Quantum Attacks on Classical Proof Systems
The Hardness of Quantum Rewinding

Andris Ambainis
University of Latvia and Institute for Advanced Study
Princeton

Ansis Rosmanis
Institute for Quantum Computing
School of Computer Science
University of Waterloo

Dominique Unruh
University of Tartu

April 29, 2014

Abstract. Quantum zero-knowledge proofs and quantum proofs of knowledge are inherently difficult to analyze because their security analysis uses rewinding. Certain cases of quantum rewinding are handled by the results by Watrous (SIAM J Comput, 2009) and Unruh (Eurocrypt 2012), yet in general the problem remains elusive. We show that this is not only due to a lack of proof techniques: relative to an oracle, we show that classically secure proofs and proofs of knowledge are insecure in the quantum setting.

More specifically, sigma-protocols, the Fiat-Shamir construction, and Fischlin’s proof system are quantum insecure under assumptions that are sufficient for classical security. Additionally, we show that for similar reasons, computationally binding commitments provide almost no security guarantees in a quantum setting.

To show these results, we develop the “pick-one trick”, a general technique that allows an adversary to find one value satisfying a given predicate, but not two.

Contents

1 Introduction 2
2 Preliminaries 6
  2.1 Security definitions 7
3 Oracle transformation techniques 8
  3.1 State creation oracles 8
  3.2 Small image oracles 10
4 The pick-one trick 11
  4.1 Additional oracles 12
5 Attacking commitments 13
6 Attacking sigma-protocols 15
  6.1 The computational case 16
7 Attacking Fiat-Shamir 17
  7.1 The computational case 18
8 Attacking Fischlin’s scheme 18
  8.1 The computational case 19

References 20
Symbol index 22
Keyword index 23
A Auxiliary lemmas 24
B Proofs for Section 3 30
  B.1 Lemmas for Section 3.1 31
  B.2 Proofs of Theorem 4 32
C Proof of Theorem 6 33
  C.1 Preliminaries 34
  C.2 Registers and symmetrization of the algorithm 35
  C.3 Representation theory of $S_X$ 37
  C.4 Framework for the proof 38
  C.5 Proof of Lemma 46 39
  C.6 Proof of Lemma 45 40
  C.7 Reduction of Lemma 47 to the $|Y| = 1$ case 43
1 Introduction

Quantum computers threaten classical cryptography. With a quantum computer, an attacker would be able to break all schemes based on the hardness of factoring, or on the hardness of discrete logarithms [28], this would affect most public key encryption and signature schemes in use today. For symmetric ciphers and hash functions, longer key and output lengths will be required due to considerable improvements in brute force attacks [20, 10]. These threats lead to the question: how can classical cryptography be made secure against quantum attacks? Much research has been done towards cryptographic schemes based on hardness assumptions not known to be vulnerable to quantum computers, e.g., lattice-based cryptography. (This is called post-quantum cryptography; see [5] for a somewhat dated survey.) Yet, identifying useful quantum-hard assumptions is only half of the problem. Even if the underlying assumption holds against quantum attackers, for many classically secure protocols it is not clear if they also resist quantum attacks: the proof techniques used in the classical setting often cannot be applied in the quantum world. This raises the question whether it is just our proof techniques that are insufficient, or whether the protocols themselves are quantum insecure. The most prominent example are zero-knowledge proofs. To show the security of a zero-knowledge proof system, one typically uses rewinding. That is, in a hypothetical execution, the adversary’s state is saved, and the adversary is executed several times starting from that state. In the quantum setting, we cannot do that: saving a quantum state means cloning it, violating the no-cloning theorem [35]. Watrous [33] showed that for many zero-knowledge proofs, security can be shown using a quantum version of the rewinding technique. (Yet this technique is not as versatile as classical rewinding. For example, the quantum security of the graph non-isomorphism proof system [19] is an open problem.) Unruh [29] noticed that Watrous’ rewinding cannot be used to show the security of proofs of knowledge; he developed a new rewinding technique to show that so-called sigma-protocols are proofs of knowledge. Yet, in [29] an unexpected condition was needed: their technique only applies to proofs of knowledge with strict soundness (which roughly means that the last message in the interaction is determined by the earlier ones); this condition is not needed in the classical case. The security of sigma-protocols without strict soundness (e.g., graph isomorphism [19]) was left open. The problem also applies to arguments as well (i.e., computationally-sound proof systems, without “of knowledge”), as these are often shown secure by proving that they are actually arguments of knowledge. Further cases where new proof techniques are needed in the quantum setting are schemes involving random oracles. Various proof techniques were developed [6, 37, 30, 8, 31], but all are restricted to specific cases, none of them matches the power of the classical proof techniques.
To summarize: For many constructions that are easy to prove secure classically, proofs in the quantum setting are much harder and come with additional conditions limiting their applicability. The question is: does this only reflect our lack of understanding of the quantum setting, or are those additional conditions indeed necessary? Or could it be that those classically secure constructions are actually insecure quantumly? We show, relative to an oracle, that the answer is indeed yes:

- Sigma-protocols are not necessarily quantum proofs of knowledge, even if they are classical proofs of knowledge. In particular, the strict soundness condition from \[29\] is necessary. (Theorem 17)
- In the computational setting, sigma-protocols are not necessarily quantum arguments, even if they are classical arguments. (Theorem 21)
- The Fiat-Shamir construction \[16\] for non-interactive proofs of knowledge in the random oracle model does not give rise to quantum proofs of knowledge. And in the computational setting, not even to quantum arguments. (Theorems 26 and 27)
- Fischlin’s non-interactive proof of knowledge in the random oracle model \[17\] is not a quantum proof of knowledge. (This is remarkable because in contrast to Fiat-Shamir, the classical security proof of Fischlin’s scheme does not use rewinding.) And in the computational setting, it is not even an argument. (Theorems 29 and 30)
- Besides proof systems, we also have negative results for commitment schemes. The usual classical definition of computationally binding commitments is that the adversary cannot provide openings to two different values for the same commitment. Surprisingly, relative to an oracle, there are computationally binding commitments where a quantum adversary can open the commitment to any value he chooses (just not to two values simultaneously). (Theorem 13)
- The results on commitments in turn allow us to strengthen the above results for proof systems. While it is known that even in the quantum case, sigma-protocols with so-called “strict soundness” (the third message is uniquely determined by the other two) are proofs and proofs of knowledge \[29\], using the computational variant of this property leads to schemes that are not even computationally secure. (Theorems 17, 21, 26, 27, 29, and 30)

Figure 1 gives an overview of the results relating to proofs of knowledge. To the best of our knowledge, these are the first cases where natural classical constructions can be shown to be actually insecure in the quantum setting (albeit relative to an oracle). Before, we only knew that our proof techniques were insufficient.

---

1We stress that negative results were known for pseudorandom functions when the adversary is not only quantum, but can also query the pseudorandom function in superposition \[36\]. Similarly for secret sharing schemes \[14\] and one-time MACs \[7\]. But, in all of these cases, the negative results are shown for the case when the adversary is allowed to interact with the honest parties in superposition. Thus, the cryptographic protocol is different in the classical case and the quantum case.

In contrast, we keep the protocols the same, with only classical communication and only change adversary’s internal power (by allowing it to be a polynomial-time quantum computer which may access quantum oracles).
Our contribution. Our main result are the separations listed in the bullet points above. Towards that goal, we additionally develop several tools that may be of independent interest in quantum cryptographic proofs:

- **Section 4** We develop the “pick-one” trick, a technique for providing the adversary with the ability to compute a value with a certain property, but not two of them. (See “our technique” below.) This technique and the matching lower bound on the adversary’s query complexity may be useful for developing further oracle separations between quantum and classical security. (At least it gives rise to all the separations listed above.)

- **Section 3.1** We show how to create an oracle that allows us to create arbitrarily many copies of a given state $|\Psi\rangle$, but that is not more powerful than having many copies of $|\Psi\rangle$, even if queried in superposition. Again, this might be useful for other oracle separations, too. (The construction of $O_{\Psi}$ in **Section 4** is an example for this.)

- **Section 3.2** We show that a random oracle (with all images chosen independently) can be replaced by an oracle with a polynomial-size image. This is a strengthening of a result by Zhandry [37] which allows us to encode challenges into the random oracle. Our result may be useful for cryptographic proofs where Zhandry’s result is not strong enough. (Zhandry’s result only allowed us to replace a fraction of the oracle’s images.) In addition, it is useful when constructing oracle separations where we want to provide access to arbitrarily many values $v$ according to some distribution: we can construct a quantum oracle that returns random values but that does not give more power than polynomially many random value $v$ (even when queried in superposition). (The construction of $O_S$ in **Section 4** is an example for this.)

Related work. Van der Graaf [32] first noticed that security definitions based on rewinding might be problematic in the quantum setting. Watrous [33] showed how the problems with quantum rewinding can be solved for a large class of zero-knowledge proofs. Unruh [29] gave similar results for proofs of knowledge; however he introduced the additional condition “strict soundness” and they did not cover the computational case (arguments and arguments of knowledge). Our work (the results on sigma-protocols, **Section 6**) shows that these restrictions are not accidental: both strict soundness and statistical security are required for the result from [29] to hold. Protocols that are secure classically but insecure in the quantum setting where given by [36, 7, 14], but only if the quantum adversary can interact with the protocol in superposition (cf. footnote 1). Boneh, Dagdelen, Fischlin, Lehmann, and Schaffner [6] first showed how to correctly define the random oracle in the quantum setting (namely, the adversary has to have superposition access to it). For the Fiat-Shamir construction (using random oracles as modeled by [5]), an impossibility result was given by Dagdelen, Fischlin, and Gagliardoni [12]. However, their impossibility only shows that security of Fiat-Shamir cannot be shown using extractors that do not perform quantum rewinding but such quantum rewinding is possible and used in the existing positive results from [33] [29] which would also not work in a model without quantum rewinding.

Our technique. The schemes we analyze are all based on sigma-protocols which have the special soundness property: In a proof of a statement $s$, given two accepting conversations $(com, ch, resp)$ and $(com, ch', resp')$, one can efficiently extract a witness for $s$. (The commitment $com$ and the response $resp$ are sent by the prover, and the challenge $ch$ by the verifier.) In the classical case, we can ensure that the prover cannot produce one accepting conversation without having enough information to produce two. This is typically proven by rewinding the prover to

---

2They do allow extractors that restart the adversary with the same classical randomness from the very beginning. But due to the randomness inherent in quantum measurements, the adversary will then not necessarily reach the same state again. They also do not allow the extractor to use a purified (i.e., unitary) adversary to avoid measurements that introduce randomness.
get two conversations. So in order to break the schemes in the quantum case, we need to give the prover some information that allows him to succeed in one interaction, but not in two.

To do so, we use the following trick (we call it the pick-one trick): Let $S$ be a set of values (e.g., accepting conversations). Give the quantum state $|\Psi\rangle := \frac{1}{\sqrt{|S|}} \sum_{x \in S} |x\rangle$ to the adversary. Now the adversary can get a random $x \in S$ by measuring $|\Psi\rangle$. However, on its own that is not more useful than just providing a random $x \in S$. So in addition, we provide an oracle that applies the unitary $O_F$ with $O_F|\Psi\rangle = -|\Psi\rangle$ and $O_F|\Psi^\perp\rangle = |\Psi^\perp\rangle$ for all $|\Psi^\perp\rangle$ orthogonal to $|\Psi\rangle$. Now the adversary can use (a variant of) Grover’s search starting with state $|\Psi\rangle$ to find some $x \in S$ that satisfies a predicate $P(x)$ of his choosing, as long as $|S|/|\{x \in S : P(x)\}|$ is polynomially bounded. Note however: once the adversary did this, $|\Psi\rangle$ is gone, he cannot get a second $x \in S$.

How do we use that to break proofs of knowledge? The simplest case is attacking the sigma-protocol itself. Assume the challenge space is polynomial. (I.e., $|ch|$ is logarithmic.) Fix a commitment $com$, and let $S$ be the set of all $(ch, \text{resp.})$ that form an accepting conversation with $com$. Give $com$ and $|\Psi\rangle$ to the malicious prover. (Actually, in the full proof we provide an oracle $O_F$ that allows us to get $|\Psi\rangle$ for a random $com$.) He sends $com$ and receives a challenge $ch'$. And using the pick-one trick, he gets $(ch, \text{resp.}) \in S$ such that $ch = ch'$. Thus sending $\text{resp.}$ will make the verifier accept.

This in itself does not constitute a break of the protocol. A malicious prover is allowed to make the verifier accept, as long as he knows a witness. Thus we need to show that even given $|\Psi\rangle$ and $O_F$, it is hard to compute a witness. Given two accepting conversations $(com, ch, \text{resp.})$ and $(com, ch', \text{resp'})$ we can compute a witness. So we need that given $|\Psi\rangle$ and $O_F$, it is hard to find two different $x, x' \in S$. We show this below (under certain assumptions on the size of $S$, see [Theorem 6], [Corollary 9]). Thus the sigma-protocol is indeed broken: the malicious prover can make the verifier accept using information that does not allow him to compute a witness. (The full counterexample will need additional oracles, e.g., for membership test in $S$ etc.) Counterexamples for the other constructions (Fiat-Shamir, Fischlin, etc.) are constructed similarly.

We stress that this does not contradict the security of sigma-protocols with strict soundness [29]. Strict soundness implies that there is only one response per challenge. Then $|S|$ is polynomial and it becomes possible to extract two accepting conversations from $|\Psi\rangle$ and $O_F$.

The main technical challenge is to prove that given $|\Psi\rangle$ and $O_F$, it is hard to find two different $x, x' \in S$. This is done using the representation-theoretic form of “quantum adversary” lower bound method for quantum algorithms [1][2]. The method is based on viewing a quantum algorithm as a sequence of transformations on a bipartite quantum system that consists of two registers: one register $H_A$ that contains the algorithm’s quantum state and another register $H_I$ that contains the information which triples $(com, ch, \text{resp})$ belong to $S$. The algorithm’s purpose is to obtain two elements $x_1, x_2 \in S$ using only a limited type of interactions between $H_A$ and $H_I$. (From a practical perspective, a quantum register $H_I$ holding the membership information about $S$ would be huge. However, we do not propose to implement such a register. Rather, we use it as a tool to prove a lower bound which then implies a corresponding lower bound in the usual model where $S$ is accessed via oracles.)

We then partition the state-space of $H_I$ into subspaces corresponding to group representations of the symmetry group of $H_I$ (the set of all permutations of triples $(com, ch, \text{resp})$ that satisfy some natural requirements). Informally, these subspaces correspond to possible states of algorithm’s knowledge about the input data: having no information about any $s \in S$, knowing one value $x \in S$, knowing two values $x_1, x_2 \in S$ and so on.

The initial state in which the algorithm has $|\Psi\rangle$ corresponds to $H_I$, being in the state “the algorithm knows one $x \in S$". (This is very natural because measuring $|\Psi\rangle$ gives one value $x \in S$ and there is no way to obtain two values $x \in S$ from this state with a non-negligible probability.) We then show that each application of the available oracles (such as $O_F$ and the membership test for $S$) can only move a tiny part of the state in $H_I$ from the “the algorithm knows one
$x \in S$ subspace of $\mathcal{H}_I$ to the “the algorithm knows two $x \in S$” subspace. Therefore, to obtain two values $x_1, x_2 \in S$, we need to apply the available oracles a large number of times.

While the main idea is quite simple, implementing it requires a sophisticated analysis of the representations of the symmetry group of $\mathcal{H}_I$ and how they evolves when the oracles are applied.

Actually, below we prove an even stronger result: We do not wish to give the state $|\Psi\rangle$ as input to the adversary. (Because that would mean that the attack only works with an input that is not efficiently computable, even in our relativized model.) Thus, instead, we provide an oracle $O_{\Psi}$ for efficiently constructing this state. But then, since the oracle can be invoked arbitrarily many times, the adversary could create two copies of $|\Psi\rangle$, thus easily obtaining two $x, x' \in S$! Instead, we provide an oracle $O_{\Psi}$ that provides a state $|\Sigma\Psi\rangle$ which is a superposition of many $|\Psi\rangle = |\Psi(y)\rangle$ for independently chosen sets $S_y$. Now the adversary can produce $|\Sigma\Psi\rangle$ and using a measurement of $y$, get many states $|\Psi(y)\rangle$ for random $y$’s, but no two states $|\Psi(y)\rangle$ for the same $y$. Taking these additional capabilities into account complicates the proof further, as does the presence of additional oracles that are needed, e.g., to construct the prover (who does need to be able to get several $x \in S$).

Organization. Section 2 introduces conventions and security definitions. Section 4 develops the pick-one trick. Section 5 shows the insecurity of computationally binding commitments, Section 6 that of sigma-protocols, Section 7 that of the Fiat-Shamir construction, and Section 8 that of Fischlin’s construction. Appendix 3 describes our additional oracle techniques: oracles for creating copies of a state $|\Psi\rangle$, and oracles with small ranges.

2 Preliminaries

Security parameter. As usual in cryptography, we assume that all algorithms are parametric in a security parameter $\eta$. Furthermore, parameters of said algorithms can also implicitly depend on the security parameter. E.g., if we say “Let $\ell$ be a superlogarithmic integer. Then $A(\ell)$ runs in polynomial time.,” then this formally means “Let $\ell$ be a superlogarithmic function. Then the running time of $A(\eta, \ell(\eta))$ is a polynomially-bounded function of $\eta$.”

Misc. $x \overset{\$}{\leftarrow} M$ means that $x$ is uniformly randomly chosen from the set $M$. $x \leftarrow A(y)$ means that $x$ is assigned the classical output of the (usually probabilistic or quantum) algorithm $A$ on input $y$.

Quantum mechanics. For space reasons, we cannot give an introduction to the mathematics of quantum mechanics used here. We refer the reader to, e.g., [25]. A quantum state is a vector of norm 1 in a Hilbert space, written $|\Psi\rangle$. Then $\langle\Psi|$ is its dual. TD($\rho, \rho'$) denotes the trace distance between mixed states $\rho, \rho'$. We write short TD($|\Psi\rangle, |\Psi'\rangle$) for TD($|\Psi\rangle\langle\Psi|, |\Psi'\rangle\langle\Psi'|$). SD($X; Y$) in contrast is the statistical distance between random variables $X$ and $Y$.

Oracles. We make heavy use of oracles in this paper. Formally, an oracle $O$ is a unitary transformation on some Hilbert space $\mathcal{H}$. An oracle algorithm $A$ with access to $O$ (written $A^O$) is then a quantum algorithm which has a special gate for applying the unitary $O$. $O$ may depend on the security parameter. $O$ may be probabilistic in the sense that at the beginning of the execution, the unitary $O$ is chosen according to some distribution (like the random oracle in cryptography). However, $O$ may not be probabilistic in the sense that $O$, when queried on the same value twice, gives two different random answers (like an encryption oracle for a probabilistic encryption scheme would). Such a behavior would be difficult to define formally when allowing queries to $O$ in superposition. When defining $O$, we use the shorthand $O(x) := f(x)$ to denote
\( \mathcal{O}[x, y] := \mathcal{O}[x, y \oplus f(x)] \). We call an oracle of this form classical. Our classical algorithms will only access oracles of this form. We stress that even for a classical oracle \( \mathcal{O} \), a quantum algorithm can query \( \mathcal{O}(x) \) in superposition of different \( x \). We often give access to several oracles \( (\mathcal{O}_1, \mathcal{O}_2, \ldots) \) to an algorithm. This can be seen as a specific case of access to a single oracle by setting \( \mathcal{O} | i \rangle | \Psi \rangle := | i \rangle \otimes \mathcal{O}_i | \Psi \rangle \).

In our setting, oracles are used to denote a relativised world in which those oracles happen to be efficiently computable. If a unitary \( U \) is implemented by an efficient quantum circuit, \( U^\dagger \) can also be implemented by an efficient quantum circuit. We would expect this also to hold in a relativised setting. Thus for any oracle \( \mathcal{O} \), algorithms should have access to their inverses, too. In our work this is ensured because all oracles defined here are self-inverse (\( \mathcal{O} = \mathcal{O}^\dagger \)).

### 2.1 Security definitions

A sigma-protocol for a relation \( R \) is a three message proof system. It is described by the lengths \( \ell_{com}, \ell_{ch}, \ell_{resp} \) of the messages, a polynomial-time prover \((P_1, P_2)\) and a polynomial-time verifier \( V \). The first message from the prover is \( com \leftarrow P_1(s, w) \) with \((s, w) \in R\) and is called commitment, the uniformly random reply from the verifier is \( ch \leftarrow \{0, 1\}^{\ell_{ch}} \) (called challenge), and the prover answers with \( resp \leftarrow P_2(ch) \) (the response). We assume \( P_1, P_2 \) to share state. Finally \( V(s, com, ch, resp) \) outputs whether the verifier accepts.

We will make use of the following standard properties of sigma-protocols. Note that we have chosen to make the definition stronger by requiring honest entities (simulator, extractor) to be classical while we allow the adversary to be quantum.

**Definition 1 (Properties of sigma-protocols)** Let \((\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R)\) be a sigma-protocol. We define:

- **Completeness**: For all \((s, w) \in R\), \( Pr[ok = 0 : com \leftarrow P_1(s, w), ch \leftarrow \{0, 1\}^{\ell_{ch}}, resp \leftarrow P_2(ch), ok \leftarrow V(s, com, ch, resp)] \) is negligible.

- **Perfect special soundness**: There is a polynomial-time classical algorithm \( E_\Sigma \) such that for any \((s, com, ch, resp, ch', resp') \) with \( ch \neq ch' \), we have that \( Pr[(s, w) \notin R \land ok = ok' = 1 : ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch', resp'), w \leftarrow E_\Sigma(s, com, ch, resp, ch', resp')] = 0 \).

- **Computational special soundness**: There is a polynomial-time classical algorithm \( E_\Sigma \) such that for any polynomial-time quantum algorithm \( A \), we have that \( Pr[(s, w) \notin R \land ch \neq ch' \land ok = ok' = 1 : (s, com, ch, resp, ch', resp') \leftarrow A, ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch', resp'), w \leftarrow E_\Sigma(s, com, ch, resp, ch', resp')] \) is negligible.

- **Statistical honest-verifier zero-knowledge (HVZK)**\(^3\) There is a polynomial-time classical algorithm \( S_\Sigma \) (the simulator) such that for any (possibly unlimited) quantum algorithm \( A \) and all \((s, w) \in R\), the following is negligible:

\[
Pr[b = 1 : com \leftarrow P_1(s, w), ch \leftarrow \{0, 1\}^{\ell_{ch}}, resp \leftarrow P_2(ch), b \leftarrow A(com, ch, resp)] - Pr[b = 1 : (com, ch, resp) \leftarrow S(s), b \leftarrow A(com, ch, resp)] = 0
\]

- **Strict soundness**: For any \((s, com, ch)\) and any \( resp \neq resp' \) we have \( Pr[ok = ok' = 1 : ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch, resp')] = 0 \).

- **Computational strict soundness**\(^4\) For any polynomial-time quantum algorithm \( A \), we have that \( Pr[ok = ok' = 1 \land resp \neq resp' : (s, com, ch, resp, resp') \leftarrow A, ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch, resp')] = 0 \).

- **Commitment entropy**: For all \((s, w) \in R \) and \( com \leftarrow P_1(s, w) \), the min-entropy of \( com \) is superlogarithmic.

---

\(^3\)In the context of this paper, HVZK is equivalent to zero-knowledge because our protocols have logarithmic challenge length \( \ell_{ch} \).  
\(^4\)Also known as unique responses in [17].
In a relativized setting, all quantum algorithms additionally get access to all oracles, and all classical algorithms additionally get access to all classical oracles.

In this paper, we will mainly be concerned with proving that certain schemes are not proofs of knowledge. Therefore, we will not need to have precise definitions of these concepts; we only need to know what it means to break them.

**Definition 2 (Total breaks)** Consider an interactive or non-interactive proof system $(P, V)$ for a relation $R$. Let $L_R := \{s : \exists w. (s, w) \in R\}$ be the language defined by $R$. A total break is a polynomial-time quantum algorithm $A$ such that the following probability is overwhelming:

$$\Pr[ok = 1 \land s \notin L_R : s \leftarrow A, ok \leftarrow \langle A, V(s) \rangle]$$

Here $\langle A, V(s) \rangle$ denotes the output of $V$ in an interaction between $A$ and $V(s)$.

A total knowledge break is a polynomial-time quantum algorithm $A$ such that for all polynomial-time quantum algorithms $E$ we have that:

- **Adversary success**: $\Pr[ok = 1 : s \leftarrow A, ok \leftarrow \langle A, V(s) \rangle]$ is overwhelming.
- **Extractor failure**: $\Pr[(x, w) \in R : s \leftarrow A, w \leftarrow E(s)]$ is negligible.

Here $E$ has access to the final state of $A$.

Note that these definitions of attacks are quite strong. In particular, $A$ does not get any auxiliary state. And $A$ needs to succeed with overwhelming probability and make the extraction fail with overwhelming probability. (Usually, proofs / proofs of knowledge are considered broken already when the adversary has non-negligible success probability.) Furthermore, we require $A$ to be polynomial-time.

In particular, a total break implies that a proof system is neither a proof nor an argument. And total knowledge break implies that it is neither a proof of knowledge nor an argument of knowledge, with respect to all definitions the authors are aware of.

### 3 Oracle transformation techniques

In this section, we show two techniques for emulating different oracles. We will need those techniques in our analysis of the pick-one trick (Section 4), but we believe that they are of independent interest and can be useful in other oracles separation results or cryptographic proofs.

#### 3.1 State creation oracles

We first show a result that shows that having access to an oracle $O_\Psi$ for creating copies of an unknown state $|\Psi\rangle$ is not more powerful than having access to a reservoir state $|R\rangle$ of polynomially-many copies of $|\Psi\rangle$ (some of them in superposition with a fixed state $|\perp\rangle$). Note that this is not immediate, because $O_\Psi$ can be queried in superposition, and its inverse applied; this might give more power than the state $|\Psi\rangle$. In fact, we know of no way to generate, e.g., $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|\perp\rangle$ for a given $|\perp\rangle$ and unknown $|\Psi\rangle$, even given many copies of $|\Psi\rangle$ (unless we have enough copies of $|\Psi\rangle$ to determine a complete description of $|\Psi\rangle$ by measuring). Yet $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|\perp\rangle$ can be generated with a single query to $O_\Psi$. This is why our reservoir $|R\rangle$ has to contain such superpositions in addition to pure states $|\Psi\rangle$.

