal}, a circuit can only *exactly* swap $|y_{f_{\text{con}}}\rangle$ to a single $|y_{f_{\text{bal}}}\rangle$. Let $\mathcal{G}$ be any gate set with $g = n^{\mathcal{O}(1)}$ possible choices of gates and qubit indices on which a single (multi-qubit) gate acts. Starting from $|y_{f_{\text{con}}}\rangle$, the total number of states that can be constructed by circuits using M gates from $\mathcal{G}$ is at most g^M. Thus, for some balanced function, the corresponding circuit must have $M = \tilde{\Omega}(N)$ gates, which is the relative complexity $\mathcal{C}(|y_{f_{\text{con}}}\rangle, |y_{f_{\text{bal}}}\rangle)$. Since the swap complexity is lower bounded by the relative complexity [1], there must exist a $|y_{f_{\text{bal}}}\rangle$ such that

$$\mathcal{S}(|y_{f_{\text{con}}}\rangle, |y_{f_{\text{bal}}}\rangle) \geq \tilde{\Omega}(N). \tag{10}$$

Note that any balanced phase state is orthogonal to the constant phase state

$$\langle y_{f_{\text{con}}} | y_{f_{\text{bal}}} \rangle = \frac{1}{N} \sum_{x\in\omega} (-1)^{f_{\text{bal}}(x)} = 0,$$

so we can use Lemma 1 with $|a\rangle := |y_{f_{\text{con}}}\rangle$ and $|b\rangle := |y_{f_{\text{bal}}}\rangle$. We have that if the exact swap complexity $\mathcal{S}(|y_{f_{\text{con}}}\rangle, |y_{f_{\text{bal}}}\rangle) = T$, then the exact distinguishability complexity $\mathcal{D}(|S\rangle, |\bar{S}\rangle) = \mathcal{O}(T)$. Hence, from Eq. (10) it follows that

$$\mathcal{D}(|S\rangle, |\bar{S}\rangle) \geq \tilde{\Omega}(N).$$

B. Approximate case

We consider the same $|y_{f_{\text{con}}}\rangle$ and $|y_{f_{\text{bal}}}\rangle$ as in Eq. (9), but will now investigate the relative circuit complexity when an error ϵ is allowed. Let us consider two balanced functions $f_{\text{bal}}, g_{\text{bal}} : \omega \to \{0,1\}$. From now on, we omit the "bal"-subscript when possible, and write $f = f_{\text{bal}}$ and $g = g_{\text{bal}}$ to ease the presentation. We define $I_{f,g} = |\{x : f(x) = g(x)\}|$ and $N_{f,g} = |\{x : f(x) \neq g(x)\}|$. The overlap between $|y_f\rangle$ and $|y_g\rangle$ is given by

$$|\langle y_f | y_g \rangle| = \frac{1}{N} |I_{f,g} - N_{f,g}| = \frac{1}{N} |2I_{f,g} - N|,$$

using the relation $I_{f,g} + N_{f,g} = N$. If we want $|\langle y_f | y_g \rangle| \leq \epsilon$, then we must have that

$$\frac{1}{N} |2I_{f,g} - N| \leq \epsilon.$$

We consider two cases:

(i) $I_{f,g} \geq \frac{N}{2}$. We find

$$\frac{1}{N}(2I_{f,g} - N) \leq \epsilon \quad \text{if and only if} \quad I_{f,g} \leq \frac{(1+\epsilon)N}{2}.$$

(ii) $I_{f,g} \leq \frac{N}{2}$. Likewise, we have

$$\frac{1}{N}(N - 2I_{f,g}) \leq \epsilon \quad \text{if and only if} \quad I_{f,g} \geq \frac{(1-\epsilon)N}{2}.$$