**Theorem 3 (Emulating state creation oracles)** Let $|\Psi\rangle$ be a state, chosen according to some distribution. Let $|\perp\rangle$ be a fixed state orthogonal to $|\Psi\rangle$. (Such a state can always be found

---

5Definitions that would not be covered would be such where the extractor gets additional auxiliary input not available to the adversary. We are, however, not aware of such in the literature.
We do not know, but the state while the tail of
The question is, can we applied
infinite state
defined, the algorithm uses finite tensor products.)

Thus we have a unitary operation
that makes \( q_\Psi \) queries to \( O_\Psi \). Then there is an oracle algorithm
\( B \) that makes the same number of queries to \( O \) as \( A \) such that:
\[
\text{TD}(B^O(|R\rangle, |\Phi\rangle), A^{O_\Phi, O}(|\Phi\rangle)) \leq \frac{\pi q_\Psi}{2\sqrt{n}} + q_\Psi \cdot o\left(\frac{1}{\sqrt{n}}\right) + \frac{2q_\Psi}{\sqrt{m+1}} \leq O\left(\frac{q_\Psi}{\sqrt{n}} + \frac{q_\Psi}{\sqrt{m}}\right).
\]

The idea behind this lemma is the following: To implement \( O_\Psi \), we need a way to convert \(|\perp\rangle\) into \(|\Psi\rangle\) and vice versa. At the first glance this seems easy: If we have a reservoir \( R \) containing \(|\Psi\rangle^{\otimes n}\) for sufficiently large \( n \), we can just take a new \(|\Psi\rangle\) from \( R \). And when we need to destroy \(|\perp\rangle\), we just move it into \( R \). This, however, does not work because the reservoir \( R \) “remembers” whether we added or removed \(|\Psi\rangle\) (because the number of \(|\Psi\rangle\)’s in \( R \) changes). So if we apply \( O_\Psi \) to, e.g., \( \frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|0\rangle \), the reservoir \( R \) essentially acts like a measurement whether we applied \( O_\Psi \) to \(|\Psi\rangle\) or \(|0\rangle\).

To avoid this, we need a reservoir \( R \) in a state that does not change when we add \(|\Psi\rangle\) or \(|\perp\rangle\) to the reservoir. Such a state would be \(|R^\infty\rangle := |\Psi\rangle^{\otimes \infty} \otimes |\perp\rangle^{\otimes \infty}\). If we add or remove \(|\Psi\rangle\) to an infinite state \(|\Psi\rangle^{\otimes \infty}\), that state will not change. Similarly for \(|\perp\rangle\). (The reader may be worried here whether an infinite tensor product is mathematically well-defined or physically meaningful. We do not know, but the state \(|R^\infty\rangle\) is only used for motivational purposes, our final proof only uses finite tensor products.)

Thus we have a unitary operation \( S \) such that \( S|\perp\rangle|R^\infty\rangle = |\Psi\rangle|R^\infty\rangle \). Can we use this operation to realize \( O_\Psi \)? Indeed, an elementary calculation reveals that the following circuit implements \( O_\Psi \) on \( X \) when \( R, Z \) are initialized with \(|R^\infty\rangle, |0\rangle\).

\[
\text{with } U_\perp := 1 - 2|\perp\rangle\langle\perp|, \quad \text{and } O_{\text{Ref}} := 1 - 2|\Psi\rangle\langle\Psi|.
\]

Note that we have introduced a new oracle \( O_{\text{Ref}} \) here. We will deal with that oracle later.

Unfortunately, we cannot use \(|R^\infty\rangle\). Even if such a state should be mathematically well-defined, the algorithm \( B \) cannot perform the infinite shift needed to fit in one more \(|\Psi\rangle\) into \(|R^\infty\rangle\). The question is, can \(|R^\infty\rangle\) be approximated with a finite state? I.e., is there a state \(|R\rangle\) such that \( S|\perp\rangle|R\rangle \approx |\Psi\rangle|R\rangle \) for a suitable \( S \)? Indeed, such a state exists, namely the state \( |R\rangle \) from

\textbf{Lemma 41}. For sufficiently large \( n \), the beginning of \(|R\rangle\) is approximately \(|\Psi\rangle \otimes |\Psi\rangle \otimes |\Psi\rangle \otimes \ldots \), while the tail of \(|R\rangle\) is approximately \( \cdots \otimes |\perp\rangle \otimes |\perp\rangle \otimes |\perp\rangle \otimes |\perp\rangle \). In between, there is a smooth transition. If \( S \) adds \(|\perp\rangle\) to the end and removes \(|\Psi\rangle\) from the beginning of \(|R\rangle\), the state still has approximately the same form (this needs to be made quantitative, of course). That is, \( S \) is a cyclic left-shift on \(|\perp\rangle|R\rangle\).

Hence \(|R\rangle\) is a good approximate drop-in replacement for \(|R^\infty\rangle\), and the circuit (1) approximately realizes \( O_\Psi \) when \( R, Z \) are initialized with \(|R\rangle, |0\rangle\).

However, we now have introduced the oracle \( O_{\text{Ref}} \). We need to show how to emulate that oracle: \( O_{\text{Ref}} \) essentially implements a measurement whether a given state \(|\Psi\rangle\) is \(|\Psi\rangle\) or orthogonal to \(|\Psi\rangle\). Thus to implement \( O_{\text{Ref}} \), we need a way to test whether a given state is \(|\Psi\rangle\) or not.
The well-known swap test [11] is not sufficient, because for |Φ⟩ orthogonal to |Ψ⟩, it gives an incorrect answer with probability $\frac{1}{2}$ and destroys the state. Instead, we use the following test that has an error probability $O(1/m)$ given $m$ copies of |Ψ⟩ as reference: Let |T⟩ := |Ψ⟩⊗m. Let $V$ be the space of all $(m + 1)$-partite states that are invariant under permutations. |Ψ⟩|T⟩ is such a state, while for |Φ⟩ orthogonal to |Ψ⟩, |Φ⟩|T⟩ is almost orthogonal to $V$ for large $m$ (up to an error of $O(1/m)$). So by measuring whether |Φ⟩|T⟩ is in $V$, we can test whether |Φ⟩ is |Ψ⟩ or not (with an error $O(1/m)$), and when doing so the state |T⟩ is only disturbed by $O(1/m)$.

We can thus simulate any algorithm that uses $\mathcal{O}_{\text{Ref}}$ up to any inversely polynomial precision using a sufficiently large state |T⟩.

We then get Theorem 3 by extending the state |R⟩ to also contain |T⟩.

Formally, the theorem is an immediate consequence of Lemmas 41 and 42 in Appendix B.1.

### 3.2 Small image oracles

In this section we show that a random function $H$ (where each $H(x)$ is independently chosen) is indistinguishable from a random function $G$ with a small image. More precisely, a suitably chosen function $G$ with $|\text{im } G| \leq \text{poly}(1/\alpha, q)$ can be distinguished from $H$ in $q$ queries only with probability $\alpha$. (Theorem 4 below.)

In particular, this implies that having access to an oracle $H$ that returns random values $H(x)$ from some distribution is no more powerful than having access to polynomially-many values from that distribution (namely $\text{im } G$). In our setting, we will use this to show that the oracle $\mathcal{O}_S$ which for each $z$ returns a simulated sigma-protocol execution $\mathcal{O}_S(z)$, is not more powerful than having access to polynomially-many such executions. This allows us to remove $\mathcal{O}_S$ from the analysis. (In Corollary 9 below.)

However, we believe Theorem 4 also to be useful in other reduction proofs, too: In [37] a similar result (Lemma 43 below) was used to perform reduction proofs in the random oracle model: In these proofs, a particular challenge value $y$ was inserted into the random oracle at random positions, and with sufficiently high probability, the adversary would then use that value for the attack (which then allows us to conclude the reduction). Our Theorem 4 goes beyond that by allowing to program the random oracle such that every value returned by the random oracle can be a challenge value. (We will not use this feature in the present paper though.)

**Theorem 4 (Small image oracles)** For any $\alpha \in (0, 1)$ and any integer $q \geq 1$, there is an integer $s := \lceil \frac{64q^4}{\alpha^2} \ln \frac{4q^4}{\alpha^2} \rceil \in O(\frac{4q^4}{\alpha^2} (\log \frac{2q^4}{\alpha^2})^2)$ and a distribution $\mathcal{D}_s$ on \{1, ..., s\} such that:

Fix sets $Z, Y$ and a distribution $\mathcal{D}_Y$ on $Y$. Let $H : Z \rightarrow Y$ be chosen as: for each $z \in Z$, $H(z) \leftarrow \mathcal{D}_Y$.

Let $G : Z \rightarrow Y$ be chosen as: Pick $y_1, \ldots, y_s \leftarrow \mathcal{D}_Y$, then for each $z \in Z$, pick $i_z \leftarrow \mathcal{D}_s$, and set $G(z) := y_{i_z}$. Let $A$ be an oracle algorithm making at most $q$ queries.

Then

$$\Delta := \left| \Pr[b = 1 : b \leftarrow A^H] - \Pr[b = 1 : b \leftarrow A^G] \right| \leq \alpha.$$
The running time of \( E \) is given in Appendix B.2.

4 The pick-one trick

In this section, we show first a basic case of the pick-one trick which focusses on the core query complexity aspects. In Section 4.1 we extend this by a number of additional oracles that will be needed in the rest of the paper.

Definition 5 (Two values problem) Let \( X, Y \) be finite sets and let \( k \leq |X| \) be a positive integer. For each \( y \in Y \), let \( S_y \) be a uniformly random subset of \( X \) of cardinality \( k \), let \( |\Psi(y)\rangle := \sum_{x \in S_y} |x\rangle/\sqrt{k} \). Let \( |\Sigma \Psi\rangle = \sum_{y \in Y} |y\rangle|\Psi(y)\rangle/\sqrt{|Y|} \) and \( |\Sigma \Phi\rangle = \sum_{y \in Y, x \in X} |y\rangle|x\rangle/\sqrt{|Y| \cdot |X|} \). The Two Values problem is to find \( y \in Y \) and \( x_1, x_2 \in S_y \) such that \( x_1 \neq x_2 \) given the following resources:

- one instance of the state \( \bigotimes_{\ell=1}^h (\alpha_{\ell,0} |\Sigma \Psi\rangle + \alpha_{\ell,1} |\Sigma \Phi\rangle) \), where \( h \) and the coefficients \( \alpha \) are independent of the \( S_y \)'s and are such that this state has unit norm;
- an oracle \( \mathcal{O}_V \) such that for all \( y \in Y \), \( x \in X \), \( \mathcal{O}_V(y, x) = 0 \) if \( x \notin S_y \) and \( \mathcal{O}_V(y, x) = 1 \) if \( x \in S_y \);
- on oracle \( \mathcal{O}_F \) that, for all \( y \in Y \), maps \( |y, \Psi(y)\rangle \) to \( -|y, \Psi(y)\rangle \) and, for any \( |\Psi^\perp\rangle \) orthogonal to \( |\Psi(y)\rangle \), maps \( |y, \Psi^\perp\rangle \) to itself.

The two values problem is at the core of the pick-one trick: if we give an adversary access to the resources described in Definition 5 he will be able to search for one \( x \in S_y \) satisfying a predicate \( P \) (shown in Theorem 7 below). But doing he will not be able to find two different \( x, x' \in S_y \) (Theorem 6 below); we will use this to foil any attempts at extracting by rewinding.

Theorem 6 (Hardness of the two values problem) Let \( \mathcal{A} \) be an algorithm for the Two Values problem that makes \( q_V \) and \( q_F \) queries to oracles \( \mathcal{O}_V \) and \( \mathcal{O}_F \), respectively. The success probability for \( \mathcal{A} \) to find \( y \in Y \) and \( x_1, x_2 \in S_y \) such that \( x_1 \neq x_2 \) is at most

\[
O \left( \frac{h}{|Y|^{1/2}} + \frac{(q_V + q_F)^{1/2} k^{1/4}}{|X|^{1/4}} + \frac{(q_V + q_F)^{1/2}}{k^{1/4}} \right).
\]

The proof uses the adversary-method from [1, 2] as described in the introduction and is given in Appendices C and D. In Section 4.1 we extend this hardness result to cover additional oracles.

Theorem 7 (Searching one value) Let \( S_y \subseteq X \) and \( \mathcal{O}_F, \mathcal{O}_V \) be as in Definition 5.

There is a polynomial-time oracle algorithm \( E_1 \) that on input \( |\Sigma \Psi\rangle \) returns a uniformly random \( y \in Y \) and \( |\Psi(y)\rangle \). There is a polynomial-time oracle algorithm \( E_2 \) such that: For any \( \delta_{\min} > 0 \), for any \( y \in Y \), for any predicate \( P \) on \( X \) with \( |\{ x \in S_y : P(x) = 1 \}|/|S_y| \geq \delta_{\min} \), and for any \( n \geq 0 \) we have

\[
\Pr[x \in S_y \land P(x) = 1 : x \leftarrow E_2^{\mathcal{O}_V, \mathcal{O}_F, P}(n, \delta_{\min}, y, |\Psi(y)\rangle)] \geq 1 - 2^{-n}.
\]

(The running time of \( E_2 \) is polynomial-time in \( n, 1/\delta_{\min}, |y| \).)

This theorem is proven with a variant of Grover’s algorithm [20]: Using Grover’s algorithm, we search for an \( x \) with \( P(x) = 1 \). However, we do not search over all \( x \in \{0, 1\}^\ell \) for some \( \ell \), but instead over all \( x \in S_y \). When searching over \( S_y \), the initial state of Grover’s algorithm needs to be \( \sum_{x \in S_y} |x\rangle = |\Psi(y)\rangle \) instead of \( \sum_{x \in \{0, 1\}^\ell} 2^{-\ell/2} |x\rangle = |\Phi\rangle \). And the diffusion operator \( I - 2|\Phi\rangle\langle\Phi| \) needs to be replaced by \( I - 2|\Psi(y)\rangle\langle\Psi(y)| \). Fortunately, we have access both to \( |\Psi(y)\rangle \) (given as input), and to \( I - 2|\Psi(y)\rangle\langle\Psi(y)| \) (through the oracle \( \mathcal{O}_F \)). To get an overwhelming success
probability, Grover’s algorithm is usually repeated until it succeeds. (In particular, when the number of solutions is not precisely known \cite{9}.) We cannot do that: we have only one copy of the initial state. Fortunately, by being more careful in how we measure the final result, we can make sure that the final state is in case of failure is also a suitable initial state for Grover’s algorithm. (Note that the necessity of repeating also occurs when the number of solutions is known precisely: since the number of iterations in Grover’s algorithm needs to be an integer, it will rarely be close enough to the optimal value.)

The full proof is given in Section E.1.

### 4.1 Additional oracles

In this section, we extend the hardness of the two values problem to cover additional oracles that we will need in various parts of the paper.

**Definition 8 (Oracle distribution)** Fix integers $\ell_{\text{com}}, \ell_{\text{ch}}, \ell_{\text{resp}}$ (that may depend on the security parameter) such that $\ell_{\text{com}}, \ell_{\text{resp}}$ are superlogarithmic and $\ell_{\text{ch}}$ is logarithmic. Let $\ell_{\text{rand}} := \ell_{\text{com}} + \ell_{\text{resp}}$.

Let $O_{\text{all}} = (O_E, O_P, O_R, O_S, O_F, O_Q, O_V)$ be chosen according to the following distribution:

- Let $s_0 := 0$ (fixed). Pick $w_0 \leftarrow \{0,1\}^{\ell_{\text{rand}}}$. 
- Choose $S_E, O_V, O_F$ as in Definition 7 with $Y := \{0,1\}^{\ell_{\text{com}}}$ and $X := \{0,1\}^{\ell_{\text{ch}}} \times \{0,1\}^{\ell_{\text{resp}}}$ and $k := 2^{\ell_{\text{ch}} + \ell_{\text{resp}} / 3}$.
- For each $z \in \{0,1\}^{\ell_{\text{rand}}}$, pick $y \leftarrow Y$ and $x \leftarrow S_y$, and set $S(z) := (y, x)$.
- Let $|\bot\rangle$ be a quantum state orthogonal to all $|\text{com}, \text{ch}, \text{resp}\rangle$ (i.e., we extend the dimension of the space in which $|\Sigma\Psi\rangle$ lives by one). $O_Q|\bot\rangle := |\Sigma\Psi\rangle$, $O_Q|\Sigma\Psi\rangle := |\bot\rangle$, and $O_Q|\Phi\rangle := |\Phi\rangle$ for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$ and $|\bot\rangle$.
- Let $O_E|\text{com}, \text{ch}, \text{resp}, \text{ch}', \text{resp}'\rangle := w_0$ iff $(\text{ch, resp}), (\text{ch}', \text{resp}') \in S_{\text{com}} \land (\text{ch, resp}) \neq (\text{ch}', \text{resp}')$ and $O_E := 0$ everywhere else.
- Let $O_R(s_0, w_0) := 1$ and $O_R := 0$ everywhere else.
- For each $\text{com} \in \{0,1\}^{\ell_{\text{com}}}, \text{ch} \in \{0,1\}^{\ell_{\text{ch}}}, z \in \{0,1\}^{\ell_{\text{rand}}}$, let $O_P(w_0, \text{com}, \text{ch}, z)$ be assigned a uniformly random resp with $(\text{ch, resp}) \in S_{\text{com}}$. (Or $\bot$ if no such resp exists.) Let $O_P(w_1, \cdot, \cdot, \cdot) := 0$ for $w \neq w_0$.

The following corollary is a strengthening of Theorem 6 to the oracle distribution from Definition 8. For later convenience, we express the soundness additionally in terms of guessing $w_0$. Since the formula would become unwieldy, we do not give a concrete asymptotic bound here. But such a bound can be easily derived from the inequalities \cite{19} \cite{20} in the proof.

**Corollary 9 (Hardness of two values 2)** Let $O_{\text{all}} = (O_E, O_P, O_R, O_S, O_F, O_Q, O_V)$, $w_0$ be as in Definition 8. Let $A$ be an oracle algorithm making at most $q_E, q_P, q_R, q_S, q_F, q_Q, q_V$ queries to $O_E, O_P, O_R, O_S, O_F, O_Q, O_V$, respectively. Assume that $q_E, q_P, q_R, q_S, q_F, q_V$ are polynomially-bounded (and $\ell_{\text{com}}, \ell_{\text{resp}}$ are superlogarithmic by Definition 8). Then:

(i) $\Pr[|w = w_0 : w \leftarrow A|] = \text{negligible}$

(ii) $\Pr[|\text{ch, resp} \neq (\text{ch}', \text{resp}') \land (\text{ch, resp}), (\text{ch}', \text{resp}') \in S_{\text{com}} : (\text{com, ch, resp}), (\text{ch}', \text{resp}') \leftarrow A|] = \text{negligible}$.

This corollary is shown by reduction to Theorem 6 (Hardness of the two values problem). Given an adversary that violates 1, we remove step by step the oracles that are not covered by Theorem 6. First, we remove the oracles $O_P, O_R$. Those do not help the adversary (much) to find $w_0$ because $O_P$ and $O_R$ only give non-zero output if their input already contains $w_0$. Next we change $A$ to output a collision $(\text{ch, resp}) \neq (\text{ch}', \text{resp}') \land (\text{ch, resp}), (\text{ch}', \text{resp}') \in S_{\text{com}}$ instead of the witness $w_0$; since $w_0$ can only be found by querying $O_E$ with such a collision, this adversary succeeds with non-negligible probability, too. Furthermore, $A$ then does not need access to $O_E$ any more since $O_E$ only helps in finding $w_0$. Next we get rid of $O_Q$: as shown
in Theorem 3 (Emulating state creation oracles), $O_\Psi$ can be emulated (up to an inversely polynomial error) using (suitable superpositions on) copies of the state $|\Sigma\Psi\rangle$. Finally we remove $O_S$: By Theorem 4 (Small image oracles), $O_S$ can be replaced by an oracle that provides only a polynomial number of triples $(com, ch, resp)$. Those triples the adversary can produce himself by measuring polynomially-many copies of $|\Sigma\Psi\rangle$ in the computational basis. Thus we have shown that without loss of generality, we can assume an adversary that only uses the oracles $O_F, O_V$ and (suitable superpositions of) polynomially-many copies of $|\Sigma\Psi\rangle$, and that tries to find a collision. But that such an adversary cannot find a collision was shown in Theorem 6.

And (ii) is shown by observing that an adversary violating (i) leads to one violating (ii) using one extra $O_E$-query. The full proof is given in Section E.2.

5 Attacking commitments

In the classical setting, a non-interactive commitment scheme is usually called computationally binding if it is hard to output a commitment and two different openings (Definition 10 below). We now show that in the quantum setting, this definition is extremely weak. Namely, it may still be possible to commit to a value and then to open the commitment to an arbitrary value (just not to two values at the same time).

Security definitions. To state this more formally, we define the security of commitments: A non-interactive commitment scheme consists of algorithms $COM, COM_{\text{verify}}$, such that $(c,u) \leftarrow COM(m)$ returns a commitment $c$ on the message $m$, and an opening information $u$. We require perfect completeness, i.e., for any $m$ and $(c,u) \leftarrow COM(m)$, $COM_{\text{verify}}(c,m,u) = 1$ with probability 1. In our setting, $c,m,u$ are all classical.

Definition 10 (Computationally binding) A commitment scheme $COM, COM_{\text{verify}}$ is computationally binding iff for any quantum polynomial-time algorithm $A$ the following probability is negligible:

$$\Pr[ok = ok' = 1 \land m \neq m' : (c,m,u,m',u') \leftarrow A,$nok \leftarrow COM_{\text{verify}}(c,m,u), \quad ok \leftarrow COM_{\text{verify}}(c,m',u')]$$

We will show below that this definition is not the right one in the quantum setting.

[29] also introduces a stronger variant of the binding property, called strict binding, which requires that also the opening information $u$ is unique (not only the message). We define a computational variant of this property here:

Definition 11 (Computationally strict binding) A commitment scheme $COM, COM_{\text{verify}}$ is computationally strict binding iff for any quantum polynomial-time algorithm $A$ the following probability is negligible:

$$\Pr[ok = ok' = 1 \land (m,u) \neq (m',u') : (c,m,u,m',u') \leftarrow A,$nok \leftarrow COM_{\text{verify}}(c,m,u), \quad ok \leftarrow COM_{\text{verify}}(c,m',u')]$$

We will show below that this stronger definition is also not sufficient.

Definition 12 (Statistically hiding) A commitment scheme $COM, COM_{\text{verify}}$ is statistically hiding iff for all $m_1,m_2$ with $|m_1| = |m_2|$ and $c_i \leftarrow COM(m_i)$ for $i = 1,2$, $c_1$ and $c_2$ are statistically indistinguishable.
The attack. We now state the insecurity of computationally binding commitments. The remainder of this section will prove the following theorem.

**Theorem 13 (Insecurity of binding commitments)** There is an oracle $O$ and a noninteractive commitment scheme $\text{COM,COM}_\text{verify}$ such that:

- The scheme is perfectly complete, computationally binding, computationally strict binding, and statistically hiding.
- There is a quantum polynomial-time adversary $B_1,B_2$ such that for all $m$,
  \[ \Pr[\text{ok} = 1 : c \leftarrow B_1(|m|), u \leftarrow B_2(m), \text{ok} \leftarrow \text{COM}_\text{verify}(c,m,u)] \]
  is overwhelming. (In other words, the adversary can open to a value $m$ that he did not know while committing.)

In the rest of this section, when referring to the sets $S_{\text{com}}$ from **Definition 8**, we will call them $S_y$ and we refer to their members as $x \in S_y$. (Not $(ch,\text{resp}) \in S_{\text{com}}$.) In particular, oracles such as $O_S$ will returns pairs $(y,x)$, not triples $(com, ch, \text{resp})$, etc.

We construct a commitment scheme relative to the oracle $O_{\text{all}}$ from **Definition 8**. (Note: that oracle distribution contains more oracles than we need for Theorem 13. However, we will need in later sections that our commitment scheme is defined relative to the same oracles as the proof systems there.)

**Definition 14 (Bad commitment scheme)** We define $\text{COM,COM}_\text{verify}$ as follows:

- $\text{COM}(m)$: For $i = 1,\ldots,|m|$, pick $z_i \leftarrow \{0,1\}^{\ell_{\text{rand}}}$ and let $(y_i,x_i) := O_S(z_i)$. Let $p_i \leftarrow \{1,\ldots,\ell_{ch} + \ell_{\text{resp}}\}$. Let $b_i := m_i \oplus \text{bit}_{p_i}(x_i)$. Let $c := (p_1,\ldots,p_{|m|},y_1,\ldots,y_{|m|},b_1,\ldots,b_{|m|})$ and $u := (x_1,\ldots,x_{|m|})$. Output $(c,u)$.
- $\text{COM}_\text{verify}(c,m,u)$ with $c = (p_1,\ldots,p_n,y_1,\ldots,y_n,b_1,\ldots,b_n)$ and $u = (x_1,\ldots,x_n)$: Check whether $|m| = n$. Check whether $O_V(y_i,x_i) = 1$ for $i = 1,\ldots,n$. Check whether $b_i = m_i \oplus \text{bit}_{p_i}(x_i)$ for $i = 1,\ldots,n$. Return 1 if all checks succeed.

For the results of the current section, there is actually no need for the values $p_i$ which select which bit of $x_i$ is used for masking the committed bit $m_i$. (E.g., we could always use the least significant bit of $x_i$.) But in Section 8 (attack on Fischlin’s scheme) we will need commitments of this particular form to enable a specific attack where we need to open commitments to certain values while simultaneously searching for these values in the first place.

**Lemma 15 (Properties of COM)** The scheme from **Definition 14** is perfectly complete, computationally binding, computationally strict binding, and statistically hiding. (Relative to $O_{\text{all}}$.)

The computational binding and computational strict binding property are a consequence of **Corollary 9 (Hardness of two values)**: to open a commitment to two different values, the adversary would need to find one $y_i$ (part of the commitment) and two $x_i \in S_{y_i}$ (part of the two openings). **Corollary 9** states that this only happens with negligible probability. Statistical hiding follows from the fact that for each $y_i$, there are superpolynomially many $x_i \in S_{y_i}$, hence $\text{bit}_{p_i}(x_i)$ is almost independent of $y_i$.

The proof is given in Section F.1.

**Lemma 16 (Attack on COM)** There is a quantum polynomial-time adversary $B_1,B_2$ such that for all $m$,

\[ \varepsilon_{\text{COM}} := \Pr[\text{ok} = 1 : c \leftarrow B_1(|m|), u \leftarrow B_2(m), \text{ok} \leftarrow \text{COM}_\text{verify}(c,m,u)] \]

is overwhelming.
6 Attacking sigma-protocols

We will now show that in general, sigma-protocols with special soundness are not necessarily proofs of knowledge. [29] showed that if a sigma-protocol additionally has strict soundness, it is a proof of knowledge. It was left as an open problem whether that additional condition is necessary. The following theorem resolves that open question by showing that the results from [29] do not hold without strict soundness (not even with computational strict soundness), relative to an oracle.

Theorem 17 (Insecurity of sigma-protocols) There is an oracle $O_{all}$ and a relation $R$ and a sigma-protocol relative to $O_{all}$ with logarithmic $\ell_{ch}$ (challenge length), completeness, perfect special soundness, computational strict soundness, and statistical honest-verifier zero-knowledge for which there exists a total knowledge break.

In contrast, a sigma-protocol relative to $O_{all}$ with completeness, perfect special soundness, and statistical honest-verifier zero-knowledge is a classical proof of knowledge.

Note that a corresponding theorem with polynomially bounded $\ell_{ch}$ follows immediately by parallel repetition of the sigma-protocol.

The remainder of this section will prove Theorem 17. As a first step, we construct the sigma-protocol.

Definition 18 (Sigma-protocol) Let $COM, COM_{\text{verify}}$ be the commitment scheme from Definition 14.

Relative to the oracle distribution from Definition 8, we define the following sigma-protocol $(\ell_{\text{com}}, \ell_{ch}, \ell_{\text{resp}}, P_{1}, P_{2}, V, R)$ for the relation $R := \{(s_{0}, w_{0})\}$:

- $P_{1}(s, w)$ picks $\text{com} \sim \{0, 1\}^{\ell_{\text{com}}}$. For each $ch \in \{0, 1\}^{\ell_{ch}}$, he picks $z_{ch} \sim \{0, 1\}^{\ell_{\text{rand}}}$ and computes $\text{resp}_{ch} := O_{P}(w, \text{com}, ch, z_{ch})$ and $(c_{ch}, u_{ch}) \leftarrow COM(\text{resp}_{ch})$. Then $P_{1}$ outputs $\text{com}^{\ast} := (\text{com}, (c_{ch})_{ch \in \{0, 1\}^{\ell_{ch}}})$.
- $P_{2}(ch)$ outputs $\text{resp}^{\ast} := (\text{resp}_{ch}, u_{ch})$.
- For $\text{com}^{\ast} = (\text{com}, (c_{ch})_{ch \in \{0, 1\}^{\ell_{ch}}})$ and $\text{resp}^{\ast} = (\text{resp}, u)$, let $V(s, \text{com}^{\ast}, ch, \text{resp}^{\ast}) := 1$ iff $O_{V}(\text{com}, ch, \text{resp}) = 1$ and $s = s_{0}$ and $COM_{\text{verify}}(c_{ch}, \text{resp}, u) = 1$.

The commitments $c_{ch}$ are only needed to get computational strict soundness. A slightly weaker Theorem 17 without computational strict soundness can be achieved using the sigma-protocol from Definition 18 without the commitments $c_{ch}$; the proofs stay the same, except that the steps relating to the commitments are omitted.

Lemma 19 (Security of the sigma-protocol) The sigma-protocol from Definition 18 has: completeness, perfect special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, commitment entropy.

The full proof is given in Appendix F.2.
Perfect special soundness follows from the existence of the oracle $O_E$. That oracle provides the witness $w_0$ given two accepting conversations, as required by perfect special soundness. Computational strict soundness stems from the fact that the message $com^*$ contains commitments $c_{ch}$ to all possible answers. Thus to break computational strict soundness (i.e., to find two different accepting $resp^*$), the adversary would need to open one of the commitments $c_{ch}$ in two ways. This happens with negligible probability since COM is computationally strict binding. Statistical honest-verifier zero-knowledge follows from the existence of the oracle $O_S$ which provides simulations. (And the commitment $c_{ch}$ that are not opened can be filled with arbitrary values due to the statistical hiding property of COM.)

The full proof is given in Appendix G.1.

**Lemma 20 (Attack on the sigma-protocol)** Assume that $\ell_{ch}$ is logarithmically bounded. Then there exists a total knowledge break (Definition 2) against the sigma-protocol from Definition 18.

To attack the sigma protocol, the malicious prover uses Theorem 7 (Searching one value) to get a $com$ and a corresponding state $|Ψ(com)\rangle$. Then, when receiving $ch$, he needs to find $(ch', resp) \in S_{com}$ with $ch' = ch$. Since an inversely polynomial fraction of $(ch', resp)$ satisfy $ch' = ch$ ($\ell_{ch}$ is logarithmic), this can be done with Theorem 7. This allows the prover to succeed in the proof with overwhelming probability. (He additionally needs to open the commitments $c_{ch}$ to suitably. This can be done using Lemma 16 (Attack on COM).) However, an extractor that has the same information as the prover (namely, access to the oracle $O_{all}$) will fail to find $w_0$ by Corollary 9 (Hardness of two values).

The full proof is given in Appendix G.2.

Note that we cannot expect to get a total break (as opposed to a total knowledge break): Since the sigma-protocol is a classical proof of knowledge, it is also a classical proof. But a classical proof is also a quantum proof, because an unlimited classical adversary can simulate a quantum adversary. However, this argument does not apply when we consider computationally limited provers, see Section 6.1 below.

### 6.1 The computational case

We now consider the variant of the impossibility result from the previous section. Namely, we consider sigma-protocols that have only computational security (more precisely, for which the special soundness property holds only computationally) and show that these are not even arguments in general (the results from the previous section only say that they are not arguments of knowledge).

**Theorem 21 (Insecurity of sigma-protocols, computational)** There is an oracle $O_{all}$ and a relation $R'$ and a sigma-protocol relative to $O_{all}$ with logarithmic $\ell_{ch}$ (challenge length), completeness, computational special soundness, and statistical honest-verifier zero-knowledge for which there exists a total break.

In contrast, a sigma-protocol relative to $O_{all}$ with completeness, computational special soundness, and statistical honest-verifier zero-knowledge is a classical argument.

Note that a corresponding theorem with polynomially bounded $\ell_{ch}$ follows immediately by parallel repetition of the sigma-protocol. The remainder of this section is dedicated to proving Theorem 21.

**Definition 22 (Sigma-protocol, computational)** We define a sigma-protocol $(\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R')$ as in Definition 18 except that the relation is $R' := \emptyset$. 

Lemma 23 (Security of the sigma-protocol, computational) The sigma-protocol from Definition 22 has: completeness. computational special soundness. computational strict soundness. statistical honest-verifier zero-knowledge. commitment entropy.

Most properties are either immediate or shown as in Lemma 19 (Security of the sigma-protocol). However, perfect special soundness does not hold for the sigma-protocol from Definition 22: There exist pairs of accepting conversations \((ch, resp)\), \((ch', resp')\) ∈ \(S_{com}\). But these do not allow us to extract a valid witness for \(s_0\) (because \(R' = \emptyset\), so no witnesses exist). However, we have computational special soundness: by Corollary 9 (Hardness of two values 2), it is computationally infeasible to find those pairs of conversations.

The full proof is given in Appendix G.3.

Lemma 24 (Attack on the sigma-protocol, computational) Assume that \(l_{ch}\) is logarithmically bounded. Then there exists a total break (Definition 2) against the sigma-protocol from Definition 22.

In this lemma, we use the same malicious prover as in Lemma 20 (Attack on the sigma-protocol). That adversary proves the statement \(s_0\). Since \(R' = \emptyset\), that statement is not in the language, thus this prover performs a total break.

The full proof is given in Appendix G.4.

Now Theorem 21 follows from Lemmas 23 and 24. (And sigma-protocols with computational special soundness are arguments of knowledge and thus arguments; we are not aware of an explicit write-up in the literature, but the proof from [13] for sigma-protocols with special soundness applies to this case, too.)

7 Attacking Fiat-Shamir

Definition 25 (Fiat-Shamir) Fix a sigma-protocol \((l_{com}, l_{ch}, l_{resp}, P_1, P_2, V, R)\) and an integer \(r > 0\). Let \(H : \{0,1\}^* \rightarrow \{0,1\}^{r \cdot l_{ch}}\) be a random oracle. The Fiat-Shamir construction \((P_{FS}, V_{FS})\) is the following non-interactive proof system:

- **Prover** \(P_{FS}(s,w)\): For \((s,w) \in R\), invoke \(com_i \leftarrow P_1(s,w)\) for \(i = 1, \ldots, r\). Let \(ch_1\|\ldots\|ch_r := H(s, com_1, \ldots, com_r)\). Invoke \(resp_i \leftarrow P_2(ch_i)\). Return \(\pi := (com_1, \ldots, com_r, resp_1, \ldots, resp_r)\).
- **Verifier** \(V_{FS}(s, (com_1, \ldots, com_r, resp_1, \ldots, resp_r))\): Let \(ch_1\|\ldots\|ch_r := H(s, com_1, \ldots, com_r)\). Check whether \(V(s, com_i, ch_i, resp_i) = 1\) for all \(i = 1, \ldots, r\). If so, return 1.

Theorem 26 (Insecurity of Fiat-Shamir) There is an oracle \(O_{all}\) and a relation \(R\) and a sigma-protocol relative to \(O_{all}\) with logarithmic \(l_{ch}\) (challenge length), completeness, perfect special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, and commitment entropy, such that there is total knowledge break on the Fiat-Shamir construction.

In contrast, the Fiat-Shamir construction based on a sigma-protocol with the same properties is a classical argument of knowledge (assuming that \(r l_{ch}\) is superlogarithmic).

As the underlying sigma-protocol, we use the one from Definition 18. The attack on Fiat-Shamir is analogous to that on the sigma-protocol itself. The only difference is that the challenge \(ch\) now comes from \(H\) and not from the verifier; this does not change the attack strategy.

The full proof is given in Appendix H.1.
7.1 The computational case

Again, we get even stronger attacks if the special soundness holds only computationally.

**Theorem 27 (Insecurity of Fiat-Shamir, computational)** There is an oracle \( O_{all} \) and a relation \( R \) and a sigma-protocol relative to \( O_{all} \) with logarithmic \( \ell_{ch} \) (challenge length), completeness, computational special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, and commitment entropy, such that there is a total break on the Fiat-Shamir construction.

In contrast, the Fiat-Shamir construction based on a sigma-protocol with the same properties is a classical argument of knowledge (assuming that \( r\ell_{ch} \) is superlogarithmic).

The proof is along the lines of those of Theorem 26 and Lemma 24 and given in Appendix \[1.2\].

8 Attacking Fischlin’s scheme

In the preceding sections we have used the pick-one trick to give negative results for the (knowledge) soundness of sigma protocols and of the Fiat-Shamir construction. Classically, both protocols are shown sound using rewinding. This leads to the conjecture that the pick-one trick is mainly useful for getting impossibilities for protocols with rewinding-based security proofs. Yet, in this section we show that this is not the case; we use the pick-one trick to give an impossibility result for Fischlin’s proof system with online-extractors \[17\]. The crucial point of that construction is that in the classical security proof, no rewinding is necessary. Instead, a witness is extracted by passively inspecting the list of queries performed by the adversary.

**Definition 28 (Fischlin’s scheme)** Fix a sigma-protocol \((\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R)\). Fix integers \( b, r, S, t \) such that \( br \) and \( 2^{\ell_{ch} - b} \) are superlogarithmic, \( b, r, t \) are logarithmic, \( S \in O(r) \) (\( S = 0 \) is permitted), and \( b \leq t \leq \ell_{ch} \).

Let \( H : \{0,1\}^* \to \{0,1\}^b \) be a random oracle. Fischlin’s construction \((P_{Fis}, V_{Fis})\) is the non-interactive proof system defined as follows:

- \( P_{Fis}(s, w) \): See \[17\]. (Omitted here since we only need to analyze \( V_{Fis} \) for our results.)
- \( V_{Fis}(s, \pi) \) with \( \pi = (com_i, ch_i, resp_i)_{i=1,...,r} \): Check if \( V(com_i, ch_i, resp_i) = 0 \) for all \( i = 1, \ldots, r \). Check if \( \sum_{i=1}^{r} H(x, (com_i), i, ch_i, resp_i) \leq S \) (where \( H(\ldots) \) is interpreted as a binary unsigned integer). If all checks succeed, return 1.

The idea (in the classical case) is that, in order to produce triples \((com_i, ch_i, resp_i)\) that make \( H(x, (com_i), i, ch_i, resp_i) \) sufficiently small, the prover needs try out several accepting \( ch_i, resp_i \) for each \( com_i \). So with overwhelming probability, the queries made to \( H \) will contain at least two \( ch_i, resp_i \) for the same \( com_i \). This then allows extraction by just inspecting the queries.

In the quantum setting, this approach towards extraction does not work: the “list of random oracle queries” is not a well-defined notion, because the argument of \( H \) is not measured when a query is performed. In fact, we show that Fischlin’s scheme is in fact not an argument of knowledge in the quantum setting (relative to an oracle):

**Theorem 29 (Insecurity of Fischlin’s construction)** There is an oracle \( O_{all} \) and a relation \( R \) and a sigma-protocol relative to \( O_{all} \) with logarithmic \( \ell_{ch} \) (challenge length), completeness, perfect special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, and commitment entropy, such that there is a total knowledge break of Fischlin’s construction.

In contrast, Fischlin’s construction based on a sigma-protocol with the same properties is a classical argument of knowledge.
As the underlying sigma-protocol, we use the one from Definition 18. The basic idea is that the malicious prover finds conversations \((\text{com}_i^*, \text{ch}_i, \text{resp}_i^*)\) by first fixing the values \(\text{com}_i^*\), and then using Theorem 7 to find \(\text{ch}, \text{resp}_i^*\) where \(\text{resp}_i^*\) contains \(\text{resp}_i\) such that \((\text{ch}_i, \text{resp}_i^*) \in S_{\text{com}_i}\) and \(H(x, (\text{com}_i^*)_i, i, \text{ch}_i, \text{resp}_i^*) = 0\). If \(\text{resp}_i^*\) would not additionally contain commitments \(c_{\text{ch}}\) (see Definition 18), this would already suffice to break Fischl’s scheme. To additionally make sure we can open the commitments to the right value, we use a specific fixpoint property of COM. See the full proof (Appendix I.1) for details.

8.1 The computational case

Theorem 30 (Insecurity of Fischlin’s construction, computational) There is an oracle \(O_{\text{all}}\) and a relation \(R\) and a sigma-protocol relative to \(O_{\text{all}}\) with logarithmic \(\ell_{\text{ch}}\) (challenge length), completeness, computational special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, and commitment entropy, such that there is a total break on Fischlin’s construction.

In contrast, Fischlin’s construction based on a sigma-protocol with the same properties is a classical argument of knowledge.

The proof is given in Appendix I.2.

Fischlin’s scheme with strict soundness. We conjecture that Theorems 29 and 30 even hold with strict soundness instead of computational strict soundness. We sketch our reasoning: Consider a variant of the oracle distribution from Definition 8 in which \(\ell_{\text{ch}}\) is superlogarithmic (not logarithmic) and in which the sets \(S_{\text{com}}\) are chosen uniformly at random from all sets \(S\) which satisfy \(\forall \text{ch} \exists_1 \text{resp}. (\text{ch}, \text{resp}) \in S\). Note that the results from Sections 5–7 do not hold in this setting, because \(\text{ch}\) must be polynomially-bounded to show the existence of successful adversaries. (Namely, when Theorem 7 (Searching one value) is invoked, the predicate \(P\) is true on a \(2^{-\ell_{\text{ch}}}\) fraction of the all values.) But the proofs of Lemma 51 (Attack on Fischlin’s construction) and Lemma 52 (Attack on Fischlin’s construction, computational) do not require this. We conjecture that Corollary 9 still holds in this modified setting (the cardinality of the \(S_{\text{com}}\) satisfies the conditions of Corollary 9 but the \(S_{\text{com}}\) have additional structure). Then the sigma-protocols from Definitions 18 and 22 (without the commitments \(c_{\text{ch}}\)) will still have the properties shown in Lemmas 19 and 23, but additionally they will have strict soundness because for any \(\text{com}, \text{ch}\), there exists only one \(\text{resp}\) such that \((\text{ch}, \text{resp}) \in S_{\text{com}}\).

We leave the proof that Corollary 9 holds even for sets \(S_{\text{com}}\) with \(\forall \text{ch} \exists_1 \text{resp}. (\text{ch}, \text{resp}) \in S_{\text{com}}\) as an open problem.

Acknowledgments. We thank Marc Fischlin and Tommaso Gagliardoni for valuable discussions and the initial motivation for this work. Andris Ambainis was supported by FP7 FET projects QCS and QALGO and ERC Advanced Grant MQC (at the University of Latvia) and by National Science Foundation under agreement No. DMS-1128155 (at IAS, Princeton). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Ansis Rosmanis was supported by the Mike and Ophelia Lazaridis Fellowship, the David R. Cheriton Graduate Scholarship, and the US ARO. Dominique Unruh was supported by the Estonian ICT program 2011-2015 (3.2.1201.13-0022), the European Union through the European Regional Development Fund through the sub-measure “Supporting the development of R&D of info and communication technology”, by the European Social Fund’s Doctoral Studies and Internationalisation Programme DoRa, by the Estonian Centre of Excellence in Computer Science, EXCS.
References


[38] Mark Zhandry. Personal communication, 2014.

**Symbol index**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>Parameter of Fischlin’s scheme: number of subproofs</td>
</tr>
<tr>
<td>$\text{COM}_{\text{verify}}$</td>
<td>Verification algorithm for COM</td>
</tr>
<tr>
<td>$O_S$</td>
<td>Oracle, enabling simulation</td>
</tr>
<tr>
<td>$k$</td>
<td>Cardinality of random sets $S_y$ (or $S_{\text{com}}$)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Security parameter</td>
</tr>
<tr>
<td>$\text{COMopen}^*$</td>
<td>A specific cheating open phase for attacking Fischlin’s scheme</td>
</tr>
<tr>
<td>$P_{FS}$</td>
<td>Prover of Fiat-Shamir</td>
</tr>
<tr>
<td>$V_{\text{Fis}}$</td>
<td>Verifier of Fischlin’s construction</td>
</tr>
<tr>
<td>$\text{COM}$</td>
<td>Commitment scheme from Definition 14</td>
</tr>
<tr>
<td>$E_\Sigma$</td>
<td>Special soundness extractor for sigma protocol $\Sigma$</td>
</tr>
<tr>
<td>$S_\Sigma$</td>
<td>Honest-verifier simulator extractor for sigma protocol $\Sigma$</td>
</tr>
<tr>
<td>$</td>
<td>\Psi\rangle$</td>
</tr>
<tr>
<td>$\langle\Psi</td>
<td>$</td>
</tr>
<tr>
<td>$\ell_{\text{com}}$</td>
<td>Length of commitments in sigma-protocol</td>
</tr>
<tr>
<td>$\ell_{\text{ch}}$</td>
<td>Length of challenges in sigma-protocol</td>
</tr>
<tr>
<td>$\ell_{\text{resp}}$</td>
<td>Length of responses in sigma-protocol</td>
</tr>
<tr>
<td>$\ell_{\text{rand}}$</td>
<td>Length of randomness in oracle queries</td>
</tr>
<tr>
<td>$\mathcal{D}$</td>
<td>Denotes a distribution</td>
</tr>
<tr>
<td>$</td>
<td>\perp\rangle$</td>
</tr>
<tr>
<td>$O_V$</td>
<td>Oracle, enabling verification</td>
</tr>
<tr>
<td>$O_{\text{Ref}}$</td>
<td>Oracle, measures $</td>
</tr>
<tr>
<td>$</td>
<td>x</td>
</tr>
<tr>
<td>$</td>
<td>x</td>
</tr>
<tr>
<td>$O_P$</td>
<td>Oracle, enabling honest proofs</td>
</tr>
<tr>
<td>$O_R$</td>
<td>Oracle, membership in relation $R$</td>
</tr>
<tr>
<td>$\mathcal{H}$</td>
<td>Denotes a Hilbert space</td>
</tr>
<tr>
<td>$\text{Var}[X]$</td>
<td>Variance of $X$</td>
</tr>
<tr>
<td>$\mathbb{R}$</td>
<td>Real numbers</td>
</tr>
<tr>
<td>$\text{bit}_i(x)$</td>
<td>$i$-th bit of bitstring $x$ (from left)</td>
</tr>
<tr>
<td>$\text{poly}(n)$</td>
<td>Polynomials-bounded in $n$</td>
</tr>
<tr>
<td>$H_\alpha(D)$</td>
<td>Renyi entropy of order $\alpha$ of distribution $D$</td>
</tr>
<tr>
<td>$\text{im } f$</td>
<td>Image of function $f$</td>
</tr>
<tr>
<td>$\text{SD}(A;B)$</td>
<td>Statistical distance between random variables or distributions $A$ and $B$</td>
</tr>
</tbody>
</table>
E[A] Expected value of random variable A

$P_{Fss}$ Prover of Fischlin's construction

$O_E$ Oracle, enabling extraction

$O_F$ Oracle, mapping $|\Psi\rangle \rightarrow -|\Psi\rangle$

COM* A specific cheating commit phase for attacking Fischlin’s scheme

$t$ Parameter of Fischlin’s scheme: number of tries performed by prover

$|x\rangle$ $x$ rounded towards $-\infty$

$\text{TD}(\rho, \rho')$ Trace distance between $\rho, \rho'$. Short $\text{TD}(|\Psi\rangle, |\Psi'\rangle)$ for $\text{TD}(|\Psi\rangle\langle\Psi|, |\Psi'\rangle\langle\Psi'|)$

$|\Psi(y)\rangle$ Superposition of all $x \in S_y$, for pick-one trick

$S_y$ Set of all “good” $x$, in pick-one trick

$x \leftarrow A$ $x$ is assigned output of algorithm $A$

$|\Sigma\Psi\rangle$ Superposition of all $|\Psi(y)\rangle$, for pick-one trick

$\text{ch}$ Challenge (second message in sigma-protocol, by verifier)

$x \leftarrow S$ $x$ chosen uniformly from set $S$ according to distribution $S$

resp Response (third message in sigma-protocol, by prover)

com Commitment (first message in sigma-protocol, by prover)

$|\text{yes}\rangle$ Superposition of no-instances in Grover search

$|\text{no}\rangle$ Superposition of no-instances in Grover search

$O_{\Psi}$ Oracle that provides $|\Psi\rangle$

$O_{\text{all}}$ The oracles $O_E, O_F, O_R, O_S, O_{\Psi}, O_V$ together

$O$ Denotes an oracle

$ok \leftarrow \langle P, V \rangle$ Joint execution of $P$ and $V$, $ok$ is $V$’s output

$L_R$ Language defined by $R$

$b$ Parameter of Fischlin’s scheme: length of $H$-outputs

$\|x\|$ Euclidean norm of $x$

$S$ Parameter of Fischlin’s scheme: maximum sum of $H$-outputs

$V_{FS}$ Verifier of Fiat-Shamir

---

**Keyword index**

- adversary success, 8
- binding
  - computationally binding, 13
  - computationally strict binding, 13
- break total, 8
- challenge, 4
  - (in sigma-protocol), 7
- commitment, 4
  - (in sigma-protocol), 7
- commitment entropy, 7
- commitment scheme
  - computationally binding, 13
  - computationally strict binding, 13
  - non-interactive, 13
  - perfect completeness, 13
  - statistically hiding, 13
- completeness
  - (of sigma-protocol), 7
  - perfect (of commitment scheme), 13
- computational special soundness
  - (of sigma-protocol), 7
- computational strict soundness, 7
- computationally binding, 13
- computationally strict binding, 13
- diffusion operator
  - (Grover’s algorithm), 11
- distance statistical, 6
- trace, 6
- entropy
  - commitment, 7
  - extractor failure, 8
- Fiat-Shamir, 17
  - (proof system), 17
- Fischlin
  - (proof system), 18

---

23
A Auxiliary lemmas

Lemma 31 \( \sqrt{2(1 - (\cos \frac{\pi}{2n})^n)} \in \frac{\pi}{2\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right) \).

Proof. By Taylor’s theorem, for \( x \to 0 \),

\[
\cos x = 1 - \frac{x^2}{2} + O(x^4),
\]

(3)

\[
\ln(1 - x) = -x + O(x^2),
\]

(4)

\[
e^x = 1 + x + O(x^2).
\]

(5)

Hence for \( n \to \infty \),

\[
\ln \cos \frac{\pi}{2n} \in \ln \left(1 - \frac{x^2}{8n^2} + O(n^{-4})\right) \subseteq -\frac{x^2}{8n^2} + O(n^{-4}).
\]

Hence

\[
2n \left(1 - (\cos \frac{\pi}{2n})^n\right) \in 2n \left(1 - e^{n \left(-\frac{\pi^2}{8n^2} + O(n^{-4})\right)}\right) \subseteq 2n \left(\frac{\pi^2}{8n^2} + O(n^{-2})\right) \subseteq \frac{\pi^2}{4} + o(1).
\]

Thus

\[
\sqrt{n} \cdot \sqrt{2(1 - (\cos \frac{\pi}{2n})^n)} \in \frac{\pi}{2} + o(1)
\]

and

\[
\sqrt{2(1 - (\cos \frac{\pi}{2n})^n)} \in \frac{\pi}{2\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right).
\]

\( \square \)
Lemma 32 Let $X$ be a set. Let $P \subseteq X$ be a set. Let $S \subseteq X$ be uniformly random with $|S| = k$. Let $\varphi := |P|/|X|$. Let $\delta_{\min} \in [0, \varphi]$. Then

$$\Pr\left[\frac{|P \cap S|}{|S|} < \delta_{\min}\right] \leq e^{-2k(\varphi - \delta_{\min})^2}.$$  

Proof. Let $N := |X|$. Let $\delta := |P \cap S|/|S|$. We can describe the choice of $S$ as sampling $k$ elements $x_i \in X$ without replacement. Let $X_i := 0$ if $x_i \in P$ and $X_i := 1$ else. Then

$$1 - \delta = \sum_{i=1}^{k} X_i/k.$$  

And the $X_i$ result from sampling $k$ elements without replacement from a population $C$ consisting of $(1 - \varphi)N$ ones and $\varphi N$ zeros. Note that $\mu := 1 - \varphi$ is the expected value of each $X_i$. Thus we get

$$\Pr[\delta < \delta_{\min}] \leq \Pr[|1 - \delta \geq 1 - \delta_{min}|] = \Pr\left[\sum_{i} X_i \geq 1 - \delta_{min}\right] = \Pr\left[\sum_{i} X_i - \mu \geq \varphi - \delta_{\min}\right] \leq e^{-2k(\varphi - \delta_{\min})^2}.$$  

Here (*) uses Hoeffding’s inequality [21] (and the fact that $0 \leq t \leq 1 - \mu$ for $t := \varphi - \delta_{\min}$). Note that Hoeffding’s inequality also holds in the case of sampling without replacement, see [21, Section 6].