Combining cases (i) and (ii), we see that a necessary and sufficient condition for $|\langle y_f|y_g\rangle| \leq \epsilon$ is

$$\frac{(1-\epsilon)N}{2} \leq I_{f,g} \leq \frac{(1+\epsilon)N}{2}. \tag{11}$$

The question now is how large the set of functions $\{f\}$ is such that for any pair $f \neq g$ we have that Eq. (11) holds. We can formulate this as a problem in extremal set theory in the following way. Let $\mathcal{F} = \{F\}$ be a family containing $N/2$-subsets from ω. We say that if $x \in F$, then a corresponding function f satisfies $f(x) = 1$. For example, if for $n = 2$ we have $f(00) = 1$, $f(01) = 0$, $f(10) = 0$ and $f(11) = 1$, then $F = \{00, 11\}$. Hence, if $|F \cap G| = l$ for some $F, G \in \mathcal{F}$ with associated functions f, g, respectively, we must have that by definition

$$|\{x : f(x) = g(x) = 1\}| = l.$$

However, we also know that for both F and G the remaining $N/2 - l$ elements in each set must all be distinct: for those we have either $f(x) = 1$ and $g(x) = 0$ or $f(x) = 0$ and $g(x) = 1$. For all remaining x which have $f(x) = g(x) = 0$, i.e., $x \notin F$ and $x \notin G$, we know that there are a total of

$$|\{x : f(x) = g(x) = 0\}| = N - l - 2(N/2 - l) = l.$$

Combining both, we have

$$I_{f,g} = |\{x : f(x) = g(x) = 0\}| + |\{x : f(x) = g(x) = 1\}| = 2l.$$

Therefore, the intersection between two sets F and G completely determines the possible overlaps between the corresponding quantum states $|y_f\rangle$ and $|y_g\rangle$: we have by the condition in Eq. (11) that if for all $F, G \in \mathcal{F}$

$$\frac{(1-\epsilon)N}{4} \leq |F \cap G| \leq \frac{(1+\epsilon)N}{4}, \tag{12}$$

then for all corresponding f, g we have that $|\langle y_f|y_g\rangle| \leq \epsilon$.

We will now use results from extremal set theory to prove our circuit lower bound. Let $\mathcal{H}$ be a family of subsets of N elements, and L be a set of non-negative integers. We say that $\mathcal{H}$ is k-uniform if $|H| = k$ for each $H \in \mathcal{H}$. We have that $\mathcal{H}$ is L-intersecting if $|H \cap G| \in L$ for each pair of distinct $G, H \in \mathcal{H}$. The following lemma was proven by Ray-Chaudhuri and Wilson and provides an upper bound on the size of $\mathcal{H}$ knowing the cardinality of L.

Lemma 4 (Ray-Chaudhuri–Wilson [10])**.** *Let $\mathcal{H}$ be a L-intersecting k-uniform family of subsets of a set of N elements, where $s = |L| \leq k$. Then*

$$|\mathcal{H}| \leq \binom{N}{s}.$$

In terms of N and s, the lemma is easily shown to be tight by considering the family of all s-subsets of any set of size N and $L = \{0, \ldots, s-1\}$.

Using Lemma 4, we can lower bound the size of our family $\mathcal{F}$. At first glance, the bound of Lemma 4 provides the wrong inequality: we are interested in a *lower* bound on the size of $\mathcal{F}$, whilst Lemma 4 provides an *upper* bound. The key idea is that we can consider a family $\mathcal{H}$ for all subsets that do *not* meet the condition of Eq. (12), and upper bound the size of that family to lower bound the size of the remaining family of subsets. After all, for a fixed k the union of L_1-intersecting and L_2-intersecting k-uniform families of subsets of N elements results in the family of all k-uniform subsets if the union of L_1 and L_2 forms the set of all possible intersection sizes. In the following statements we will also use asymptotic $\omega(\cdot)$-notation, which is not to be confused with the universe ω (which does not have an argument).

Lemma 5. *For $\epsilon = \omega(1/N) > 0$, let $\mathcal{F}$ be an L-intersecting set of $N/2$-subsets of a set of N elements, where*

$$L = \{\lceil (1-\epsilon)N/4 \rceil, \lceil (1-\epsilon)N/4 \rceil + 1, \ldots, \lfloor (1+\epsilon)N/4 \rfloor\}.$$

Then we have that

$$|\mathcal{F}| \geq \Omega(2^{N/2}).$$

Proof. Let $\mathcal{H}$ be an L'-intersecting $N/2$-uniform family of subsets of a set of N elements with $L' = \{0, 1, \ldots, N/2\} \setminus L$. We have that the cardinality of L', i.e., $s = |L'|$, can be upper bounded as

$$\begin{aligned} s &\leq 1 + N/2 - ((1+\epsilon)N/4 - (1-\epsilon)N/4 - 2) \\ &= N(1-\epsilon)/2 + 2 \\ &= 3 + \frac{N}{2}(1-\epsilon)\,. \end{aligned}$$