Lemma 33 Let $X$ be a finite and $Y$ a countable set. Let $\mathcal{D}$ be a distribution over $Y$. Let $H_{\frac{1}{2}}(\mathcal{D})$ denote the Rényi entropy of order $1/2$ of $\mathcal{D}$. For each $x \in X$, let $\mathcal{O}(x)$ be an independently chosen $y \leftarrow \mathcal{D}$. Let $y_1 \leftarrow \mathcal{D}$, and $y_2 := \mathcal{O}(x)$ for $x \succeq X$. Then

$$\SD((\mathcal{O}, y_1); (\mathcal{O}, y_2)) \leq \frac{1}{2\sqrt{n}} \sqrt{\frac{1}{\gamma} H_{\frac{1}{2}}(\mathcal{D})} \leq \frac{1}{\gamma} \sqrt{|Y|/|X|}.$$  

(I.e., we bound the statistical distance between an element $y_1$ chosen according to $\mathcal{D}$, and an element $y_2$ chosen by evaluating $\mathcal{O}$ on a random input, when the function $\mathcal{O}$ is known.)

Proof. Let $n := |X|$. For a function $f : X \rightarrow Y$, let $\mathcal{D}_f$ denote the empirical distribution of $f$, i.e., $\mathcal{D}_f(y) = \frac{1}{n}|\{x : f(x) = y\}|$. Let $j(f) := 2 \SD(\mathcal{D}, \mathcal{D}_f)$. And let $J_n := j(\mathcal{O})$, i.e., $J_n$ is a real-valued random variable. Then [21, Lemma 8] proves that $E[J_n] \leq \frac{1}{\sqrt{n}} \sum_{y \in Y} \sqrt{\mathcal{D}(y)} =: \gamma$. Since $H_{\frac{1}{2}}(\mathcal{D}) = \frac{1}{2\gamma} \log \left(\sum_{y \in Y} \mathcal{D}(y)^{\frac{1}{2}}\right)$ by definition, we have $\gamma = \frac{1}{\sqrt{n}} 2^{\frac{1}{2} H_{\frac{1}{2}}(\mathcal{D})}$. Since $H_{1/2}(\mathcal{D}) \leq \log |Y|$ for any distribution $\mathcal{D}$ on $Y$, we furthermore have $\gamma \leq \frac{1}{\sqrt{n}} 2^\frac{1}{2} \log |Y| = \sqrt{|Y|/|X|}$. Let $\SD(y_1, y_2|E)$ denote the statistical distance between $y_1$ and $y_2$ conditioned on an event $E$. We can finally compute:

$$\SD((\mathcal{O}, y_1); (\mathcal{O}, y_2)) = \sum_{f : X \rightarrow Y} \Pr[\mathcal{O} = f] \cdot SD(y_1, y_2|\mathcal{O} = f) = \sum_{f : X \rightarrow Y} \Pr[\mathcal{O} = f] \cdot SD(\mathcal{D}, \mathcal{D}_f) = \sum_{f : X \rightarrow Y} \Pr[\mathcal{O} = f] \cdot j(f) = \frac{1}{\gamma} E[J_n] \leq \frac{1}{\gamma} \gamma. \quad \Box$$

Lemma 34 Let $\text{bit}_p(x)$ denote the $p$-th bit of $x$. Let $X = \{0, 1\}^\ell$ for some $\ell$, and $k \geq 1$, $p \in \{1, \ldots, \ell\}$ be integers. Let $S \subseteq X$ be uniformly random with $|S| = k$. Let $x \leftarrow S$. Let $b^* \leftarrow \{0, 1\}$. Then $\SD((S, \text{bit}_p(x)); (S, b^*)) \leq 1/2\sqrt{k}.$

Proof. Let $P := \{x \in S : \text{bit}_p(x) = 1\}$. Let $\SD(X; Y|S)$ denote the statistical distance between $X$ and $Y$ conditioned on a specific choice of $S$. And $\Pr[S]$ denote the probability of a specific
choice of \( S \). Then

\[
\text{SD}((S, \text{bit}_p(x)); (S, b^*)) = \sum_S \Pr[S] \text{SD}(\text{bit}_p(x); b^* | S)
\]

\[
= \sum_S \Pr[S] \cdot \left| \Pr[x \in P : x \not\in S] - \Pr[b^* = 1 : b^* \not\in \{0, 1\}] \right|
\]

\[
= \sum_S \Pr[S] \cdot \left| \frac{|P|}{|S|} - \frac{1}{2} \right| \leq \sqrt{\sum_S \Pr[S] \left( \frac{|P|}{|S|} - \frac{1}{2} \right)^2} = \sqrt{\mathbb{E} \left[ \left( \frac{|P|}{|S|} - \frac{1}{2} \right)^2 \right]}
\]

\[
= \sqrt{\mathbb{E} \left[ \left( \frac{|P|}{|S|} - \mathbb{E} \left[ \frac{|P|}{|S|} \right] \right)^2 \right]} = \sqrt{\text{Var}[|P|/|S|]} \leq \frac{1}{\kappa} \sqrt{\text{Var}[|P|]}
\]

Here (*) uses Jensen’s inequality. And (**) that \(|S| = k\).

\(|P|\) is the number of successes when sampling \( k \) times without replacement from a population of size \( 2^k \) containing \( 2^{k-1} \) successes (the elements \( x \in \{0, 1\}^k \) with \( \text{bit}_p(x) = 1 \)). That is, \(|P|\) has hypergeometric distribution with parameters \( m = n = 2^{k-1} \) and \( N := k \) (in the notation of [34]). Thus (see [34]):

\[
\text{Var}[|P|] = \frac{mnN(m+n-N)}{(m+n)^2(m+n-1)} = \frac{k}{4} \frac{2^k-k}{2^k-1} \leq \frac{k}{4}.
\]

Summarizing,

\[
\text{SD}((S, \text{bit}_p(x)); (S, b^*)) \leq \frac{\sqrt{\text{Var}[|P|]}}{k/4} = \frac{1}{2\sqrt{k}}.
\]

**Lemma 35** Let \( C \) and \( R \) be finite sets, let \( k \geq 1 \) be an integer. Let \( S \) be a uniformly chosen subset of \( C \times R \) with \(|S| = k\). Let \( c' \overset{\$}{\leftarrow} C \), and \( r \overset{\$}{\leftarrow} S_{c'} := \{ r : (c', r) \in S \} \) (with \( r := \bot \notin R \) iff \( S_{c'} = \emptyset \)). Let \( (c'', r'') \overset{\$}{\leftarrow} S \).

Then \( \sigma := \text{SD}((S, c', r'); (S, c'', r'')) \leq \frac{2k^2}{|C| \times |R|} + \frac{\sqrt{|C|}}{2\sqrt{k}}. \)

**Proof.** In the following calculation, \( G \overset{\$}{\approx} H \) means that the distribution of \((S, c)\) when picked according to \( G \) has statistical distance \( \leq \varepsilon \) from the distribution of \((S, c)\) when picked according to \( H \). And \( G \equiv H \) means equality of these distributions \((G \overset{0}{\approx} H)\). And \([C \times R]_k\) denotes the set of all \( S \subseteq C \times R \) with \(|S| = k\). And \( x_1, \ldots, x_k \overset{\$}{\leftarrow} M \) means that the \( x_i \) are chosen uniformly but distinctly from \( M \) (drawn without replacing).

\[
S \overset{\$}{\leftarrow} [C \times R]_k, \ (c, r) \overset{\$}{\leftarrow} S
\]

\[
\equiv F(1), \ldots, F(k) \overset{\$}{\leftarrow} C \times R, \ S := \text{im} F, \ j \overset{\$}{\leftarrow} \{1, \ldots, k\}, \ (c, r) := F(j)
\]

\[
\overset{\$}{\approx} F(1), \ldots, F(k) \overset{\$}{\leftarrow} C \times R, \ S := \text{im} F, \ j \overset{\$}{\leftarrow} \{1, \ldots, k\}, \ (c, r) := F(j)
\]

\[
\equiv F_1(1), \ldots, F_1(k) \overset{\$}{\leftarrow} C, \ F_2(1), \ldots, F_2(k) \overset{\$}{\leftarrow} R, \ S := \text{im}((F_1, F_2)), \ j \overset{\$}{\leftarrow} \{1, \ldots, k\}, \ c := F_1(j), \ r := F_2(j)
\]

\[
\overset{\$}{\equiv} F_1(1), \ldots, F_1(k) \overset{\$}{\leftarrow} C, \ j \overset{\$}{\leftarrow} \{1, \ldots, k\}, \ c := F_1(j), \ F_2(1), \ldots, F_2(k) \overset{\$}{\leftarrow} R, \ S := \text{im}((F_1, F_2))
\]

\[
\overset{\$}{\equiv} F_1(1), \ldots, F_1(k) \overset{\$}{\leftarrow} C, \ F_2(1), \ldots, F_2(k) \overset{\$}{\leftarrow} R, \ S := \text{im}(F_1, F_2)
\]

\[
\overset{\$}{\equiv} F(1), \ldots, F(k) \overset{\$}{\leftarrow} C \times R, \ S := \text{im} F, \ c \overset{\$}{\leftarrow} C
\]

\[
\overset{\$}{\equiv} F(1), \ldots, F(k) \overset{\$}{\leftarrow} C \times R, \ S := \text{im} F, \ c \overset{\$}{\leftarrow} C
\]

\[
\equiv S \overset{\$}{\leftarrow} [C \times R]_k, \ c \overset{\$}{\leftarrow} C
\]
We restate an auxiliary lemma from [30, full version, Lemma 7]:

With (6), the lemma follows.

Proof. Fix a basis such that

Thus $TD((S, c'; (S, c''))) \leq \varepsilon_1 + \varepsilon_2$.

Since $r'$ given $S, c'$ has the same distribution as $r''$ given $S, c''$, it follows

$$SD((S, c', r'); (S, c'', r'')) \leq \varepsilon_1 + \varepsilon_2.$$  \hspace{1cm} (6)

We have $\varepsilon_1 \leq \sum_{i \neq j} \Pr[F(i) = F(j)] = \sum_{i \neq j} 1/|C \times R| \leq k^2/|C \times R|$. For a function $f : \{1, \ldots, k\} \rightarrow C$, let $D_f$ denote the empirical distribution of $f$, i.e., $D_f(c) = \frac{1}{k}\sum_{i : f(i) = c}$. Let $U$ denote the uniform distribution on $C$. Let $j(f) := 2 SD(U, D_f)$. And let $J_k := j(F_1)$ for $F_1(1), \ldots, F_1(k) \sim C$, i.e., $J_k$ is a real-valued random variable. Then [1] Lemma 8] proves that $E[J_k] \leq \frac{1}{\sqrt{k}} \sum_{c \in C} \sqrt{U(c)} = \sqrt{|C|/k}$. Then

$$\varepsilon_2 = SD((F_1, c); (F_1, u)) = \sum_{f} \Pr[F_1 = f] \cdot SD(D_f, U) = \sum_{f} \Pr[F_1 = f] \cdot \frac{1}{2} j(f) = \frac{1}{2} E[J_k] \leq \frac{1}{2} \sqrt{|C|/k}.$$  

With (6), the lemma follows. \hspace{1cm} $\square$

We restate an auxiliary lemma from [30] full version, Lemma 7):

**Lemma 36** Let $|\Psi_1\rangle, |\Psi_2\rangle$ be quantum states that can be written as $|\Psi_i\rangle = |\Psi_i^+\rangle + |\Phi^+\rangle$ where both $|\Psi_i^+\rangle$ are orthogonal to $|\Phi^+\rangle$. Then $TD(|\Psi_1\rangle, |\Psi_2\rangle) \leq 2 |||\Psi_2^+|||.$

**Lemma 37** Let $|\Psi_1\rangle, |\Psi_2\rangle$ be quantum states. Then $TD(|\Psi_1\rangle, |\Psi_2\rangle) \leq |||\Psi_1\rangle - |\Psi_2\rangle||$.

**Proof.** Fix a basis such that $|\Psi_1\rangle = |0\rangle$ and $|\Psi_2\rangle = \alpha|0\rangle + \beta|1\rangle$. Then $|\alpha|^2 + |\beta|^2 = 1$ and

$$TD(|\Psi_1\rangle, |\Psi_2\rangle)^2 \leq 1 - |\langle \Psi_1 | \Psi_2 \rangle|^2 = 1 - |\alpha|^2 = |\beta|^2 \leq |1 - |\alpha|^2 + |\beta|^2| = |||\Psi_1\rangle - |\Psi_2\rangle||^2.$$  

Here (*) uses that the trace distance is bounded in terms of the fidelity (e.g., [25] (9.101)). Thus $TD(|\Psi_1\rangle, |\Psi_2\rangle) \leq |||\Psi_1\rangle - |\Psi_2\rangle||$. \hspace{1cm} $\square$

**Lemma 38 (Preimage search in a random function)** Let $\gamma \in [0, 1]$. Let $Z$ be a finite set. Let $q \geq 0$ be an integer. Let $F : Z \rightarrow \{0, 1\}$ be the following function: For each $z$, $F(z) := 1$ with probability $\gamma$, and $F(z) := 0$ else. Let $N$ be the function with $\forall z : N(z) = 0$.

If an oracle algorithm $A$ makes at most $q$ queries, then

$$\left| \Pr[b = 1 : b \leftarrow A^f] - \Pr[b = 1 : b \leftarrow A^N] \right| \leq 2q\sqrt{7}.$$  

**Proof.** We can assume that $A$ uses three quantum registers $A, K, V$ for its state, oracle inputs, and oracle outputs. For a function $f$, let $O_f[a, k, v] := |a, k, v \oplus f(k)\rangle$. Then the final state of $A^f$ is $(UO_f)^q|\Psi_0\rangle$ for some unitary $U$ and some initial state $|\Psi_0\rangle$. The output $b$ of $A^f$ is then obtained by applying $f$ to the projective measurement $P_{\text{final}}$ on that final state.

Let $|\Psi_f^i\rangle := (UO_f)^i|\Psi_0\rangle$ and $|\Phi_i\rangle := (UO_N)^i|\Psi_0\rangle = U^i|\Psi_0\rangle$. (Recall: $N$ is the constant-zero function.)

We compute:

$$D_i^f := TD(|\Psi_f^i\rangle, |\Phi_i\rangle) = TD(O_f |\Psi_f^{i-1}\rangle, |\Phi^{i-1}\rangle) \leq TD(O_f |\Psi_f^{i-1}\rangle, O_f |\Phi^{i-1}\rangle) + TD(O_f |\Phi^{i-1}\rangle, |\Psi_f^{i-1}\rangle) = D_{i-1}^f + TD(O_f |\Phi^{i-1}\rangle, |\Psi_f^{i-1}\rangle).$$  

27
Furthermore $D_0^f = TD(|\Psi_0\rangle, |\Psi_0\rangle) = 0$, thus $D_0^f \leq \sum_{i=0}^{q-1} TD(O_f|\Psi^i\rangle, |\Psi^i\rangle)$.

Let $Q_z$ be the projector projecting $K$ onto $|z\rangle$ (i.e., $Q_z = I \otimes |z\rangle\langle z| \otimes I$). $Q_f$ is the projector projecting $K$ onto all $|z\rangle$ with $f(z) = 1$ (i.e., $Q_f = \sum_{z:f(z)=1} Q_z$). Let $\alpha_f := Pr[F = f]$.

We then have
\[
\sum_f \alpha_f \| Q_f |\Psi_i\rangle \|^2 = \sum_f \alpha_f \sum_{z:f(z)=1} \| Q_z |\Psi_i\rangle \|^2 = \sum_{z \in Z} \sum_{f:f(z)=1} \alpha_f \| Q_z |\Psi_i\rangle \|^2 \\
(=) \lambda \sum_{z} \| Q_z |\Psi_i\rangle \|^2 = \lambda \| |\Psi_i\rangle \|^2 = \lambda. (7)
\]

Here $(*)$ uses that $Q_f = \sum_{z:f(z)=1} Q_z$ and all $Q_z|\Psi_i\rangle$ are orthogonal. And $(**)\) uses that $\sum_{z:f(z)=1} \alpha_f = Pr[F(x) = 1] = \lambda$.

Then
\[
\sum_f \alpha_f \text{TD}(|\Psi^q_f\rangle, |\Psi^q\rangle) = \sum_f \alpha_f D^q_f \leq \sum_{f,i} \alpha_f \text{TD}(O_f|\Psi^i\rangle, |\Psi^i\rangle) \\
= \sum_{f,i} \alpha_f \text{TD}(O_fQ_f|\Psi^i\rangle + (1 - Q_f)|\Psi^i\rangle, Q_f|\Psi^i\rangle + (1 - Q_f)|\Psi^i\rangle) \\
\overset{(\ref{eq:td})}{\leq} \sum_{f,i} \alpha_f 2\| Q_f |\Psi^i\rangle \| \leq 2 \sum_i \sqrt{\sum_f \alpha_f \| Q_f |\Psi^i\rangle \|^2} \\
\overset{(*)}{=} 2 \sum_i \sqrt{\lambda} = 2q\sqrt{\lambda}. (8)
\]

Here $(*)$ uses Lemma 36 and $(**\) uses Jensen’s inequality. Finally,
\[
|\Pr[b = 1 : b \leftarrow A^f] - \Pr[b = 1 : b \leftarrow A^N]| \\
\leq \sum_f \alpha_f |\Pr[b = 1 : b \leftarrow A^f] - \Pr[b = 1 : b \leftarrow A^N]| \\
\leq \sum_f \alpha_f \text{TD}(|\Psi^q_f\rangle, |\Psi^q\rangle) \overset{(*)}{\leq} 2q\sqrt{\lambda}. \quad \square
\]

The following lemma formalizes that an oracle $O_1$ does not help (much) in finding a value $w$ if $O_1$ only gives answers when $w$ is already contained in its input.

**Lemma 39 (Removing redundant oracles 1)** Let $w$, $O_1$, $O_2$ be chosen according to some joint distribution. Here $w$ is a bitstring, and $O_1$, $O_2$ are oracles, and $O_1$ is classical (i.e., $\forall x, y, \exists y', O_1|x\rangle|y\rangle = |x\rangle|y'\rangle$). Fix a function $f$. Assume that for all $x$ with $f(x) \neq w$, $O_1(x) = 0$. (In other words, $O_1|x\rangle|y\rangle = |x\rangle|y\rangle$ for $f(x) \neq w$.)

Let $A$ be an oracle machine that makes at most $q$ queries to $O_1$ and $q'$ queries to $O_2$. Then there is another oracle machine $\hat{A}$ that makes at most $q'$ queries to $O_2$ such that:
\[
\text{Pr}[w = w' : w' \leftarrow A^{O_1,O_2}] \leq 2(q + 1) \sqrt{\text{Pr}[w = w : w' \leftarrow \hat{A}^{O_2}]}.
\]

*Proof.* We can assume that $A$ is unitary until the final measurement of its output. Then the final state of $A$ before that measurement is $|\Psi^q\rangle := (U_2O_1)^q(U_2)|\Psi\rangle$ for some unitary $U_2$ depending only on $O_2$, and $O_1$ operating on quantum registers $K, V$ for oracle input and output, and $|\Psi\rangle$ being some initial state independent of $O_1, O_2, w$. Let $|\Psi_i\rangle := (U_2O_1)^q-iU_2^{q+i}|\Psi\rangle$. Note that $|\Psi_0\rangle = |\Psi\rangle$. Let $P_X := \sum_{x:f(x)=w} |x\rangle \langle x| \otimes I$ and $P_X := 1 - P_X$. Note that since $O_1|x\rangle|y\rangle = |x\rangle|y\rangle$.
for \( f(x) \neq w \), we have \( O_1 = O_1P_X + \bar{P}_X \). We have for \( i = 1, \ldots, q \):
\[
\text{TD}(|\Psi_{i-1}⟩, |\Psi_i⟩) = \text{TD}((U_2O_1)^{q-i}(U_2O_1)U_2^i |Ψ⟩, (U_2O_1)^{q-i}U_2^i |Ψ⟩)
= \text{TD}(O_1U_2^i |Ψ⟩, U_2^i |Ψ⟩)
= \text{TD}(O_1^2XU_2^i |Ψ⟩ + \bar{P}_XU_2^i |Ψ⟩, P_XU_2^i |Ψ⟩ + \bar{P}_XU_2^i |Ψ⟩)
\leq 2\|P_XU_2^i |Ψ⟩\|.
\]

Here \( (\ast) \) uses Lemma 36 (using that \( |Ψ_i⟩ := O_1P_XU_2^i |Ψ⟩ \) and \( |Ψ^*_i⟩ := P_XU_2^i |Ψ⟩ \) are both orthogonal to \( |Ψ^∗⟩ := \bar{P}_XU_2^i |Ψ⟩ \) because \( O_1 \) is classical and therefore does not leave the image of \( P_X \).

Thus \( \text{TD}(|Ψ^∗⟩, |Ψ_q⟩) \leq \sum_{i=1}^q 2\|P_XU_2^i |Ψ⟩\| \). For \( i = 1, \ldots, q \), let \( A_i^{O_2} \) be the oracle algorithm that computes \( U_2^i |Ψ⟩ \) and measures register \( K \) in the computational basis, giving outcome \( x \), and then outputs \( f(x) \). (Note that \( A_i \) does not need access to \( O_1 \) because \( U_2 \) does not depend on \( O_1 \).) Then \( \Pr[w = w' : w' ← A_i] = \|P_XU_2^i |Ψ⟩\|^2 \). Let \( A_0 \) be the oracle machine that performs the same operations as \( A \), except that it omits all calls to \( O_1 \). That is, its state before measuring the output is \( |Ψ_q⟩ \). Thus
\[
\Pr[w = w' : w' ← A^{O_1,O_2}] - \Pr[w = w' : w' ← A_0^{O_2}] \leq \text{TD}(|Ψ^∗⟩, |Ψ_q⟩) \leq \sum_{i=1}^q 2\sqrt{\Pr[w = w' : w' ← A_i^{O_2}]}
\]

Let \( \hat{A}^{O_2} \) be the algorithm that picks \( i \leftarrow \{0, \ldots, q\} \) and runs \( A_i \). Then
\[
\Pr[w = w' : w' ← A^{O_1,O_2}] \leq \sum_{i=1}^q 2\sqrt{\Pr[w = w' : w' ← A_i^{O_2}]} + \Pr[w = w' : w' ← A_0^{O_2}]
\leq 2(q + 1) \sum_{i=0}^q \sqrt{\frac{1}{q+1} \Pr[w = w' : w' ← A_i^{O_2}]}
\leq 2(q + 1) \left( \sum_{i=0}^q \frac{1}{q+1} \Pr[w = w' : w' ← A_i^{O_2}] \right)
\leq 2(q + 1) \sqrt{\Pr[w = w' : w' ← \hat{A}^{O_2}]}.
\]

Here \( (\ast) \) uses Jensen’s inequality.

The following lemma formalizes that if \( w \) is a random bitstring that can be accessed only by querying an oracle \( O_1 \) on some input \( x \in X \), then the probability of finding \( w \) using \( O_1 \) is bounded in terms of the probability of finding some \( x \in X \) without using \( O_1 \).

**Lemma 40 (Removing redundant oracles 2)** Let \( w, X, O_1, O_2 \) be chosen according to some joint distribution such that \( w \) and \( O_2 \) are stochastically independent. Here \( X \) is a set of bitstrings, and \( O_1, O_2 \) are oracles, and \( O_1 \) is classical (\( i.e., \forall x, y. \exists y'. O_1|x⟩|y⟩ = |x⟩|y'⟩ \)). And \( w \) is uniformly distributed on \( \{0, 1\}^\ell \). Assume that for all \( x \notin X, O_1(x) = 0 \). (In other words, \( O_1|x⟩|y⟩ = |x⟩|y⟩ \) for \( x \notin X \).)

Let \( A \) be an oracle machine that makes at most \( q \) queries to \( O_1 \) and \( q' \) queries to \( O_2 \). Then there is another oracle machine \( \hat{A} \) that makes at most \( q \) queries to \( O_2 \) such that:
\[
\Pr[w = w' : w' ← A^{O_1,O_2}] \leq 2q\sqrt{\Pr[x \in X : x ← \hat{A}^{O_2}]} + 2^{-\ell}
\]
Proof. Let \( P_X := \sum_{x \in X} |x\rangle \langle x| \otimes I \) and \( \bar{P}_X := 1 - P_X \). Note that since \( \mathcal{O}_1|x\rangle = |x\rangle \) for \( x \notin X \), we have \( \mathcal{O}_1 = \mathcal{O}_1 P_X + \bar{P}_X \).

Let \( |\Psi\rangle, |\Psi^*\rangle, |\Psi_q\rangle \) and \( U_2 \) be defined as in the proof of Lemma 39. (Remember that all of these only depend on \( \mathcal{O}_2 \), not \( \mathcal{O}_1 \).) Exactly as in Lemma 39 we get \( \text{TD}(\langle \Psi^* |, |\Psi_q\rangle) \leq \sum_{i=1}^q 2\|P_X U_2^i |\Psi\rangle\| \). For \( i = 1, \ldots, q \), let \( A_i^{O_2} \) be the oracle algorithm that computes \( U_2^i |\Psi\rangle \) and measures register \( K \) in the computational basis and outputs the outcome. Then \( \Pr[|x \in X : x \leftarrow A_i^{O_2} = ||P_X U_2^i |\Psi\rangle||^2] \).

Like in the proof of Lemma 39 let \( A_0 \) be the oracle machine that performs the same operations as \( A \), except that it omits all calls to \( \mathcal{O}_1 \). That is, its state before measuring the output is \( |\Psi_q\rangle \). Let \( A_0^{O_2} \) pick a random \( i \in \{1, \ldots, q\} \) (not \( i \in \{0, \ldots, q\} \) as in Lemma 39) and run \( A_i^{O_2} \). Then

\[
\Pr[w = w' : w' \leftarrow A_0^{O_2}] - \Pr[w = w' : w' \leftarrow A_0^{O_2}] \leq \text{TD}(\langle \Psi^* |, |\Psi_q\rangle) \\
\leq 2q \sum_{i=1}^q \frac{1}{q} \sqrt{\Pr[x \in X : x \leftarrow A_i^{O_2}]} \leq 2q \sqrt{\sum_{i=1}^q \frac{1}{q} \Pr[x \in X : x \leftarrow A_i^{O_2}]} \\
= 2q \sqrt{\Pr[x \in X : x \leftarrow A_0^{O_2}].} \tag{9}
\]

Here \((*)\) uses Jensen’s inequality.

Since \( w \) and \( O_2 \) are independent and \( w \) is uniform on \( \{0, 1\}^\ell \), \( \Pr[w = w' : w' \leftarrow A_0^{O_2}] \leq 2^{-\ell} \). With \((*)\), we get \( 2q \sqrt{\Pr[x \in X : x \leftarrow A_0^{O_2}] \geq \Pr[w = w' : w' \leftarrow A_0^{O_2}] = 2^{-\ell}. \)

\[
\Box
\]

\section*{B Proofs for Section 3}

\subsection*{B.1 Lemmas for Section 3.1}

**Lemma 41** Let \( |\Psi\rangle \) be a state, chosen according to some distribution. Let \( |\perp\rangle \) be a fixed state orthogonal to \( |\Psi\rangle \). (Such a state can always be found by extending the dimension of the Hilbert space containing \( |\Psi\rangle \) and using the new basis state as \( |\perp\rangle \).)

Let \( O_y \) be an oracle with \( O_y |\Psi\rangle = |\perp\rangle, O_y |\perp\rangle = |\Psi\rangle \), and \( O_y |\psi^-\rangle = |\psi^-\rangle \) for any \( |\psi^-\rangle \) orthogonal to both \( |\Psi\rangle \) and \( |\perp\rangle \). Let \( O_{\text{Ref}} := I - 2 |\Psi\rangle \langle \Psi| \).

Let \( O \) be an oracle, not necessarily independent of \( |\Psi\rangle \). Let \( |\Phi\rangle \) be a quantum state, not necessarily independent of \( |\Psi\rangle \).

Let \( n \geq 0 \) be an integer. Let \( |R\rangle := |\alpha_1\rangle \otimes \cdots \otimes |\alpha_n\rangle \) where \( |\alpha_j\rangle := (\cos \frac{\pi}{2n}) |\Psi\rangle + (\sin \frac{\pi}{2n}) |\perp\rangle \).

Then there is an oracle algorithm \( B \) that makes \( q_y \) queries to \( O_{\text{Ref}} \) and makes the same number of queries to \( O \) as \( A \) such that:

\[
\text{TD}(B^{O_{Ref}, O}(|R\rangle, |\Phi\rangle), A^{O_{Ref}, O}(|\Phi\rangle)) \leq \frac{2q_y}{\sqrt{n}} + q_y o(\frac{1}{\sqrt{n}}).
\]

**Proof.** In this proof, we use the following shorthand notation: \( |\Phi\rangle = |\Phi^\prime\rangle \pm \varepsilon \) means that \( |||\Phi^\prime\rangle - |\Phi^\prime\rangle|| \leq \varepsilon \).

We first show that

\[
S|\perp\rangle |R\rangle = |\Psi\rangle |R\rangle \pm \varepsilon_n \quad \text{and} \quad S^\dagger|\Psi\rangle |R\rangle = |\perp\rangle |R\rangle \pm \varepsilon_n \quad \text{with} \quad \varepsilon_n := \frac{\pi}{2\sqrt{n}} + o(\frac{1}{\sqrt{n}}), \tag{10}
\]

where \( S|\Phi_0\rangle \cdots |\Phi_n\rangle := |\Phi_1\rangle \cdots |\Phi_n\rangle |\Phi_0\rangle \) (cyclic shift) and \( |R\rangle \) is as in the statement of the lemma (the reservoir state).
We have
\[
(S|\bot)|R\rangle^\dagger (|\Psi\rangle|R\rangle) = \left(\alpha_1|\alpha_2\rangle\ldots|\alpha_n\rangle|\bot\rangle\right)^\dagger \left(\alpha_1|\alpha_2\rangle\ldots|\alpha_n\rangle|\bot\rangle\right)
\]
\[= \langle \alpha_1|\Psi\rangle \cdot \prod_{j=1}^{n-1} \langle \alpha_{j+1}|\alpha_j\rangle \cdot \langle \bot|\alpha_{n-1}\rangle \]
\[= \cos \frac{\pi}{n} \cdot \prod_{j=1}^{n-1} \cos \left(\frac{j+1}{2n} - \frac{\pi}{2n}\right) \cdot \sin \frac{\pi}{2n} = (\cos \frac{\pi}{2n})^n.
\]

Here (*) uses that $|\Psi\rangle$ and $|\bot\rangle$ are orthogonal (and the definition of $|\alpha_j\rangle$ from the statement of the lemma). For any quantum states $|\Phi\rangle, |\Phi\rangle'$ we have $||\Phi\rangle - |\Phi\rangle'||^2 = \langle |\Phi\rangle - |\Phi\rangle'\rangle^\dagger (|\Phi\rangle - |\Phi\rangle') = 1 - \langle |\Phi\rangle'\rangle - \langle |\Phi\rangle\rangle + 1 = 2(1 - \Re(|\Phi\rangle |\Phi\rangle'))$ where $\Re$ denote the real part. Thus $||S|\bot\rangle|R\rangle - |\Psi\rangle|R\rangle|| \leq \sqrt{2(1 - \cos \frac{\pi}{2n})^n} \leq \frac{\pi}{2\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right) = \varepsilon_n$. (The asymptotic bound uses Lemma 31). This shows the lhs of (10). The rhs follows from the rhs by applying the unitary $S^\dagger$ on both sides.

Let $U_\Psi$ denote the unitary computed by circuit $\Pi$ on page 9. We will show that for any $|\Phi\rangle$,
\[
U_\Psi |\Phi\rangle|R\rangle|0\rangle = (O_\Psi |\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n. \tag{11}
\]

By linearity of $U_\Psi, O_\Psi$ and the triangle inequality, it is sufficient to verify this for $|\Phi\rangle = |\Psi\rangle$, $|\Phi\rangle = |\bot\rangle$, and $|\Phi\rangle$ orthogonal to both $|\Psi\rangle, |\bot\rangle$. In an execution of circuit $\Pi$ on state $|\Phi\rangle|R\rangle|0\rangle$, we denote the state before $S$ with $|\Phi_1\rangle$, the state after $S$ with $|\Phi_2\rangle$, the state before $S^\dagger$ with $|\Phi_3\rangle$, and the state after $S^\dagger$ with $|\Phi_4\rangle$. We denote the final state with $|\Phi\rangle' = U_\Psi |\Phi\rangle|R\rangle|0\rangle$.

For $|\Phi\rangle = |\Psi\rangle$, we have
\[
|\Phi_1\rangle = |\Psi\rangle|R\rangle|0\rangle, \quad |\Phi_2\rangle = |\Psi\rangle|R\rangle|0\rangle, \\
|\Phi_3\rangle = |\Psi\rangle|R\rangle|1\rangle, \quad |\Phi_4\rangle = (|\bot\rangle|R\rangle|1\rangle \pm \varepsilon_n, \\
|\Phi\rangle' = (|\bot\rangle|R\rangle|0\rangle \pm \varepsilon_n = (O_\Psi |\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n.
\]

For $|\Phi\rangle = |\bot\rangle$, we have
\[
|\Phi_1\rangle = (|\bot\rangle|R\rangle|1\rangle, \quad |\Phi_2\rangle = (|\bot\rangle|R\rangle|1\rangle \pm \varepsilon_n, \\
|\Phi_3\rangle = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n, \quad |\Phi_4\rangle = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n, \\
|\Phi\rangle' = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n = (O_\Psi |\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n.
\]

And for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$ and $|\bot\rangle$, we have
\[
|\Phi_1\rangle = |\Phi\rangle|R\rangle|0\rangle, \quad |\Phi_2\rangle = |\Phi\rangle|R\rangle|0\rangle, \\
|\Phi_3\rangle = |\Phi\rangle|R\rangle|0\rangle, \quad |\Phi_4\rangle = |\Phi\rangle|R\rangle|0\rangle, \\
|\Phi\rangle' = |\Phi\rangle|R\rangle|0\rangle = (O_\Psi |\Phi\rangle)|R\rangle|0\rangle.
\]

Thus (11) holds.

Without loss of generality, we assume that the algorithm $A$ is unitary and only (optionally) performs a final measurement at the end. Let $B$ be like $A$, except that $B$ has additional register $R, Z$ initialized with $|R\rangle, |0\rangle$, and that $B$ computes circuit $\Pi$ on $X, R, Z$ whenever $A$ invokes $O_\Psi$ on $X$. (And when $A$ performs a controlled invocation of $O_\Psi$, then $B$ executes the circuit with all operations accordingly controlled.) Let $|\Phi_0\rangle$ be the initial state of $A$ and $B$, and let $|\Phi_A\rangle, |\Phi_B\rangle$ be the final state of $A, B$ (right before the final measurement), respectively. Then by induction, from (11) we get $||\Phi_A\rangle - |\Phi_B\rangle|| \leq q_\Psi \varepsilon_n$. By Lemma 37, $TD(|\Phi_A\rangle - |\Phi_B\rangle) \leq q_\Psi \varepsilon_n$. Thus
\[
TD\left(B^{O_{Ref},O,\Psi}(|R\rangle, |\Phi\rangle), A^{O_{Ref},O,\Psi}(|\Phi\rangle)\right) \leq q_\Psi \varepsilon_n \leq \frac{\pi q_\Psi}{2\sqrt{n}} + q_\Psi o\left(\frac{1}{\sqrt{n}}\right). \quad \Box
\]
Lemma 42 Let $|\Psi\rangle$ be a state, chosen according to some distribution. Let $O_{\text{Ref}} := I - 2|\Psi\rangle\langle\Psi|$. Let $O$ be an oracle, not necessarily independent of $|\Psi\rangle$. Let $|\Phi\rangle$ be a quantum state, not necessarily independent of $|\Psi\rangle$. Let $A$ be an oracle algorithm that makes $q_{\text{Ref}}$ queries to $O_{\text{Ref}}$. Let $m \geq 0$ be an integer. Then there is an oracle algorithm $B$ that makes the same number of queries to $O$ as $A$ such that:

$$\text{TD}(B^O(|\Psi\rangle^{\otimes m}, |\Phi\rangle), A^{O_{\text{Ref}}, O}(|\Phi\rangle)) \leq \frac{2q_{\text{Ref}}}{\sqrt{m + 1}}.$$ 

Proof. Let $H$ be the space in which $|\Psi\rangle$ lives (i.e., $|\Psi\rangle \in H$). Let $S$ denote a cyclic shift on $(m + 1)$-partite states. That is, $S(|\Phi_0\rangle|\Phi_1\rangle \cdots |\Phi_m\rangle := |\Phi_1\rangle \cdots |\Phi_m\rangle |\Phi_0\rangle$ for all $|\Phi_i\rangle \in H$. (extended linearly to all of $H^{\otimes m+1}$). $S$ is unitary.

Let $V \subseteq H^{\otimes m+1}$ be the space of states invariant under $S$. I.e., $|\Phi\rangle \in V$ iff $S|\Phi\rangle = |\Phi\rangle$.

Let $U_V$ be the unitary with $U_V|\Phi\rangle = -|\Phi\rangle$ for $|\Phi\rangle \in V$, and $U_V|\Phi\rangle = |\Phi\rangle$ for $|\Phi\rangle$ orthogonal to $V$. (That is, $U_V = I - 2P_V$ where $P_V$ is the orthogonal projector onto $V$.)

In this proof, we use the following shorthand notation: $|\Phi\rangle = |\Phi\rangle' \pm \varepsilon$ means that $|||\Phi\rangle - |\Phi\rangle'|| \leq \varepsilon$.

Let $|T\rangle := |\Psi\rangle^{\otimes m}$.

We show that for any $|\Phi\rangle \in H$,

$$U_V|\Phi\rangle|T\rangle = \langle O_{\text{Ref}}|\Phi\rangle|T\rangle \pm \frac{2}{\sqrt{m+1}}.$$ \hspace{1cm} (12)

We first show this for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$. We decompose $|\Phi\rangle|T\rangle = \alpha|\chi\rangle + \beta|\kappa\rangle$ for quantum states $|\chi\rangle \in V$, and $|\kappa\rangle$ orthogonal to $V$. Since $|\chi\rangle \in V$, we have $\langle\chi| = \langle\chi|S^j$ for any $j$. Thus

$$|\alpha| = |\langle\chi|(|\Phi\rangle|T\rangle)| = \left| \frac{1}{m+1} \sum_{j=0}^{m} \langle\chi|S^j(|\Phi\rangle|T\rangle) \right|$$

$$= \frac{1}{m+1} \left| \langle\chi| \left( \sum_{j=0}^{m} S^j|\Phi\rangle|T\rangle \right) \right| \leq \frac{1}{m+1} \sqrt{m+1} = \frac{1}{\sqrt{m+1}}.$$ 

Here (*) follows from the fact that $|\Psi\rangle$ and $|\Phi\rangle$ are orthogonal, and hence all $S^j|\Phi\rangle|T\rangle$ ($j = 0, \ldots, m$) are orthogonal, and thus $\|\sum_j S^j|\Phi\rangle|T\rangle\| = \sqrt{m+1}$. Thus

$$\|U_V|\Phi\rangle|T\rangle - \langle O_{\text{Ref}}|\Phi\rangle|T\rangle\| = \||\Phi\rangle|T\rangle - 2\alpha|\chi\rangle - |\Phi\rangle|T\rangle\| = 2|\alpha| \leq \frac{2}{\sqrt{m+1}}.$$ 

Thus shows (12) for the case that $|\Phi\rangle$ is orthogonal to $|\Psi\rangle$. If $|\Phi\rangle = |\Psi\rangle$, (12) follows since $|\Phi\rangle|T\rangle = |\Psi\rangle^{\otimes m} \in V$ and thus $U_V|\Phi\rangle|T\rangle = -|\Phi\rangle|T\rangle = O_{\text{Ref}}|\Phi\rangle|T\rangle$. By linearity and the triangle inequality, (12) then holds for all $|\Phi\rangle \in H$.

Without loss of generality, we assume that the algorithm $A$ is unitary and only (optionally) performs a final measurement at the end. Let $B$ be like $A$, except that $B$ has additional register $T$ initialized with $|T\rangle$ (which is given as input), and that $B$ applies $U_V$ to $X,T$ whenever $A$ invokes $O_{\text{Ref}}$ on $X$. (And when $A$ performs a controlled invocation of $O_{\text{Ref}}$, then $B$ executes the circuit with all operations accordingly controlled.) Let $|\Phi_0\rangle$ be the initial state of $A$ and $B$, and let $|\Phi_A\rangle, |\Phi_B\rangle$ be the final state of $A, B$ (right before measuring the output), respectively. Then by induction, from (12) we get $\| |\Phi_A\rangle - |\Phi_B\rangle \| \leq \frac{2q_{\text{Ref}}}{\sqrt{m+1}}$. By Lemma 37 $\text{TD}(|\Phi_A\rangle - |\Phi_B\rangle) \leq \frac{2q_{\text{Ref}}}{\sqrt{m+1}}$. Thus

$$\text{TD}(B^O(|T\rangle, |\Phi\rangle), A^{O_{\text{Ref}}, O}(|\Phi\rangle)) \leq \frac{2q_{\text{Ref}}}{\sqrt{m+1}}.$$
B.2 Proofs of Theorem 4

Before proving Theorem 4, we first restate a result by Zhandry on which we build.

Lemma 43 (Distinguishing semi-constant distributions [37]) Let \( Z, Y \) be finite sets, let \( \lambda \in [0, 1] \), and let \( q \geq 0 \) be an integer. Let \( D_Y \) be a distribution on \( Y \).

Let \( H : Z \to Y \) be chosen as: \( H(z) \sim D_Y \) for each \( z \in Z \).

Let \( G : Z \to Y \) be chosen as: first pick \( y \sim D_Y \), then for all \( z \in Z \), with probability \( \lambda \), let \( G(z) := y \), and with probability \( 1 - \lambda \), let \( G(z) := D_Y \). (\( G \) is “semi-constant” in the language of [37].)

Then for any oracle algorithm making at most \( q \) queries:

\[
\left| \Pr[b = 1 : b \sim A^H] - \Pr[b = 1 : b \sim A^G] \right| \leq \frac{\sqrt{q}}{2} \lambda^2.
\]

Note that [37, Coro. 4.3] shows this only for \( D_Y \) being the uniform distribution on \( Y \). However, the general case is an immediate consequence because the uniform outputs of \( H \) or \( G \), respectively, can be seen as random coins for producing values distributed according to \( D_Y \) [38].

Proof of Theorem 4 Fix some integer \( s \) and some \( \lambda \in (0, 1) \). (We will provide concrete values at the end of this proof.)

For an integer \( \nu \geq 0 \), let \( D^s_\nu \) be the following distribution on \( \{1, \ldots, \nu, \perp\} \): Return \( j \in \{1, \ldots, \nu\} \) with probability \( (1 - \lambda)^{j-1} \lambda \). Return \( \perp \) with probability \( (1 - \lambda)^\nu \).

Let \( H_\nu : Z \to Y \) be the function chosen as follows: Pick \( y_1, \ldots, y_\nu \sim D_Y \). For each \( z \in Z \), pick \( j \sim D^s_\nu \), and set \( H_\nu(z) := y_j \) if \( j \neq \perp \) and \( H_\nu(z) := D_Y \) if \( j = \perp \).

Note that \( H_0 \) has the same distribution as \( H \). And \( H_1 \) is semi-constant in the sense of Lemma 43.

We will now show that \( \left| \Pr[b = 1 : b \sim A^{H_\nu}] - \Pr[b = 1 : b \sim A^{H_{\nu+1}}] \right| \leq \frac{\sqrt{q}}{2} \lambda^2 \) for all \( \nu \geq 0 \).

Fix \( \nu \geq 0 \). Let \( H^* : Z \to Y \cup \{\perp\} \) be chosen as follows: Pick \( y_1, \ldots, y_\nu \sim D_Y \). For each \( z \in Z \), pick \( j \sim D^s_\nu \) and set \( H^*(z) := y_j \) if \( j \neq \perp \) and \( H^*(z) := \perp \) if \( j = \perp \). For functions \( f, g \), let \( f|g \) be the function with \( (f|g)(z) = f(z) \) if \( f(z) \neq \perp \) and \( (f|g)(z) = g(z) \) if \( f(z) = \perp \). Then \( H^*|H \) has the same distribution as \( H_\nu \). And \( H^*|H_1 \) has the same distribution as \( H_{\nu+1} \). Let \( B^f \) be the algorithm that picks \( H^* \) and then executes \( A^{H^*}\).

Note that \( B^f \) makes at most \( q \) queries to \( f \).

\[
\Pr[b = 1 : b \sim A^{H_\nu}] = \Pr[b = 1 : b \sim A^{H^*|H}] = \Pr[b = 1 : b \sim B^H]
\]

\[
\overset{\text{\simeq}}{=} \Pr[b = 1 : b \sim B^{H_1}] = \Pr[b = 1 : b \sim A^{H^*|H_1}] = \Pr[b = 1 : b \sim A^{H_{\nu+1}}].
\]

Here \( p \overset{\text{\simeq}}{=} q \) means that \( |p - q| \leq \varepsilon_1 \) with \( \varepsilon_1 := \frac{\sqrt{q}}{2} \lambda^2 \). And the \( \overset{\text{\simeq}}{=} \)-equation follows from Lemma 43 (with \( H := H \) and \( G := H_1 \)).

Thus \( \left| \Pr[b = 1 : b \sim A^{H_\nu}] - \Pr[b = 1 : b \sim A^{H_{\nu+1}}] \right| \leq \frac{\sqrt{q}}{2} \lambda^2 \) for all \( \nu \geq 0 \). Using that \( H \) and \( H_0 \) have the same distribution, we get \( \left| \Pr[b = 1 : b \sim A^H] - \Pr[b = 1 : b \sim A^{H_1}] \right| \leq \frac{\sqrt{q}}{2} \lambda^2 =: \varepsilon_2 \).

Let \( D_s \) be the distribution on \( \{1, \ldots, s\} \) that picks \( j \in \{1, \ldots, s\} \) with probability \( \frac{(1 - \lambda)^{j-1} \lambda}{1 - (1 - \lambda)^s} \).

That is, \( D_s \) is \( D^s_\nu \) conditioned on not returning \( \perp \). Let \( G : Z \to Y \) be chosen as in the statement of the lemma. (I.e., \( G(z) = y_i \) with \( i \sim D_s \).)

Let \( F : Z \to \{0, 1\} \) be the following function: For each \( z \), \( F(z) := 1 \) with probability \( (1 - \lambda)^s \), and \( F(z) := 0 \) else. Let \( N \) be the function with \( \forall z : N(z) = 0 \).

For a function \( f : Z \to \{0, 1\} \), let \( H_f \) be the following function: For all \( z \in Z \), if \( f(z) = 0 \) then \( H_f(z) := G(z) \), and if \( f(z) = 1 \) then \( H_f(z) := H(z) \).
Then $H_N = G$. And $H_F$ has the same distribution as $H_s$. Let $C^F$ be an oracle algorithm that picks $G, H$ himself and then runs $A^{H_F}$. Then

$$\Pr [ b = 1 : b \leftarrow A^H ] \leq \Pr [ b = 1 : b \leftarrow A^{H_s} ]$$

with $\varepsilon_2 = \frac{4}{3} s q^4 \lambda^2$

$$= \Pr [ b = 1 : b \leftarrow A^{H_F} ] = \Pr [ b = 1 : b \leftarrow C^F ]$$

$$\leq \Pr [ b = 1 : b \leftarrow C^N ]$$

with $\varepsilon_3 := 2q(1 - \lambda)^{s/2}$

$$= \Pr [ b = 1 : b \leftarrow A^{H_N} ] = \Pr [ b = 1 : b \leftarrow A^G ].$$

Here $\leq \gamma s$ follows from Lemma 38 with $\gamma := (1 - \lambda)^s$ and $A := C$.

Thus $\Delta = \Pr [ b = 1 : b \leftarrow A^H ] - \Pr [ b = 1 : b \leftarrow A^G ] \leq \varepsilon_2 + \varepsilon_3$.

We now fix $s := \left( \frac{b q^4}{2s} \ln \frac{b}{2s} \right)^2$ and $\lambda := 2s^{-1} \ln \frac{b}{2s}$. It is then immediate to verify that $\varepsilon_2 \leq \alpha/2$. And

$$\varepsilon_3 = 2q(1 - \lambda)^{s/2} \leq 2q e^{-\lambda s/2} = 2q e^{-\ln \frac{b}{2s}} = \frac{q}{2}.$$

Here $(*)$ uses the fact that $(1 - 1/x)^x$ converges to $e^{-1}$ from below. Thus $\Delta = \varepsilon_2 + \varepsilon_3 \leq \alpha$ as required by the statement of the lemma.

\[\square\]

C Proof of Theorem 6

C.1 Preliminaries

Let $M = |Y|$ and $N = |X|$ and, without loss of generality, let $Y = \{1, \ldots, M\}$ and $X = \{1, \ldots, N\}$. Let $D \subset \{0,1\}^N$ be the set of all $\binom{N}{k}$ $N$-bit strings of Hamming weight $k$. For every $y$, we associate $S_y$ with a string $z_y \in D$ whose $x$-th entry $z_{y,x} := (z_y)_x$ is 1 if and only if $x \in S_y$. This association is one-to-one. The black-box oracles essentially hide an input $z = (z_1, \ldots, z_M) \in D^M$. Let us write $|\Psi (z_y)\rangle$ and $|\Sigma \Psi (z)\rangle$ instead of $|\Psi (y)\rangle$ and $|\Sigma \Psi \rangle$, respectively, to emphasize that these states depend on $z$.

Let $S_L$ denote the symmetric group of a finite set $L$, that is, the group with the permutations of $L$ as elements and the composition as a group operation. For a positive integer $n$, let $S_n$ denote the isomorphism class of the symmetric groups $S_L$ with $|L| = n$. A permutation $\sigma \in S_X$ acts on $z_y \in D$ in a natural way: we define

$$\sigma (z_y) := (z_{y,\sigma^{-1}(1)}, \ldots, z_{y,\sigma^{-1}(N)}),$$

so that $(\sigma (z_y))_{\sigma (x)} = z_{y,x}$ holds. A permutation $\pi \in S_Y$ acts on $z \in D^M$ in the same way: we define $\pi (z) := (z_{\pi^{-1}(1)}, \ldots, z_{\pi^{-1}(M)})$.

Consider a pair $(\sigma, \pi)$, where $\sigma = (\sigma_1, \ldots, \sigma_M) \in S_X^M$ and $\pi \in S_Y$. Let this pair act on $z \in D^M$ by first permuting the entries of $z$ with respect to $\pi$ and then permuting entries within each $(\pi (z))_y$ with respect to $\sigma_y$. Namely, let

$$(\sigma, \pi) : (z_1, \ldots, z_M) \mapsto (\sigma_1 (z_{\pi^{-1}(1)}), \ldots, \sigma_M (z_{\pi^{-1}(M)})).$$

This action defines a (linear) representation of the wreath product $\mathbb{W} := S_X \wr S_Y$.

Definition 44 ([22, Chapter 4]) The wreath product $G \wr S_M$ of groups $G$ and $S_M$ is the group whose elements are $(\sigma, \pi) \in G^M \times S_M$ and whose group operation is

$$(\sigma_1, \ldots, \sigma_M, \pi')((\sigma_1, \ldots, \sigma_M, \pi) := ((\sigma'_1 \sigma_{(\pi')^{-1}(1)}), \ldots, \sigma'_M \sigma_{(\pi')^{-1}(M)}), \pi' \pi).$$
Let $X_2$ be the set of all $\binom{N}{2}$ size-two subsets of $X$. In addition to (14), we are also interested in the following two representations of $\mathcal{W}$ defined by its action on the sets $Y \times X$ and $Y \times X_2$, respectively:

$$
(\sigma, \pi) : (y, x) \mapsto (\pi(y), \sigma_{\pi(y)}(x)), \quad (15)
$$

$$
(\sigma, \pi) : (y, \{x_1, x_2\}) \mapsto (\pi(y), \{\sigma_{\pi(y)}(x_1), \sigma_{\pi(y)}(x_2)\}). \quad (16)
$$

The former representation concerns oracle queries and the latter—the output of the algorithm.