For $\epsilon \geq \frac{6}{N}$ we have $s \leq N/2$ and, using Lemma 4, we get

$$|\mathcal{H}| \leq \binom{N}{s} \leq \binom{N}{3 + \frac{N}{2}(1-\epsilon)}.$$

This means that

$$|\mathcal{F}| = \binom{N}{N/2} - |\mathcal{H}| \geq \binom{N}{N/2} - \binom{N}{3 + \frac{N}{2}(1-\epsilon)} = \Omega(2^{N/2}),$$

if $N(1-\epsilon)/2 + 3 < N/2$, which holds asymptotically if $\epsilon = \omega(1/N)$. □

This implies that for $\epsilon = \omega(1/N) > 0$ there are a total of $\Omega(2^{N/2})$ functions $f = f_{\text{bal}}$ such that the corresponding balanced phase states from the set $\Psi^{\epsilon}_{\text{bal}} = \{|y_{f_{\text{bal}}}\rangle\}$ all have at most ϵ pairwise overlap.

Theorem 9. *For any $\epsilon = \omega(1/N)$, there exists a subset $S \in \mathcal{S}_K$, $K = N/2$, such that*

$$\mathcal{D}_{\epsilon}(|S\rangle, |\bar{S}\rangle) = \tilde{\Omega}(N).$$

Proof. By Lemma 5, for any $\epsilon = \omega(1/N)$, there are $\Omega(2^{N/2})$ possible balanced functions f_{bal} such that all states $|y_{f_{\text{bal}}}\rangle$ have pairwise overlap $\leq \epsilon$. Since any circuit can only ϵ-approximately swap $|y_{f_{\text{con}}}\rangle$ to a single state $|y_{f_{\text{bal}}}\rangle \in \Psi^{\epsilon}_{\text{bal}}$, we need to consider $\Omega(2^{N/2})$ different possible circuits. Again, let $\mathcal{G}$ be any gate set with $g = n^{\mathcal{O}(1)}$ possible choices of gates and qubit indices on which a single (multi-qubit) gate acts. Starting from $|y_{f_{\text{con}}}\rangle$, the total number of states that can be constructed by circuits using M gates from $\mathcal{G}$ is at most g^M. Thus, for some balanced function, the swapper circuit must have $M = \tilde{\Omega}(N)$ gates, which implies $\mathcal{S}_{\epsilon} \geq \Omega(N)$. By Lemma 1, the result immediately follows. □

From Proposition 1, the following corollary immediately follows.

Corollary 2. *For any $\epsilon = \omega(1/N)$, there exists a subset $S \in \mathcal{S}_K$, $K = N/2$, such that*

$$\mathcal{C}_{\epsilon}(|S\rangle) = \tilde{\Omega}(N).$$

As a final remark, Corollary 2 can also qualitatively be obtained as a corollary of Theorem 8, as any provided samples can be hardcoded into a circuit.

[1] S. Aaronson, Y. Atia, and L. Susskind, On the hardness of detecting macroscopic superpositions, arXiv preprint arXiv:2009.07450 (2020).

[2] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **439**, 553 (1992).

[3] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (1996) pp. 212–219.

[4] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics **1**, 231 (1969).

[5] C. Zalka, A Grover-based quantum search of optimal order for an unknown number of marked elements, arXiv preprint quant-ph/9902049 (1999).

[6] A. C.-C. Yao, Probabilistic computations: Toward a unified measure of complexity, in *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)* (IEEE Computer Society, 1977) pp. 222–227.

[7] A. F. Mendelson, M. A. Zuluaga, B. F. Hutton, and S. Ourselin, What is the distribution of the number of unique original items in a bootstrap sample?, arXiv preprint arXiv:1602.05822 (2016).

[8] M. Luby and C. Rackoff, How to construct pseudorandom permutations from pseudorandom functions, SIAM Journal on Computing **17**, 373 (1988).

[9] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*, Chapman & Hall/CRC Cryptography and Network Security Series (CRC Press, 2020).

[10] D. K. Ray-Chaudhuri and R. M. Wilson, On t-designs, Osaka Journal of Mathematics **12**, 737 (1975).