For $w = (y, x) \in Y \times X$, let $z_w = z_{y,x}$. Note that the representations (14) and (15) are such that, for $\tau \in \mathcal{W}$, we have $(\tau(z))_{\tau(w)} = z_w$.

C.2 Registers and symmetrization of the algorithm

Let $\mathcal{H}_A$ be the workspace on which $\mathcal{A}$ operates. We express

$$\mathcal{H}_A = \mathcal{H}_Q \otimes \mathcal{H}_B \otimes \mathcal{H}_O \otimes \mathcal{H}_R \otimes \mathcal{H}_W, \quad (17)$$

where the tensor factors are defined as follows.

- $\mathcal{H}_Q := \mathcal{H}_{Q_Y} \otimes \mathcal{H}_{Q_X}$ and $\mathcal{H}_B$ are the “query” registers that the oracles $\mathcal{O}_V$ and $\mathcal{O}_F$ use, where $\mathcal{H}_{Q_Y}$, $\mathcal{H}_{Q_X}$, and $\mathcal{H}_B$ correspond to the sets $Y$, $X$, and $\{0,1\}$, respectively. For all $(y, x, b) \in Y \times X \times \{0,1\}$, we have

$$\mathcal{O}_V|y, x, b \rangle := |y, x, b \oplus z_{y,x}\rangle \quad (18)$$

and $\mathcal{O}_F$ maps $|y, \Psi(y), b\rangle$ to $-|y, \Psi(y), b\rangle$ and, for every $|\Psi^\perp\rangle$ orthogonal to $|\Psi(y)\rangle$, maps $|y, \Psi^\perp, b\rangle$ to itself.

- $\mathcal{H}_O := \mathcal{H}_{O_Y} \otimes \mathcal{H}_{O_X}$ is the “output” register, where $\mathcal{H}_{O_Y}$ and $\mathcal{H}_{O_X}$ correspond to the sets $Y$ and $X$, respectively.

- $\mathcal{H}_R := \bigotimes_{\ell=1}^h \mathcal{H}_{R(\ell)}$ is the (initial) “resource” register, where $\mathcal{H}_{R(\ell)} = \mathcal{H}_{R_Y(\ell)} \otimes \mathcal{H}_{R_X(\ell)}$, in which $\mathcal{H}_{R_Y(\ell)}$ and $\mathcal{H}_{R_X(\ell)}$ correspond to the sets $Y$ and $X$, respectively. At the beginning of the algorithm, the register $\mathcal{H}_R$ is initialized to the resource state

$$|\xi'(z)\rangle := \bigotimes_{\ell=1}^h (\alpha_{\ell,0}|\Sigma\Psi(z)\rangle + \alpha_{\ell,1}|\Sigma\Psi\rangle). \quad (19)$$

Also, let $\mathcal{H}_{R_Y} := \bigotimes_{\ell=1}^h \mathcal{H}_{R_Y(\ell)}$ and $\mathcal{H}_{R_X} := \bigotimes_{\ell=1}^h \mathcal{H}_{R_X(\ell)}$.

- $\mathcal{H}_W$ is the rest of the workspace.

Let us also define $\mathcal{H}_{A-Q}$, $\mathcal{H}_{A-O}$, and $\mathcal{H}_{A-R}$ to be the space corresponding to all the registers of the algorithm except $\mathcal{H}_Q$, $\mathcal{H}_O$, and $\mathcal{H}_R$, respectively. Let $I$ be the identity operator. We frequently write subscripts below states and unitary transformations to clarify, respectively, which registers they belong to or act on. For example, we may write $|\xi'(z)\rangle_R$ instead of $|\xi'(z)\rangle$. We do this especially when the order of registers is not that of (17). We may also concatenate subscripts when we use multiple registers at once. For example, we may write $I_Q \otimes I_B$ instead of $I_Q \otimes I_B$.

Let $|\xi_0(z)\rangle_A := |\xi'(z)\rangle_R \otimes |\xi''\rangle_{A-R}$ be the initial state of the algorithm, where $|\xi''\rangle_{A-R}$ is independent from $z$. The algorithm makes in total $q_T := q_V + q_F$ oracle calls. For $q \in \{0,1,\ldots,q_T - 1\}$, let

$$|\xi_q(z)\rangle_A = \sum_{w \in Y \times X} |w\rangle_Q |\xi_{q,w}(z)\rangle_{A-Q}.$$
be the state of the algorithm \( \mathcal{A} \), as a sequence of transformations on \( \mathcal{H}_A \), just before \((q+1)\)-th oracle call, \( \mathcal{O}_V \) or \( \mathcal{O}_F \), where \( |\xi_{q,w}(z)\rangle_{A-Q} \) are unnormalized. Similarly, for \( q=q_T \), let

\[
|\xi_{q_T}(z)\rangle_A = \sum_{w \in Y \times X_2} |w\rangle_{O} |\xi_{q_T,w}(z)\rangle_{A-O}
\]

be the final state of the algorithm.

Let \( U_{\mathcal{I}} \), and \( U_{\mathcal{Q}} \), and \( U_{\mathcal{O}} \) be unitary transformations corresponding to representations (13), (15), and (16) of \( \mathcal{W} \), respectively, where the register \( \mathcal{H}_S \) is yet to be defined. (That is, \( U_{\mathcal{I}}, U_{\mathcal{Q}}, U_{\mathcal{O}} \) are actually families of unitaries, indexed by elements \( \tau \in \mathcal{W} \).) We add a subscript \( \tau \in \mathcal{W} \) when we want to specify that we are considering the representation of the element \( \tau \), for example, we may write \( U_{\mathcal{Q},\tau} \). Since \( \mathcal{H}_R \) is essentially the \( h \)-th tensor power of \( \mathcal{H}_Q \), we define \( U_R := U_{\mathcal{Q},h} \). The tensor product of two (or more) representations of \( \mathcal{W} \) is also a representation of \( \mathcal{W} \). Let \( U_{\mathcal{I}Q} := U_{\mathcal{I}} \otimes U_{\mathcal{Q}} \) and \( U_{\mathcal{IO}} := U_{\mathcal{I}} \otimes U_{\mathcal{O}} \), and we later use analogous notation for other “concatenations”.

We first “symmetrize” \( \mathcal{A} \) by adding an extra register \( \mathcal{H}_S \) holding a “permutation” \( \tau \in \mathcal{W} \). Initially, \( \mathcal{H}_S \) holds a uniform superposition over all permutations:

\[
|\mathcal{W}\rangle_S := \frac{1}{\sqrt{M!(N)!^M}} \sum_{\tau \in \mathcal{W}} |\tau\rangle_S.
\]

Then, at specific points in the algorithm, we insert unitary transformations controlled by the content \( \tau \) of \( \mathcal{H}_S \).

1. At the beginning of the algorithm, we insert the controlled transformation \( U_{R,\tau} \) on the register \( \mathcal{H}_R \). Recall that, if (and only if) \( z_{y,x} = 1 \), then \((\tau(z))_{\tau(y,x)} = 1 \). Hence,

\[
\sum_{\tau \in \mathcal{W}} |\tau\rangle_S |\xi(z)\rangle_A \xrightarrow{\tau \text{ on } \mathcal{H}_R} \sum_{\tau \in \mathcal{W}} |\tau\rangle_S |\xi(\tau(z))\rangle_A.
\]

2. Before each oracle call, \( \mathcal{O}_V \) or \( \mathcal{O}_F \), we insert the controlled transformation \( U^{-1}_{Q,\tau} \) on the register \( \mathcal{H}_Q \). Note that \( (\tau(z))_{y,x} = 1 \) if and only if \( z_{\tau^{-1}(y,x)} = 1 \), and \( \mathcal{O}_V \) and \( \mathcal{O}_F \) use \( z \) as the input. After the oracle call, we insert the controlled \( U_{Q,\tau} \).

3. At the end of the algorithm, we insert the controlled transformation \( U^{-1}_{O,\tau} \) on the register \( \mathcal{H}_O \) containing the output of \( \mathcal{A} \) because, again, \( z_{\tau^{-1}(y,x)} = 1 \) if and only if \( (\tau(z))_{y,x} = 1 \).

The effect of the symmetrization is that, on the subspace \( |\tau\rangle_S \), the algorithm is effectively running on the input \( \tau(z) \). If the original algorithm \( \mathcal{A} \) succeeds on every input \( z \) with average success probability \( p \), the symmetrized algorithm succeeds on every input with success probability \( p \).

Next, we recast \( \mathcal{A} \) into a different form, using an “input” register \( \mathcal{H}_I \) that stores \( z \in D^M \). Namely, let \( \mathcal{H}_I := \bigotimes_{y=1}^M \mathcal{H}(y) \) be an \( \binom{N}{k}^M \)-dimensional Hilbert space whose basis states correspond to possible inputs \( z \), where we define \( \mathcal{H}(y) \) to be \( \binom{N}{k} \)-dimensional Hilbert space whose basis states correspond to \( z_y \in D \). Since all the spaces \( \mathcal{H}(y) \) are essentially equivalent, we write \( \mathcal{H}_I \) instead of \( \mathcal{H}(y) \) when we do not care which particular \( y \in Y \) we are talking about, and \( \mathcal{H}_I = \mathcal{H}_I^{\otimes M} \).

Initially, \( \mathcal{H}_I \) is in the uniform superposition of all the basis states of \( \mathcal{H}_I \). More precisely, \( \mathcal{H}_I \otimes \mathcal{H}_S \otimes \mathcal{H}_A \) takes the following initial state (before applying the controlled transformation \( U_{R,\tau} \) in step 1 of the symmetrisation above):

\[
\left(\binom{N}{k}\right)^{-M/2} \sum_{z \in D^M} |z\rangle_I \otimes |\mathcal{W}\rangle_S \otimes |\xi_0(z)\rangle_A.
\]
We transform the symmetrised version of $A$ into a sequence of transformations on a Hilbert space $\mathcal{H} = \mathcal{H}_I \otimes \mathcal{H}_S \otimes \mathcal{H}_A$. A black-box transformation $O$ (where $O = O_V$ or $O = O_F$) is replaced by a transformation $O' = \sum_{z \in D^M} |z\rangle \langle z| \otimes O(z)$, where $O(z)$ is the transformation $O$ for the case when the input is equal to $z$.

At the end, the algorithm measures the input register $\mathcal{H}_I$ and the output register $\mathcal{H}_O = \mathcal{H}_{O_1} \otimes \mathcal{H}_{O_2}$ in the computational basis, and outputs the result of this measurement: $z \in D^M$, $y \in Y$, and $\{x_1, x_2\} \in X_2$. The algorithm is successful if $z_{y, x_1} = z_{y, x_1} = 1$.

For $q \in \{0, \ldots, q_T - 1\}$, let $|\phi_q^-\rangle$ be the state of the algorithm just before the controlled $U^{-1}_{q,\tau}$ transformation preceding the $(q + 1)$-th oracle call, and let $|\phi_q\rangle$ be the state just after we apply this $U^{-1}_{q,\tau}$ and still before the oracle call. Due to the symmetrization, we have

$$|\phi_q^-\rangle = \gamma \sum_{z \in D^m} |z\rangle \sum_{\tau \in \mathcal{W}} |\tau\rangle_S \sum_{w \in Y \times X} |w\rangle_Q |\xi_{q,w}(\tau(z))\rangle_{A-Q},$$

where $\gamma = 1/\sqrt{M! (N!)^M}$, and, after we apply $U^{-1}_{q,\tau}$, we have

$$|\phi_q\rangle = \gamma \sum_{z \in D^m} |z\rangle \sum_{\tau \in \mathcal{W}} |\tau\rangle_S \sum_{w \in Y \times X} |\tau^{-1}(w)\rangle_Q |\xi_{q,w}(\tau(z))\rangle_{A-Q}. \tag{20}$$

Recall the representations $U_I$ and $U_Q$ of $\mathcal{W}$. Let us also consider the right regular representation of $\mathcal{W}$ acting on $\mathcal{H}_S$: for $\kappa \in \mathcal{W}$, let $U_{S,\kappa}|\tau\rangle := |\tau \kappa^{-1}\rangle$. Let $U_{IS} := U_I \otimes U_S \otimes U_Q$, and, for all $\kappa \in \mathcal{W}$, we have

$$(U_{IS}, \kappa \otimes I_{A-Q})|\phi_q\rangle = \gamma \sum_{z \in D^m} |\kappa(z)\rangle \sum_{\tau \in \mathcal{W}} |\tau \kappa^{-1}\rangle_S \sum_{w \in Y \times X} |\kappa \tau^{-1}(w)\rangle_Q |\xi_{q,w}(\tau(z))\rangle_{A-Q}$$

$$= \gamma \sum_{z \in D^m} |\kappa(z)\rangle \sum_{\tau \in \mathcal{W}} |\tau \kappa^{-1}\rangle_S \sum_{w \in Y \times X} |(\tau \kappa^{-1})^{-1}(w)\rangle_Q |\xi_{q,w}((\tau \kappa^{-1})(\tau(z)))\rangle_{A-Q} = |\phi_q\rangle. \tag{21}$$

For $q \in \{0, 1, \ldots, q_T - 1\}$, let $\rho_q'$ be the density matrix obtained from $|\phi_q\rangle \langle \phi_q|$ by tracing out the $\mathcal{H}_S$ and $\mathcal{H}_A-Q$ registers and, in turn, let $\rho_q$ be obtained from $\rho_q'$ by tracing out the register $\mathcal{H}_Q$. Due to (21), we have

$$U_{I,Q,\tau} \rho_q' U_{I,Q,\tau}^{-1} = \rho_q' \quad \text{and} \quad U_{I,\tau} \rho_q U_{I,\tau}^{-1} = \rho_q \quad \text{for all } \tau \in \mathcal{W}. \tag{22}$$

Similarly, for $q = q_T$, let $|\phi_{q_T}\rangle$ be the final state of the algorithm (i.e., the state after the controlled $U^{-1}_{O,1}$), and it satisfies an analogous symmetry to (21): for all $\kappa \in \mathcal{W}$, we have

$$(U_{IS}, \kappa \otimes I_{A-O})|\phi_{q_T}\rangle = |\phi_{q_T}\rangle.$$ Let $\rho_{q_T}'$ be the density matrix obtained from $|\phi_{q_T}\rangle \langle \phi_{q_T}|$ by tracing out all the registers but $\mathcal{H}_I$ and $\mathcal{H}_O$, let $\rho_{q_T}$ be obtained from $\rho_{q_T}'$ by tracing out the register $\mathcal{H}_I$. Again, we have

$$U_{I,O,\tau} \rho_{q_T}' U_{I,O,\tau}^{-1} = \rho_{q_T}' \quad \text{and} \quad U_{I,\tau} \rho_{q_T} U_{I,\tau}^{-1} = \rho_{q_T} \quad \text{for all } \tau \in \mathcal{W}. \tag{23}$$

Note that, throughout the algorithm, the density matrix of the $\mathcal{H}_I$ part of the state of the algorithm can be affected only by oracle calls. Therefore, for $q \in \{0, 1, \ldots, q_T\}$, this density matrix equals $\rho_q$ just after $q$-th oracle call (at the very beginning of the algorithm, if $q = 0$) and remains such till $(q + 1)$-th oracle call (till the end of the algorithm, if $q = q_T$).

### C.3 Representation theory of $S_X$

Consider a positive integer $n$. The representation theory of $S_n$ is closely related to partitions. A partition $\lambda$ of $n$ is a non-increasing list $(\lambda_1, \ldots, \lambda_k)$ of positive integers satisfying $\lambda_1 + \cdots + \lambda_k = n$. There is one-to-one correspondence between irreducible representations (irreps, for short) of $S_n$, and partitions $\lambda \vdash n$, and we will use these terms interchangeably. For example, $(n)$ corresponds
to the trivial representation and \((1^n) = (1, 1, \ldots, 1)\) to the sign representation. (One may refer to [27] for more background on the representation theory of finite groups and to [22, 26] for the representation theory of the symmetric group and the wreath product.)

The group action of \(S_N\) on \(H_I\) is given by (13), which defines a representation \(U_I\) of \(S_N\) (this representation is independent from \(y\)). In order to decompose \(U_I\) into a direct sum of irreps of \(S_N\) (recall that \(X = \{1, \ldots, N\}\)), first consider the subgroup \(S_k \times S_{N-k}\) of \(S_N\), where \(S_k\) permutes \(\{1, \ldots, k\}\) and \(S_{N-k}\) permutes \(\{k+1, \ldots, N\}\). Let \(V_{I,\sigma}\) be \(U_{I,\sigma}\) restricted to \(\sigma \in S_k \times S_{N-k}\) and the one-dimensional space \(\{1^k 0^{N-k}\}\). \(V_I\) is a representation of \(S_k \times S_{N-k}\) and, since it acts trivially on \(1^k 0^{N-k}\), we have \(V_I \cong (k) \times (N-k)\). And, since

\[
|S_N|/|S_k \times S_{N-k}| = |D|/|\{1^k 0^{N-k}\}|
\]

\(U_I\) is equal to the induced representation when we induce \(V_I\) from \(S_k \times S_{N-k}\) to \(S_N\). For shortness, we write \(U_I = V_I \uparrow S_N\). The Littlewood-Richardson rule then implies

\[
((k) \times (N-k)) \uparrow S_N = (N) \oplus (N-1, 1) \oplus (N-2, 2) \oplus \ldots \oplus (N-k, k).
\]

Hence, we have

\[
H_I = \bigoplus_{i=0}^{k} H_i^{(N-i, i)}
\]

where \(U_I\) restricted to \(H_i^{(N-i, i)}\) is an irrep of \(S_N\) corresponding to the partition \((N-i, i)\) of \(N\).

It is also known (see [18, 24]) that \(H_i^{(N-i, i)} = T_i \cap (T_i^{-1})\), where \(T_i\) is the space spanned by all \(\binom{N}{i}\) states

\[
|\psi_{x_1, \ldots, x_i}\rangle = \frac{1}{\sqrt{\binom{N-i}{k-i}}} \sum_{y \in D} z_y \otimes \prod_{y, x_i = 1}^N |y\rangle
\]

(the value of \(y\) is irrelevant here). When \(i = 0\), let us denote this state by \(|\psi_0\rangle\).

C.4 Framework for the proof

We use the representation-theoretic framework developed in [11] (and used in [8] and [2]). Let

\[
H_{I,a} := T_I = H_I^{(N)} \oplus H_I^{(N-1, 1)}, \quad H_{I,b} := H_I \cap (H_{I,a})^\perp, \quad H_{I,a} := H_I \cap (H_{I,a})^\perp, \quad H_{I,b} := H_I \cap (H_{I,a})^\perp.
\]

And let \(\Pi_{I,a}, \Pi_{I,b}, \Pi_{I,a},\) and \(\Pi_{I,b}\) denote the projections to the spaces \(H_{I,a}, H_{I,b}, H_{I,a},\) and \(H_{I,b}\), respectively.

Recall that \(\rho_q\) is the density matrix of the \(H_I\) part of the state of the algorithm anywhere between \(q\)-th and \((q+1)\)-th oracle calls (interpreting \((-1)\)-st and \((q+1)\)-th oracle calls as the beginning and the end of the algorithm, respectively). Recall that \(\rho_q\) is fixed under the action of \(W\)—for all \(\tau \in W\), we have \(U_{\tau, \rho_q} U_{\tau}^{-1} = \rho_q\)—and so are \(\Pi_{I,a}\) and \(\Pi_{I,b}\). Let

\[
p_{a,q} := \text{Tr}(\rho_q \Pi_{I,a}) \quad \text{and} \quad p_{b,q} := 1 - p_{a,q} = \text{Tr}(\rho_q \Pi_{I,b}).
\]

**Theorem 6** (Hardness of the two values problem) then follows from the following three lemmas.

**Lemma 45** The success probability of the algorithm is at most \(2(k-1)/N \cdot \sqrt{2p_{b,q}}\).

**Lemma 46** (At the very beginning of the algorithm) we have \(p_{b,0} < h^2/(2M)\).

**Lemma 47** For all \(q \in \{0, \ldots, q_T - 1\}\), we have \(|p_{b,q} - p_{b,q+1}| = O(\max\{\sqrt{k/N}, \sqrt{1/k}\})\).

One can see that \(M\), the size of the set \(Y\), does not appear in the statements of Lemmas 45 and 47. The size of \(Y\) indeed does not matter for them, as in we will eventually reduce the general case for Lemmas 45 and 47 to the case when \(|Y| = 1\).
C.5 Proof of Lemma 46

Let us rewrite (19) as

\[ |\zeta'(z)_R| = \bigotimes_{\ell=1}^{h} \left( \frac{1}{\sqrt{M}} \sum_{y \in Y} |y\rangle_{R_Y(\ell)} \langle \alpha_{\ell,0} | \Psi(z_g) \rangle + \langle \alpha_{\ell,1} | \Phi \rangle_{R_X(\ell)} \right) \]

\[ = \frac{1}{\sqrt{M^h}} \sum_{y_1, \ldots, y_h \in Y} |y_1, \ldots, y_h\rangle_{R_Y} |\zeta'(y_1, \ldots, y_h)\rangle_{R_X}, \]

where \(|\Phi\rangle := \sum_{x \in X} |x\rangle / \sqrt{|X|}\) and

\[ |\zeta'(y_1, \ldots, y_h)\rangle_{R_X} := \bigotimes_{\ell=1}^{h} \langle \alpha_{\ell,0} | \Psi(z_g) \rangle_{R_X(\ell)} + \langle \alpha_{\ell,1} | \Phi \rangle_{R_X(\ell)} \]

has unit norm for \langle \Psi(z_g) | \Phi \rangle = \langle \Sigma | \Psi(z) \rangle | \Sigma \Phi \rangle = \sqrt{k/N}. \]

Let \(Y_h\) be the set of all \((y_1, \ldots, y_h) \in Y^h\) such that \(y_\ell \neq y_{\ell'}\) whenever \(\ell \neq \ell'\). Let us write \(|\zeta'(z)_R| = |\zeta'_\alpha(z)_R| + |\zeta'_\beta(z)_R|\), where the unnormalized state \(|\zeta'_{\alpha_\ell}(z)\rangle_R\) corresponds to all \((y_1, \ldots, y_h) \in Y_h\) in the register \(H_{R_Y}\). Then, \(|||\zeta'_\beta(z)\rangle||^2\) equals the probability that among \(h\) numbers chosen independently and uniformly at randomly from \(\{1, \ldots, M\}\) at least two numbers are equal. Analysis of the birthday problem tells us that this probability is at most \(h(h-1)/(2M)\) \[23\]. For \(c \in \{a, b\}, \]

\[ |\phi_c\rangle := \binom{N}{k}^{-M/2} \sum_{z \in D^m} |z\rangle_{W_S} |\Psi(z)_{R_Y(S)}\rangle \left| \zeta'_{\alpha,\beta}(z) \right\rangle_{R_X} \right|_{A-R}, \]

and note that \(|||\phi_c|| = |||\zeta'_c(z)||. \)

The initial state of the algorithm is \(|\phi_a\rangle + |\phi_b\rangle\). (Note: in this proof, the subscript of \(\phi\) does not denote the number of queries.)

**Claim 1** We have \((\Pi_{\mathcal{I}_a} \otimes 1_{SA})|\phi_a\rangle = |\phi_a\rangle\).

Claim 1 implies that \((\Pi_{\mathcal{I}_b} \otimes 1_{SA})|\phi_a\rangle = 0, \) and therefore

\[ p_{b,0} = \text{Tr}(\rho_0 \Pi_{\mathcal{I}_b}) = \text{Tr}(\text{Tr}_{SA}((|\phi_a\rangle + |\phi_b\rangle)(\langle \phi_a | + \langle \phi_b |)) \Pi_{\mathcal{I}_b}) \]

\[ = \langle \phi_b | (\Pi_{\mathcal{I}_b} \otimes 1_{SA})|\phi_b\rangle \leq \langle \phi_b | \phi_b \rangle < h^2/(2M). \]

**Proof of Claim 1.** First, let \(|\Omega_0(z_{y_1})\rangle := |\Psi(z_{y_1})\rangle\) and \(|\Omega_1(z_{y_1})\rangle := |\Phi\rangle\), so that

\[ |\zeta'(y_1, \ldots, y_h)\rangle_{R_X} = \sum_{\beta=(\beta_1, \ldots, \beta_h) \in \{0,1\}^h} \langle \alpha_{\beta_1} \ldots \alpha_{\beta_h} | \Omega_{\beta_1}(z_{y_1}), \ldots, \Omega_{\beta_h}(z_{y_h}) \rangle_{R_X}. \]

For all \(\beta \in \{0,1\}^h\) and all \((y_1, \ldots, y_h) \in Y_h\), let

\[ |\phi_{a,\beta}(y_1, \ldots, y_h)\rangle := \gamma \sum_{z \in D^m} |z\rangle_{W_S} |\Omega_{\beta_1}(z_{y_1}), \ldots, \Omega_{\beta_h}(z_{y_h})\rangle_{R_X} \left| \zeta''_{A-R} \right\rangle_{A-R}, \]

where \(\gamma = \binom{N}{k}^{-M/2} (\alpha_{1,\beta_1} \ldots \alpha_{h,\beta_h}) / \sqrt{M^h}\). We have

\[ |\phi_a\rangle = \sum_{\beta \in \{0,1\}^h} \sum_{(y_1, \ldots, y_h) \in Y_h} |\phi_{a,\beta}(y_1, \ldots, y_h)\rangle, \]

and it is enough to show that

\[ (\Pi_{\mathcal{I}_a} \otimes 1_{SA})|\phi_{a,\beta}(y_1, \ldots, y_h)\rangle = |\phi_{a,\beta}(y_1, \ldots, y_h)\rangle \]

39
for all $\beta \in \{0, 1\}^h$ and $(y_1, \ldots, y_h) \in Y_h$.

Notice that, if $\beta_i = 1$, then the register $R_{X,i}$ contains the state $|\Phi\rangle$ and this register is not entangled with any the other registers. Therefore, it suffices to consider the case when $\beta = 0^h$. Without loss of generality, let $(y_1, \ldots, y_h) = (1, \ldots, h)$.

For simplicity, let $|\hat{\phi}\rangle$ be the the state $|\phi_{a,0^h}(1, \ldots, h)\rangle/\gamma$ restricted to registers $H_I$ and $H_{X,R}$, for these registers are not entangled with the other registers and we have

$$\text{Tr}_{SA}(|\phi_{a,0^h}(1, \ldots, h)\rangle\langle\phi_{a,0^h}(1, \ldots, h)|) = \gamma^2 \text{Tr}_{R_X}(|\hat{\phi}\rangle \langle\hat{\phi}|).$$

We have

$$|\hat{\phi}\rangle = \sum_{z \in D^m} |z\rangle_I |\Psi(z_1), \ldots, \Psi(z_h)\rangle_{R_X} = \bigotimes_{y=1}^h \left( \sum_{z_y \in D} |z_y\rangle_{I(y)} |\Psi(z_y)\rangle_{R_X(y)} \right) \otimes \bigotimes_{y=h+1}^M \left( \sum_{z_y \in D} |z_y\rangle_{I(y)} \right).$$

Recall the states $|\psi_{x_1, \ldots, x_t}\rangle \in H_I$ from \cite{25}. We have

$$\sum_{z_y \in D} |z_y\rangle |\Psi(z_y)\rangle \propto \sum_{z_y \in D} \sum_{x \in X} |x\rangle = \sum_{x \in X} \left( \sum_{z_y \in D} |z_y\rangle \right) |x\rangle \propto \sum_{x \in X} |\psi_x\rangle |x\rangle \in T_{I(y)} \otimes H_{R_X(y)};$$

$$\sum_{z_y \in D} |z_y\rangle \propto |\psi_y\rangle \in T_{I(y)}^0 = H_{I(y)}^{(N)}.$$ 

The claim follows from the definition of $H_{I,a}$ (Section C.4).

C.6 Proof of Lemma 45

Reduction to the $p_{qr,b} = 0$ case. Let us first reduce the lemma to its special case when $p_{qr,b} = 0$. This reduction was used in \cite{1} for a very similar problem. Recall that the final state of the algorithm $|\phi_{qr}\rangle$ satisfies the symmetry $(U_{SO,\tau} \otimes I_{A-O})|\phi_{qr}\rangle = |\phi_{qr}\rangle$ for all $\tau \in \mathbb{W}$, and note that, for $c \in \{a, b\}$, the state

$$|\phi^c_{qr}\rangle := \frac{(1 - p_{qr}) |\phi_{qr}\rangle}{\| (1 - p_{qr}) |\phi_{qr}\rangle \|} = \frac{1}{\sqrt{1 - p_{qr}}} (1 - p_{qr}) |\phi_{qr}\rangle$$

satisfies the same symmetry. We have

$$|\phi_{qr}\rangle = \sqrt{1 - p_{qr}} |\phi^a_{qr}\rangle + \sqrt{p_{qr}} |\phi^b_{qr}\rangle.$$ 

Since $|\phi^a_{qr}\rangle$ and $|\phi^b_{qr}\rangle$ are orthogonal, we have

$$\| |\phi_{qr}\rangle - |\phi^a_{qr}\rangle \| = \sqrt{(1 - \sqrt{1 - p_{qr}})^2 + (\sqrt{p_{qr}})^2} \leq \sqrt{2 p_{qr}}.$$ 

From now on, let us assume that $p_{qr} = 0$ and, thus, $|\psi_{qr}\rangle = |\psi^a_{qr}\rangle$. \cite{37} and (27) states that this changes the success probability at most $\sqrt{2 p_{qr}}$.

Reduction to the $|Y| = 1$ case. Recall that $\rho''_{qr} = \text{Tr}_{S,A-O} |\phi_{qr}\rangle \langle\phi_{qr}|$, and we have

$$(\Pi_{I,a} \otimes I_{O}) \rho''_{qr} = \rho''_{qr} \quad \text{and} \quad \forall \tau \in \mathbb{W}: U_{SO,\tau} \rho''_{qr} U_{SO,\tau}^{-1} = \rho''_{qr}.$$ 

The algorithm makes its final measurement of the $H_I$ and $H_O$ registers, ignoring all the other registers, therefore the success probability is completely determined by $\rho''_{qr}$. Let us assume that the algorithm measures (and then discards) the $H_O$ register first, before measuring $H_I$ and $H_{O,X_2}$, and that the outcome of this measurement is $y \in Y$. Due to the symmetry, we get each outcome $y$ with the same probability $1/M$. 

40
Now the algorithm can discard the registers $H_{I(y')}$ for all $y' \neq y$, as their content do not affect the success probability. We are left with
\[
\rho''_{q_I,y} = M \text{Tr}_{I(y')} : y' \neq y \left( |H_{O_{X_2}} \rangle \langle y|_{O_Y} \rho''_{q_I} (|H_{O_{X_2}} \rangle \langle y|_{O_Y}) \right),
\]
which is a density matrix on the registers $H_{I(y)}$ and $H_{O_{X_2}}$, and it satisfies
\[
(\Pi_{I,a} \otimes I_{O_{X_2}})\rho''_{q_I,y} = \rho''_{q_I,y},
\]
\[
\forall \sigma \in \mathcal{S}_X : (U_{I,\sigma} \otimes U_{O_{X_2,\sigma}})\rho''_{q_I,y} (U_{I,\sigma} \otimes U_{O_{X_2,\sigma}})^{-1} = \rho''_{q_I,y}
\]
(we use the subscript $I$ instead of $I(y)$ as $y$ is fixed from now on). The success probability of the algorithm equals the probability that we measure the state $\rho''_{q_I,y}$ in the computational basis and obtain $z_y \in D$ and $\{x_1, x_2\} \in X_2$ such that $z_{y,x_1} = z_{y,x_2} = 1$. Hence, we have reduced the proof to the case when $|Y| = 1$.

**The $|Y| = 1$ case.** Since $y \in Y$ is fixed, to lighten the notation, in the remainder of the proof of Lemma 4.5 let us write $z'$ instead of $z_y$ and $z'_z$ instead of $z_{y,z}$.

Let us now assume that the algorithm measures the $H_{O_{X_2}}$ register, obtaining $\{x_1, x_2\} \in X_2$, and only then measures $H_I$. Due to the symmetry, the measurement yields each outcome $\{x_1, x_2\}$ with the same probability $1/(N^2)$, and let
\[
\hat{\rho} := \left( \frac{N}{2} \right) (I_I \otimes \langle \{x_1, x_2\}|_{O_{X_2}} \rho''_{q_I,y} (I_I \otimes \langle \{x_1, x_2\})_{O_{X_2}})
\]
be the density matrix of the register $H_I$ after the measurement. Without loss of generality, let $\{x_1, x_2\} = \{1, 2\}$, and let $\hat{\mathcal{S}} := \mathcal{S}_{\{1,2\}} \times \mathcal{S}_{\{3,...,N\}} < \mathcal{S}_X$ be the group of all permutations $\sigma \in \mathcal{S}_X$ that map $\{1, 2\}$ to itself. Now we have
\[
\Pi_{I,a} \hat{\rho} = \hat{\rho} \quad \text{and} \quad \forall \sigma \in \hat{\mathcal{S}} : U_{I,\sigma} \hat{\rho} U_{I,\sigma}^{-1} = \hat{\rho}.
\]
Let $\hat{\Pi}$ denote the projection to the subspace of $H_I$ spanned by all $|z'\rangle$ such that $z'_1 = z'_2 = 1$. We note that $U_{I,\sigma} \hat{\Pi} U_{I,\sigma}^{-1} = \hat{\Pi}$ for all $\sigma \in \hat{\mathcal{S}}$. One can see that the success probability of the algorithm is $\text{Tr}(\hat{\Pi} \hat{\rho})$, and it is left to show

**Claim 2** $\text{Tr}(\hat{\Pi} \hat{\rho}) \leq 2(k-1)/(N-1)$.

**Proof.** We can express $\hat{\rho}$ as a mixture of its eigenvectors $|\chi_i\rangle$, with probabilities that are equal to their eigenvalues $\chi_i$:
\[
\hat{\rho} = \sum_i \chi_i |\chi_i\rangle \langle \chi_i|.
\]
Hence we have
\[
\text{Tr}(\hat{\Pi} \hat{\rho}) = \sum_i \chi_i \text{Tr}(\hat{\Pi} |\chi_i\rangle \langle \chi_i|) = \sum_i \chi_i \|\hat{\Pi} |\chi_i\rangle\|^2,
\]
which is at most
\[
\max_{|\chi|} \left( \|\hat{\Pi} |\chi\rangle\|^2/\| |\chi\rangle\|^2 \right)
\]
where the maximization is over all eigenvectors of $\hat{\rho}$ with non-zero eigenvalues. Due to the symmetry (28), we can calculate the eigenspaces of $\hat{\rho}$ by inspecting the restriction of $U_I$ to the subspace $T_I^1$, namely, $\hat{U}_I := \Pi_{I,a} U_I$. Recall that we defined $T_I^1$ to be the space spanned by all $|\psi_x\rangle = \frac{1}{\sqrt{k-1}} \sum_{z' \in D} |z'\rangle$. We note that $\langle \psi_{x_1} | \psi_{x_2}\rangle = \frac{k-1}{N-1}$ for all $x_1, x_2 : x_1 \neq x_2$. 


Both $U_I$ and $\hat{U}_I$ are representations of both $S_X$ and its subgroup $\hat{S}$. We already studied $U_I$ as a representation of $S_X$ in Section C.3. Since $T_I = \mathcal{H}_I^{(N)} \oplus \mathcal{H}_I^{(N-1, 1)}$, the representation $\hat{U}_I$ of $S_X$ consists of only two irreps: one-dimensional $(N)$ and $(N - 1)$-dimensional $(N - 1, 1)$, which correspond to the spaces $\mathcal{H}_I^{(N)}$ and $\mathcal{H}_I^{(N-1, 1)}$, respectively.

In order to see how $\hat{U}_I$ decomposes into irreps of $\hat{S}$, we need to restrict $(N)$ and $(N - 1, 1)$ from $S_N$ to $S_2 \times S_{N-2}$. The Littlewood-Richardson rule gives us the decomposition of these restrictions:

$$(N) \downarrow (S_2 \times S_{N-2}) = (2) \times (N - 2));$$

$$(N - 1, 1) \downarrow (S_2 \times S_{N-2}) = (2) \times (N - 2)) \oplus ((1, 1) \times (N - 2)) \oplus (2) \times (N - 3, 1)).$$

Hence, Schur’s lemma and (28) imply that that eigenspaces of $\hat{\rho}$ are invariant under $U_{I, \sigma}$ for all $\sigma \in \hat{S}$, and they have one of the following forms:

1. one-dimensional subspace spanned by $|\psi(\alpha, \beta)\rangle = \alpha(|\psi_1\rangle + |\psi_2\rangle) + \beta \sum_{x=3}^N |\psi_x\rangle$ for some coefficients $\alpha, \beta$;

2. one-dimensional subspace spanned by $|\psi_1\rangle - |\psi_2\rangle$;

3. $(N - 3)$-dimensional subspace consisting of all $\sum_{i=3}^N \alpha_x |\psi_x\rangle$ with $\sum_x \alpha_x = 0$ (spanned by all $|\psi_x\rangle - |\psi_{x'}\rangle$, $x, x' \in \{3, \ldots, N\}$);

4. a direct sum of subspaces of the above form.

In the first case,

$$\hat{\Pi}|\psi(\alpha, \beta)\rangle = \frac{2\alpha + (k - 2)\beta}{\sqrt{\binom{N-1}{k-1}}} \sum_{\substack{z_1', \ldots, z_N' \in \{0, 1\} \left. \sum_{i=3}^N z_i' = k - 2 \}} |1, 1, z_3', \ldots, z_N\rangle.$$

Therefore,

$$||\hat{\Pi}|\psi(\alpha, \beta)\rangle||^2 = \frac{(N-2)}{(N-1)} \frac{2\alpha + (k - 2)\beta)}{2k-2} = \frac{k-1}{k-1} \frac{2\alpha + (k - 2)\beta)}{2k-2}.$$

We also have

$$|||\psi(\alpha, \beta)\rangle||^2 = \langle \psi(\alpha, \beta) | \psi(\alpha, \beta) \rangle$$

$$= \left(2 + \frac{k-1}{N-1}\right) |\alpha|^2 + \left(N - 2 + (N - 2)(N - 3) \frac{k-1}{N-1}\right) |\beta|^2 + 4(N - 2)\alpha\beta$$

$$\geq \frac{(2\alpha + (k - 2)\beta)^2}{2}$$

with the last inequality following by showing that coefficients of $|\alpha|^2$, $|\beta|^2$ and $\alpha\beta$ on the left hand side are all larger than similar coefficients on the right hand side. Therefore,

$$\frac{||\hat{\Pi}|\psi(\alpha, \beta)\rangle||^2}{|||\psi(\alpha, \beta)\rangle||^2} \leq \frac{2(k-1)}{N-1}.$$
and
\[ \|\hat{\Pi}(|ψ_3⟩ - |ψ_4⟩)\|^2 = 2\binom{N-4}{k-3} \binom{N-1}{k-1} = \frac{(k-2)(k-3)(N-k)}{(N-2)(N-3)(N-4)}. \]

We also have
\[ \|ψ_3⟩ - |ψ_4⟩\|^2 = 2 - ⟨ψ_3|ψ_4⟩ = 2 - \frac{k-1}{N-1} = \frac{2N-k}{N-1}. \]

Hence,
\[ \frac{\|\hat{\Pi}(|ψ_3⟩ - |ψ_4⟩)\|^2}{\|ψ_3⟩ - |ψ_4⟩\|^2} = \frac{2(k-2)(k-3)(N-1)}{(N-2)(N-3)(N-4)} = O\left(\frac{k^2}{N^2}\right). \]

C.7 Reduction of Lemma 47 to the \(|Y| = 1\) case

First, instead of the oracle \(O_V\) given by \((18)\), we define
\[ O_V(z)|y, x, b⟩ := (-1)^{b_{x,y}}|y, x, b⟩. \]

Both definitions are equivalently powerful as one is obtained from another by two Hadamard gates on the register \(H_B\).

For all \(z_y \in D\), let
\[ O'_V(z_y) := I_{Q_X} - 2 \sum_{x \in X} |x⟩⟨x|_{Q_X} \quad \text{and} \quad O'_F(z_y) := I_{Q_X} - 2|Ψ(z_y)⟩⟨Ψ(z_y)|_{Q_X} \]

act on \(H_{Q_X}\), so that we have
\[ O'_V = \sum_{z \in D^M} |z⟩⟨z|_I \otimes \sum_{y \in Y} |y⟩⟨y|_{Q_Y} \otimes O'_V(z_y) ⊗ |1⟩⟨1|_B + I_{Q} \otimes |0⟩⟨0|_B, \]
\[ O'_F = \sum_{z \in D^M} |z⟩⟨z|_I \otimes \sum_{y \in Y} |y⟩⟨y|_{Q_Y} \otimes O'_F(z_y) ⊗ I_B. \]  

Let
\[ ρ'' = ρ'_{q,00} ⊗ |0⟩⟨1|_B + ρ'_{q,01} ⊗ |0⟩⟨0|_B + ρ'_{q,10} ⊗ |1⟩⟨0|_B + ρ'_{q,11} ⊗ |0⟩⟨1|_B \]
be the state of the algorithm corresponding to the \(H_I, H_Q, \) and \(H_B\) registers right before the \((q + 1)\)-th oracle call \((O_V \lor O_F)\). Note that \(ρ_q = Tr_{QB}(ρ''')\) and, since oracles are the only gates of the algorithm that interact with the \(H_I\) register, \(ρ_{q+1} = Tr_{QB}(O'′′′ρ''′′O')\).

Notice that \(|p_{a,q} - p_{a,q+1}| = |p_{a,q} - p_{a,q+1}|\), therefore let us deal with \(p_{a,q}\) instead. We have
\[ |p_{a,q} - p_{a,q+1}| = Tr((Π_{I,a}ρ_q - ρ_{q+1})) = Tr((Π_{I,a} ⊗ I_Q)(ρ'' - O'′′ρ''′O')) \]
which for the oracle \(O_V\) equals
\[ Tr((Π_{I,a} ⊗ I_Q)(ρ'_q,11 - O'_Vρ'_q,11O'_V)), \]

where \(O'_V = (I_{Q} \otimes |1⟩⟨1|_B)O'_V(I_{Q} \otimes |1⟩⟨1|_B). \) Therefore, without loss of generality, we assume that the state of \(H_B\) is always \(|1⟩\) throughout the execution of the algorithm. In turn, we assume that \(O'_V\) and \(O'_F\) in \((29)\) act only on \(H_I \otimes H_Q, \) and we take \(ρ_q\) instead of \(ρ''\) and \(I_{Q}\) instead of \(I_{Q_B}\) in \((30)\).

Since \(τ(z)I_{Q_Y} = 1\) if and only if \(z_{x,y} = 1\), we have \(U_{Q,τO′U_Q}^{-1} = O'\) for all \(τ \in W,\) and recall that the same symmetry holds for \(ρ_q,\) namely, \((22)\). Hence, for all \(y \in Y,\)
\[ ρ'_{q,y} = M(I_{Q_X} \otimes |y⟩⟨y|_{Q_Y})ρ'_q(I_{Q_X} \otimes |y⟩⟨y|_{Q_Y}). \]
has trace one and \( (30) \) equals

\[
M \text{Tr}(\mathbb{I}_{QX} \otimes |y\rangle\langle Q, y|)(\Pi_{I,a} \otimes \mathbb{I}_Q)(\rho'_q - \mathcal{O}' \rho'_q \mathcal{O}')(\mathbb{I}_{QX} \otimes |y\rangle\langle Q, y|))
\]

\[
= \text{Tr}\left((\Pi_{I,a} \otimes \mathbb{I}_{QX})\left(\rho'_{q,y} - \left( \sum_{z \in D^M} |z\rangle\langle z| \otimes \mathcal{O}'(z_y)\right)\rho'_{q,y}\left( \sum_{z \in D^M} |z\rangle\langle z| \otimes \mathcal{O}'(z_y)\right)\right)\right). \tag{31}
\]

Without loss of generality, let \( y = 1 \), and let us write

\[
\sum_{z \in D^M} |z\rangle\langle z| = \sum_{z \in D} |z_1\rangle\langle z_1| \otimes \mathbb{I}^{(M-1)}.
\]

Recall that \( \Pi_{I,a} = \Pi^{(M)}_{I,a} \). Therefore, for

\[
\hat{\rho}'_{q,1} := \text{Tr}_{I(2),\ldots,I(M)}((\mathbb{I}_{I(1)} \otimes \Pi^{(M-1)}_{I,a} \otimes \mathbb{I}_{QX})\rho'_{q,y}),
\]

\( (30) \) and \( (31) \) are equal to

\[
\text{Tr}\left((\Pi_{I(1),a} \otimes \mathbb{I}_{QX})\left(\rho'_{q,1} - \left( \sum_{z \in D} |z_1\rangle\langle z_1| \otimes \mathcal{O}'(z_1)\right)\rho'_{q,1}\left( \sum_{z \in D} |z_1\rangle\langle z_1| \otimes \mathcal{O}'(z_1)\right)\right)\right). \tag{32}
\]

Since \( \hat{\rho}'_{q,1} \) is a positive semidefinite operator of trace at most one and it acts on \( \mathcal{H}_{I(1)} \otimes \mathcal{H}_{QX} \), we have reduced the lemma to the case when \( |Y| = 1 \). We consider this case in Section D.

**D Proof of Lemma 47 when \( |Y| = 1 \)**

Since \( |Y| = \{y\} \), let us use notation \( \mathcal{H}_Q \) instead of \( \mathcal{H}_{Q_X} \) to denote the register corresponding to the query index \( x \in X \). Also, now we have \( z = (z_y) \), so let us use \( z \) instead of \( z_y \) and \( z_x \) instead of \( z_y \). Also, now we denote the permutations in \( S_N \) with \( \pi \) instead of \( \sigma \).

We will consider the following representations of \( S_N \):

1. The computational basis of \( \mathcal{H}_Q \) is labeled by \( x \in \{1, \ldots, N\} = X \). We define the action of \( \pi \in S_N \) on \( \mathcal{H}_Q \) via the unitary \( U_{Q,\pi}|x\rangle := |\pi(x)\rangle \). \( \mathcal{H}_Q \) is known as the natural representation of \( S_N \), and we can decompose \( \mathcal{H}_Q = \mathcal{H}_Q^{(N)} \oplus \mathcal{H}_Q^{(N-1,1)} \) so that \( U_{Q} \) restricted to \( \mathcal{H}_Q^{(N)} \) and \( \mathcal{H}_Q^{(N-1,1)} \) are irreps of \( S_N \) isomorphic to \( (N) \) and \( (N-1,1) \), respectively.

2. The computational basis of \( \mathcal{H}_I \) is labeled by \( z \in D \), that is, \( z = (z_1, \ldots, z_N) \in \{0, 1\}^N \) such that \( \sum_{i=1}^N z_i = k \). In Section C.3 we already defined and studied the representation \( U_I \) for \( \pi \in S_N \),

\[
U_{I,\pi}|z_1 \ldots z_N\rangle = U_I,\pi|z_{\pi^{-1}(1)} \ldots z_{\pi^{-1}(N)}\rangle.
\]

We showed that we can decompose \( \mathcal{H}_I = \bigoplus_{i=0}^k \mathcal{H}_I^{(N-i,i)} \) so that \( U_I \) restricted to \( \mathcal{H}_I^{(N-i,i)} \) is an irrep of \( S_N \) isomorphic to \( (N-i,i) \).

3. Finally, let \( U := U_{Q} \otimes U_I \), which acts on \( \mathcal{H} := \mathcal{H}_Q \otimes \mathcal{H}_I \) and is also a representation of \( S_N \).

Let \( \Pi_Q^{(N)} \) and \( \Pi_Q^{(N-1,1)} \) denote, respectively, the projectors on \( \mathcal{H}_Q^{(N)} \) and \( \mathcal{H}_Q^{(N-1,1)} \). \( \Pi_Q^{(N)} \) is the \( N \)-dimensional matrix with all entries equal to \( 1/N \), and \( \Pi_Q^{(N-1,1)} \) is the \( N \)-dimensional matrix with \( 1-1/N \) on the diagonal and \( -1/N \) elsewhere. Let \( \Pi_I^{(N)}, \Pi_I^{(N-1,1)}, \ldots, \Pi_I^{(N-k,k)} \) denote, respectively, the projectors on \( \mathcal{H}_I^{(N)}, \mathcal{H}_I^{(N-1,1)}, \ldots, \mathcal{H}_I^{(N-k,k)} \). The entries of these \( \binom{N}{k} \)-dimensional matrices can be calculated using the fact that they project on the eigenspaces of the Johnson scheme (see [18]).
Let us also denote
\[
\Pi_{H_Q \otimes S_{z2}} := \mathbb{I}_Q \otimes \sum_{j=2}^k \Pi_I^{(N-j,j)} = (\Pi_Q^{(N)} + \Pi_Q^{(N-1,1)}) \otimes \sum_{j=2}^k \Pi_I^{(N-j,j)},
\]
\[
\Pi_{H_Q \otimes S_{z2}} := \mathbb{I}_Q - \Pi_{H_Q \otimes S_{z2}} = (\Pi_Q^{(N)} + \Pi_Q^{(N-1,1)}) \otimes (\Pi_I^{(N)} + \Pi_I^{(N-1,1)}),
\]
which are equal to \(\mathbb{I}_Q \otimes \Pi_{I,b}\) and \(\mathbb{I}_Q \otimes \Pi_{I,a}\), respectively.

D.1 Statement of the lemma

For the oracles, let us write \(\mathcal{O}\) instead of \(\mathcal{O}'\) (where \(\mathcal{O} = \mathcal{O}_V\) or \(\mathcal{O} = \mathcal{O}_F\)). Similarly to (29), we have to consider
\[
\mathcal{O}_V = \sum_{z \in D} \left( \sum_{x \in X_{z,0}} |x\rangle\langle x| - \sum_{x \in X_{z,1}} |x\rangle\langle x| \right)_Q \otimes |z\rangle\langle z|_I,
\]
\[
\mathcal{O}_F = \sum_{z \in D} \left( |\Psi(z)\rangle\langle \Psi(z)| \right)_Q \otimes |z\rangle\langle z|_I,
\]
where \(|\Psi(z)\rangle = \sum_{x: z_x = 1} |x\rangle / \sqrt{k}\). Note that \(\mathcal{O}\) acts on \(\mathcal{H}\) and is satisfies \(U_\pi \mathcal{O} U_\pi^{-1} = \mathcal{O}\) for all \(\pi \in S_N\). Equivalently to (32), it suffices to prove that
\[
|\text{Tr}(\Pi_{H_Q \otimes S_{z2}} (\rho - \mathcal{O}_\rho \mathcal{O}))| \leq O(\max\{\sqrt{k/N}, 1/k\})
\]
for every density operator \(\rho\) on \(\mathcal{H}\) that satisfies \(U_\pi \rho U_\pi^{-1} = \rho\) for all \(\pi \in S_N\) and both oracles \(\mathcal{O} = \mathcal{O}_V\) and \(\mathcal{O} = \mathcal{O}_F\).

For a subspace \(\mathcal{H}' \subset \mathcal{H}\) such that \(\mathcal{H}'\) is invariant under \(U\) (i.e., under \(U_\pi\) for all \(\pi \in S_N\)), let \(U|_{\mathcal{H}'}\) be \(U\) restricted to this subspace (note: \(U|_{\mathcal{H}'}\) is a representation of \(S_N\)). Let \(\Pi_{\mathcal{H}'}\) denote the projector on \(\mathcal{H}'\). Due to Schur’s lemma, there is a spectral decomposition
\[
\rho = \sum_\mu \chi_\mu \frac{\Pi_\mu}{\dim \mu},
\]
where \(\sum_\mu \chi_\mu = 1\), every \(\mu\) is invariant under \(U\), and \(U|_\mu\) in an irrep of \(S_N\). Hence, it suffices to show the following.

**Lemma 48** For every subspace \(\mu \subset \mathcal{H}\) such that \(U|_\mu\) is an irrep and for \(\mu'\) being the subspace that \(\mu\) is mapped to by \(\mathcal{O}_V\) or \(\mathcal{O}_F\), we have
\[
\frac{1}{\dim \mu} |\text{Tr}(\Pi_{H_Q \otimes S_{z2}} (\Pi_\mu - \Pi_{\mu'}))| \leq O(\max\{\sqrt{k/N}, 1/k\}).
\]

In order to prove Lemma 48, we need to inspect the representation \(U\) in more detail.

D.2 Decomposition of \(U\)

Let us decompose \(U\) into irreps. We consider two approaches how to do that. That is, the list of irreps contained in \(U\) cannot depend on which approach we take, but we can choose the way we address individual instances of irreps. For example, we will show that \(U\) contains four instances of \((N-1,1)\), and we have as much freedom in choosing a projector on a single instance of \((N-1,1)\) as in choosing (up to global phase) a unit vector in \(\mathbb{C}^4\).

For an irrep \(\theta\) present in \(U\), let \(\Pi_\theta\) be a projector on the space corresponding to all instances of \(\theta\) in \(U\).
We can see that, for every \(\theta\) which is the projector on the unique instance of \((N-j, j)\), and, similarly, for \(O\) and \(I\) here to specify which spaces these irreps act on, namely, \(H_Q\) and \(H_I\), respectively, but we will drop these subscripts most of the time later). Note that (\(N\) \(\otimes\) (\(N-j, j\)) \(\cong\) (\(N-j, j\)) and \((N-1, 1) \otimes (N) \cong (N-1, 1)\) as \((N)\) is the trivial representation. And, for \(j \in \{1, \ldots, k\}\), the decomposition of \((N-1, 1) \otimes (N-j, j)\) is given by the following claim.

**Claim 3** For \(j \in \{1, \ldots, k\}\), we have

\[
(N-1, 1) \otimes (N-j, j) = (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j-1, j+1) \oplus (N-j-1, j, 1),
\]

where we omit the term \((N-j, j-1, 1)\) when \(j = 1\).

**Proof.** We use Expression 2.9.5 of [22], which, for \(j \in \{2, \ldots, k\}\), gives us

\[
(N-1, 1) \otimes (N-j, j) = (N-j, j) \downarrow (S_{N-1} \times S_1) \uparrow S_N \ominus (N-j, j) \downarrow S_N \uparrow S_N
\]

\[
= ((N-j, j-1) \times (1)) \uparrow S_N \ominus ((N-j-1, j) \times (1)) \uparrow S_N \ominus (N-j, j)
\]

\[
= (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j, j)
\]

\[
\oplus (N-j-1, j+1) \oplus (N-j-1, j, 1) \oplus (N-j, j)
\]

\[
= (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j-1, j+1)
\]

\[
\oplus (N-j-1, j, 1)
\]

and, similarly, for \(j = 1\), gives us

\[
(N-1, 1) \otimes (N-1, 1) = (N-1, 1) \downarrow (S_{N-1} \times S_1) \uparrow S_N \ominus (N-1, 1) \downarrow S_N \uparrow S_N
\]

\[
= (N) \oplus (N-1, 1) \oplus (N-2, 2) \oplus (N-2, 1, 1).
\]

We can see that, for every \(\ell \in \{0, 1\}\) and \(j \in \{0, \ldots, k\}\), the representation \((N-\ell, \ell)Q \otimes (N-j, j)_I\) is *multiplicity-free*, that is, it contains each irrep at most once. For an irrep \(\theta\) present in \((N-\ell, \ell)Q \otimes (N-j, j)_I\), let

\[
\Pi^{(N-\ell, \ell)Q \otimes (N-j, j)_I}_\theta := \Pi\theta(\Pi^{(N-\ell, \ell)Q}_Q \otimes \Pi^{(N-j, j)_I}_I),
\]

which is the projector on the unique instance of \(\theta\) in \((N-\ell, \ell)Q \otimes (N-j, j)_I\). For example, for \(\theta = (N-1, 1)\), we have projectors \(\Pi^{(N-1, 1)Q \otimes (N-1, 1)_I}_{(N-1, 1)}\), \(\Pi^{(N-1, 1)Q \otimes (N-1, 1)_I}_{(N-1, 1)}\), and \(\Pi^{(N-1, 1)Q \otimes (N-2, 2)_I}_{(N-1, 1)}\).

**Approach 2: via spaces invariant under queries \(O_V\) and \(O_F\).** Let us decompose \(H\) as the direct sum of four subspaces, each invariant under the action of \(U\), \(O_V\), and \(O_F\). First, let \(H = H^{(0)} \oplus H^{(1)}\), where \(H^{(0)}\) and \(H^{(1)}\) are spaces corresponding to, respectively, the subsets

\[
H_0 = \{ (x, z) \in X \times D : z_x = 0 \} \quad \text{and} \quad H_1 = \{ (x, z) \in X \times D : z_x = 1 \},
\]

of the standard basis \(X \times D\). Let us further decompose \(H^{(0)}\) and \(H^{(1)}\) as

\[
H^{(0)} = H^{(0, s)} \oplus H^{(0, t)} \quad \text{and} \quad H^{(1)} = H^{(1, s)} \oplus H^{(1, t)},
\]

where

\[
H^{(0, s)} := \text{span}\left\{ \sum_{x: z_x=0} |x, z\rangle : z \in D \right\} \quad \text{and} \quad H^{(1, s)} := \text{span}\left\{ \sum_{x: z_x=1} |x, z\rangle : z \in D \right\},
\]

46
We noted in Section D.2 that $O_{F}$ acts on $\mathcal{H}^{(1,s)}$ as the minus identity and on $\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1,t)}$ as the identity. Meanwhile, $O_{V}$ acts on $\mathcal{H}^{(1)}$ as the minus identity and on $\mathcal{H}^{(0)}$ as the identity.

For every superscript $\sigma \in \{(0), (1), (0, s), (0, t), (1, s), (1, t)\}$, let $\Pi^{\sigma}$ be the projector on the space $\mathcal{H}^{\sigma}$, and let $U^{\sigma}$ be the restriction of $U$ to $\mathcal{H}^{\sigma}$. Let $V^{\sigma}_{\pi}$ be $U^{\sigma}_{\pi}$ restricted to $\pi \in S_{k} \times S_{N-k}$ and the space

$$\tilde{\mathcal{H}}^{\sigma} := \mathcal{H}^{\sigma} \cap (\mathcal{H}_{Q} \otimes |1^{k}0^{N-k}\rangle).$$

$V^{\sigma}$ is a representation of $S_{k} \times S_{N-k}$. One can see that

$$|S_{N}|/|S_{k} \times S_{N-k}| = \dim \mathcal{H}^{\sigma}/\dim \tilde{\mathcal{H}}^{\sigma},$$

so we have $U^{\sigma} = V^{\sigma} \uparrow S_{N}$. In order to see how $U^{\sigma}$ decomposes into irreps, we need to see how $V^{\sigma}$ decomposes into irreps, and then apply the Littlewood-Richardson rule.

We have $\dim \tilde{\mathcal{H}}^{(0,s)} = \dim \tilde{\mathcal{H}}^{(1,s)} = 1$, and it is easy to see that $V^{(0,s)}$ and $V^{(1,s)}$ act trivially on $\mathcal{H}^{(0,s)}$ and $\mathcal{H}^{(1,s)}$, respectively. That is, $V^{(0,s)} \cong V^{(1,s)} \cong (k) \times (N-k)$. Now, note that

$$\tilde{\mathcal{H}}^{(0)} = \text{span} \{ |x\rangle \otimes |1^{k}0^{N-k}\rangle : x \in \{k+1, \ldots, N\} \}.$$

The group $S_{k}$ (in $S_{k} \times S_{N-k}$) acts trivially on $\tilde{\mathcal{H}}^{(0)}$, while and the action of $S_{N-k}$ on $\tilde{\mathcal{H}}^{(0)}$ defines the natural representation of $S_{N-k}$. Hence, $V^{(0)} \cong (k) \times ((N-k) \oplus (N-k-1, 1))$, and $V^{(0)} = V^{(0,s)} \oplus V^{(0,t)}$, in turn, gives us $V^{(0,t)} \cong (k) \times (N-k-1, 1)$. Analogously we obtain $V^{(1,t)} \cong (k-1, 1) \times (N-k)$. The decompositions of $U^{(0,s)} = V^{(0,s)} \uparrow S_{N}$ and $U^{(1,s)} = V^{(1,s)} \uparrow S_{N}$ into irreps are given via \(^{(24)}\). For $U^{(0,t)} = V^{(0,t)} \uparrow S_{N}$ and $U^{(1,t)} = V^{(1,t)} \uparrow S_{N}$, the Littlewood-Richardson rule gives us, respectively,

$$((k) \times (N-k-1, 1)) \uparrow S_{N} = (N-1, 1) \oplus (N-2, 2) \oplus (N-2, 1, 1) \oplus (N-3, 3) \oplus (N-3, 2, 1) \oplus (N-4, 4) \oplus (N-4, 3, 1) \oplus \ldots \oplus (N-k, k) \oplus (N-k, k-1, 1) \oplus (N-k-1, k+1) \oplus (N-k+1, k, 1) \oplus (N-k-1, k, 1)$$

and

$$((k-1, 1) \times (N-k)) \uparrow S_{N} = (N-1, 1) \oplus (N-2, 2) \oplus (N-2, 1, 1) \oplus (N-3, 3) \oplus (N-3, 2, 1) \oplus (N-4, 4) \oplus (N-4, 3, 1) \oplus \ldots \oplus (N-k+1, k-1) \oplus (N-k+1, k-2, 1) \oplus (N-k, k-1, 1).$$

Note that all $U^{(0,s)}$, $U^{(0,t)}$, $U^{(1,s)}$, and $U^{(1,t)}$ are multiplicity-free. For a superscript $\sigma \in \{(0), (0, t), (1, s), (1, t)\}$ and an irrep $\theta$ present in $U^{\sigma}$, let $\Pi^{\sigma}_{\theta} := \Pi^{\sigma}_{\theta} \Pi^{\sigma}$, which is the projector on the unique instance of $\theta$ in $U^{\sigma}$. For example, for $\theta = (N-1, 1)$, we have all the projectors $\Pi^{(0,s)}_{(N-1,1)}$, $\Pi^{(0,t)}_{(N-1,1)}$, $\Pi^{(1,s)}_{(N-1,1)}$, and $\Pi^{(1,t)}_{(N-1,1)}$.

### D.3 Significant irreps

We noted in Section \[\text{[D2]}\] that $O_{F}$ acts on $\mathcal{H}^{(1,s)}$ as the minus identity and on $\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1,t)}$ as the identity and $O_{V}$ acts on $\mathcal{H}^{(1)}$ as the minus identity and on $\mathcal{H}^{(0)}$ as the identity. This means that, if $\mu$ is a subspace of one of the spaces $\mathcal{H}^{(0)}$, $\mathcal{H}^{(1,s)}$, or $\mathcal{H}^{(1,t)}$, then $\mu' = \mu$. In turn, even if that is not the case, we still have that $U|_{\mu}$ and $U|_{\mu'}$ are isomorphic irreps.

Also note that

$$|\text{Tr}(\Pi_{Q} \otimes \mathbb{S}_{>2}(\Pi_{\mu} - \Pi_{\mu'}))| = |\text{Tr}(\Pi_{Q} \otimes \mathbb{S}_{<2}(\Pi_{\mu} - \Pi_{\mu'}))|.$$  

(36)
Hence we need to consider only $\mu$ such that $U|_{\mu}$ is isomorphic to an irrep present in both

\[((N) \oplus (N - 1, 1))_{Q} \otimes ((N) \oplus (N - 1, 1))_{I} \text{ and } ((N) \oplus (N - 1, 1))_{Q} \otimes \bigoplus_{j=2}^{k} (N - j, j)_{I}, \]

as otherwise the expression equals 0. From Section D.2 we see that the only such irreps are $(N - 1, 1)$, $(N - 2, 2)$, and $(N - 2, 1, 1)$.

The representation $U$ contains four instances of irrep $(N - 1, 1)$, four of $(N - 2, 2)$, and two of $(N - 2, 1, 1)$. Projectors on them, according to Approach 1 in Section D.2, are

\begin{align*}
\Pi^{(N)_{Q} \oplus (N - 1, 1)_{I}}, & \Pi^{(N - 1)_{Q} \oplus (N - 1, 1)_{I}}, \Pi^{(N - 1)_{Q} \oplus (N - 2, 2)_{I}}, \Pi^{(N - 1)_{Q} \oplus (N - 3, 3)_{I}}, \\
\Pi^{(N - 2)_{Q} \oplus (N - 2, 2)_{I}}, & \Pi^{(N - 1)_{Q} \oplus (N - 2, 2)_{I}}, \Pi^{(N - 1)_{Q} \oplus (N - 3, 3)_{I}}.
\end{align*}

or, according to Approach 2 in Section D.2, are

\begin{align*}
\Pi^{(0, s)_{(N - 1, 1)}}, & \Pi^{(0, t)_{(N - 1, 1)}}, \Pi^{(1, s)_{(N - 1, 1)}}, \Pi^{(1, t)_{(N - 1, 1)}}, \\
\Pi^{(0, s)_{(N - 2, 2)}}, & \Pi^{(0, t)_{(N - 2, 2)}}, \Pi^{(1, s)_{(N - 2, 2)}}, \Pi^{(1, t)_{(N - 2, 2)}}, \\
\Pi^{(0, t)_{(N - 2, 1, 1)}}, & \Pi^{(1, t)_{(N - 2, 1, 1)}}.
\end{align*}

One thing we can see from this right away is that, if $U|_{\mu} \cong (N - 2, 1, 1)$, then $\mu \subset \mathcal{H}^{(0)} \oplus \mathcal{H}^{(1, t)}$, so the application of the query $\mathcal{O}_{F}$ fixes $\mu$, and the expression equals 0.

### D.4 Necessary and sufficient conditions for irrep $(N - 1, 1)$

We would like to know what are necessary and sufficient conditions for inequality (35) to hold. First, let us consider the irrep $(N - 1, 1)$; later, the argument for the other two irreps will be very similar.

**Transporters as the standard basis for irreps.** For $a_1, a_2 \in \{0, 1\}$ and $b_1, b_2 \in \{s, t\}$, let $\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)}$ be, up to a global phase, the unique operator of rank $\dim(N - 1, 1)$ such that

\[ (U^{(a_1, b_1)}_{(N - 1, 1)})_{\pi} = \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} (U^{(a_2, b_2)}_{(N - 1, 1)})_{\pi} (\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)})^{*}, \]

for all $\pi \in S_{N}$. We call $\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)}$ the transporter from irrep $U^{(a_2, b_2)}_{(N - 1, 1)}$ to $U^{(a_1, b_1)}_{(N - 1, 1)}$. One can see that all non-zero singular values of $\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)}$ are 1. We also have

\[ \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} (\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)})^{*} = \Pi^{(a_1, b_1)}_{(N - 1, 1)}, \quad (\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)})^{*} \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} = \Pi^{(a_2, b_2)}_{(N - 1, 1)}. \]

We can and do choose global phases of these transporters in a consistent manner so that

\[ (\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)})^{*} = \Pi^{(a_2, b_2) \leftarrow (a_1, b_1)}_{(N - 1, 1)} \quad \text{and} \quad \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} = \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} \]

for all $a_3 \in \{0, 1\}$ and $b_3 \in \{s, t\}$. Together they imply $\Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} = \Pi^{(a_1, b_1)}_{(N - 1, 1)}$.

Fix $a_3$ and $b_3$, and note that

\[ (\Pi^{(a_3, b_3) \leftarrow (a_1, b_1)}_{(N - 1, 1)})^{*} \Pi^{(a_3, b_3) \leftarrow (a_2, b_2)}_{(N - 1, 1)} = \Pi^{(a_1, b_1) \leftarrow (a_2, b_2)}_{(N - 1, 1)} \]

is independent of our choice of $(a_3, b_3)$. Therefore, let us introduce the notation

\[ \Pi^{(a_1, b_1)}_{(N - 1, 1)} := \Pi^{(a_3, b_3) \leftarrow (a_1, b_1)}_{(N - 1, 1)}. \]
The norm of the vector $\mu$ is given by

$$\|\mu\| = \sqrt{\text{tr}(\mu^\dagger \mu)}.$$

For $b \in \mathbb{N}$ and note that

$$\mathcal{O}_V = \left\{ \gamma = (\gamma_{0,s}, \gamma_{0,t}, \gamma_{1,s}, \gamma_{1,t}) \mid \gamma_{0,s}, \gamma_{0,t}, \gamma_{1,s}, \gamma_{1,t} \in \mathbb{C} \right\}.$$

The norm of the vector $\gamma$ is 1. The converse also holds: for any unit vector $\gamma$, $\Pi^\dagger \Pi$ is a projector to an irrep isomorphic to $(N-1,1)$.

From now on, let us work in this basis of transporters, because in this basis, queries $\mathcal{O}_V$ and $\mathcal{O}_F$ restricted to $\Pi_{(N-1,1)}$ are, respectively,

$$\mathcal{O}_V|_{(N-1,1)} = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right) \quad \text{and} \quad \mathcal{O}_F|_{(N-1,1)} = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right).$$

**Necessary and sufficient condition for the query $\mathcal{O}_V$.** In the basis of transporters we have

$$\Pi_{(N-1,1)} \bar{\Pi} \Pi_{(N-1,1)} = \left( \begin{array}{cccc} \gamma_{0,s}^2 & \gamma_{0,s} \gamma_{0,t} & \gamma_{0,s} \gamma_{1,s} & \gamma_{0,s} \gamma_{1,t} \\ \gamma_{0,t} \gamma_{0,s} & \gamma_{0,t}^2 & \gamma_{0,t} \gamma_{1,s} & \gamma_{0,t} \gamma_{1,t} \\ \gamma_{1,s} \gamma_{0,s} & \gamma_{1,s} \gamma_{0,t} & \gamma_{1,s}^2 & \gamma_{1,s} \gamma_{1,t} \\ \gamma_{1,t} \gamma_{0,s} & \gamma_{1,t} \gamma_{0,t} & \gamma_{1,t} \gamma_{1,s} & \gamma_{1,t}^2 \end{array} \right),$$

and note that

$$|a,b|^2 = \text{Tr}(\Pi_{(N-1,1)} (a,b)) / \text{dim}(N-1,1).$$

From (37), one can see that

$$\bar{\Pi}_{(N-1,1)} \Pi_{(N-1,1)} = \Pi_{(N-1,1),Q \otimes S_2}.$$

Hence, for the space $\mu$, the desired inequality (35) becomes

$$\frac{1}{\text{dim}(N-1,1)} |\text{Tr}(\Pi_{(N-1,1),Q \otimes S_2} (\Pi_{(N-1,1)} - \Pi_{(N-1,1)})| \leq O(\max\{ \sqrt{k/\sqrt{N}}, \sqrt{1/k} \}).$$

Let us first obtain a necessary condition if we want this to hold for all $\mu$.

In the same transporter basis, let

$$\Pi_{(N-1,1),Q \otimes S_2} = \left( \begin{array}{cccc} |\beta_{0,s}|^2 & \beta_{0,s} \beta_{0,t} & \beta_{0,s} \beta_{1,s} & \beta_{0,s} \beta_{1,t} \\ \beta_{0,t} \beta_{0,s} & |\beta_{0,t}|^2 & \beta_{0,t} \beta_{1,s} & \beta_{0,t} \beta_{1,t} \\ \beta_{1,s} \beta_{0,s} & \beta_{1,s} \beta_{0,t} & |\beta_{1,s}|^2 & \beta_{1,s} \beta_{1,t} \\ \beta_{1,t} \beta_{0,s} & \beta_{1,t} \beta_{0,t} & \beta_{1,t} \beta_{1,s} & |\beta_{1,t}|^2 \end{array} \right).$$

For $b_0, b_1 \in \{ s, t \}$ and a phase $\phi \in \mathbb{R}$, define the space $\xi_{b_0, b_1, \phi}$ via the projector on it:

$$\Pi_{\xi_{b_0, b_1, \phi}} := \frac{1}{2} (\Pi_{(N-1,1)}^{(0, b_0)} + e^{i\phi} \Pi_{(N-1,1)}^{(0, b_0)} \leftrightarrow (1, b_1)) + e^{-i\phi} \Pi_{(N-1,1)}^{(1, b_1)} \leftrightarrow (0, b_0) + \Pi_{(N-1,1)}^{(1, b_1)} \leftrightarrow (0, b_0)).$$

We have

$$\Pi_{\xi_{b_0, b_1, \phi}} \mathcal{O}_V \Pi_{\xi_{b_0, b_1, \phi}} = e^{i\phi} \Pi_{(N-1,1)}^{(0, b_0)} \leftrightarrow (1, b_1) + e^{-i\phi} \Pi_{(N-1,1)}^{(1, b_1)} \leftrightarrow (0, b_0),$$

so, for this space, the inequality (39) becomes

$$|e^{i\phi} \beta_{1,b_1} \beta_{0,b_0} + e^{-i\phi} \beta_{1,b_1} \beta_{0,b_0}| \leq O(\max\{ \sqrt{k/\sqrt{N}}, \sqrt{1/k} \}).$$
Since this has to hold for all \(b_0, b_1\), and \(\phi\) (in particular, consider \(b_0\) and \(b_1\) that maximize \(|\beta_{1,b_0}\beta_{0,b_0}\)|), we must have either
\[
|\beta_{1,s}|^2 + |\beta_{1,t}|^2 \leq O(\max\{k/N, 1/k\}) \quad \text{or} \quad |\beta_{1,s}|^2 + |\beta_{1,t}|^2 \geq 1 - O(\max\{k/N, 1/k\}),
\]
and note that
\[
|\beta_{1,s}|^2 + |\beta_{1,t}|^2 = \text{Tr}(\Pi^{(N-1,1)}_1 \otimes (N-2,2)_t \cdot \Pi^{(1)}_1) / \text{dim}(N-1,1).
\]

The condition (41) is necessary, but it is also sufficient for (39). Because, if it holds, then
\[
|\beta_{1,b_0}\beta_{0,b_0}| \leq O(\max\{\sqrt{k}/N, \sqrt{1/k}\})
\]
for all \(b_0, b_1 \in \{s, t\}\) and, clearly, \(\gamma_{1,b_0}^s \gamma_{0,b_0}^t \in O(1)\) for all unit vectors \(\gamma\). Therefore, if we plug (38) and (40) into (39), the inequality is satisfied.

**Necessary and sufficient condition for the query \(O_F\).** Almost identical analysis shows that, in order for the main conjecture hold when \(U|_{\mu}\) is isomorphic to \((N-1,1)\) and we apply \(O_F\), it is necessary and sufficient that
\[
|\beta_{1,s}|^2 \leq O(\max\{k/N, 1/k\}) \quad \text{or} \quad |\beta_{1,s}|^2 \geq 1 - O(\max\{k/N, 1/k\}).
\]

Note that
\[
|\beta_{1,s}|^2 = \text{Tr}(\Pi^{(N-1,1)}_1 \otimes (N-2,2)_t \cdot \Pi^{(1)}_1) / \text{dim}(N-1,1).
\]

**D.5 Conditions for irreps \((N-2,2)\) and \((N-2,1,1)\)**

For irreps \((N-2,2)\) and \((N-2,1,1)\), let us exploit equation (36). Mainly, we do that because the space \(H_Q \otimes S_{>2}\) contains three instances of irrep \((N-2,2)\), while \(H_Q \otimes S_{<2}\) contains only one. From (37) we get
\[
\hat{\Pi}_{(N-2,2)}^{(N-1,1)} \Pi_{H_Q \otimes S_{<2}}^{(N-1,1)} = \Pi_{(N-2,2)}^{(N-1,1,1)} \cdot \Pi_{(N-1,1)}^{(1)} \quad \text{and} \quad \hat{\Pi}_{(N-2,1,1)}^{(N-1,1)} \Pi_{H_Q \otimes S_{<2}}^{(N-1,1)} = \Pi_{(N-2,1,1)}^{(N-1,1,1)} \cdot \Pi_{(N-1,1)}^{(1)}.
\]

**Condition for the query \(O_V\).** An analysis analogous to that of the irrep \((N-1,1)\) shows that, in order for the desired inequality (35) to hold for query \(O_V\) and irreps \((N-2,2)\) and \((N-2,1,1)\), it is sufficient to have
\[
\frac{\text{Tr}(\Pi^{(N-1,1)}_1 \otimes (N-2,2)_t \cdot \Pi^{(1)}_1)}{\text{dim}(N-2,2)} \leq O(k/N) \quad \text{and} \quad \frac{\text{Tr}(\Pi^{(N-1,1)}_1 \otimes (N-2,1,1)_t \cdot \Pi^{(1)}_1)}{\text{dim}(N-2,1,1)} \leq O(k/N).
\]

Let us prove this. Consider irrep \((N-2,2)\) and the hook-length formula gives us \(\text{dim}(N-2,2) = N(N-3)/2\). We have
\[
\text{Tr}(\Pi^{(N-1,1)}_1 \otimes (N-2,2)_t \cdot \Pi^{(1)}_1) \leq \text{Tr}(\Pi^{(N-1,1)}_Q \otimes \Pi^{(N-1,1)}_I \cdot \Pi^{(1)}_1),
\]
and we can evaluate the right-hand side of this exactly. \(\Pi^{(1)}_1\) is diagonal (in the standard basis), and, on the diagonal, it has \((N-k)\binom{N}{k}\) zeros and \(k\binom{N}{k}\) ones. The diagonal entries of \(\Pi^{(N-1,1)}_Q\) are all the same and equal to \(N-1\). The diagonal entries of \(\Pi^{(N-1,1)}_I\) are also all the same, because \(\Pi^{(N-1,1)}_I\) projects to an eigenspace of the Johnson scheme. More precisely, we have \(\text{Tr}(\Pi^{(N-1,1)}_I) = \text{dim}(N-1,1) = N-1\), therefore the diagonal entries of \(\Pi^{(N-1,1)}_I\) are \((N-1)/\binom{N}{k}\). Hence, the diagonal entries of \(\Pi^{(N-1,1)}_Q \otimes \Pi^{(N-1,1)}_I\) are \((N-1)^2/(N\binom{N}{k})\), implying that
\[
\text{Tr}\left(\Pi^{(N-1,1)}_Q \otimes \Pi^{(N-1,1)}_I \cdot \Pi^{(1)}_1\right) = \frac{k(N-1)^2}{N}
\]
and, in turn,
\[
\frac{\text{Tr}(\Pi^{(N-1,1)}_Q \otimes (N-2,2)_t \cdot \Pi^{(1)}_1)}{\text{dim}(N-2,2)} \leq \frac{2k(N-1)^2}{N^2(N-3)} \in O(k/N)
\]
as required. The same argument works for irrep \((N-2,1,1)\) as, by the hook-length formula, \(\text{dim}(N-2,1,1) = (N-1)(N-2)/2 = \text{dim}(N-2,2) + 1\).
Condition for the query $\mathcal{O}_F$. As we mentioned in the very end of Section D.3, $\mathcal{O}_F$ affects no space $\mu$ such that $U|\mu$ is isomorphic to irrep $(N−2, 1, 1)$. However, the following argument for irrep $(N−2, 2)$ actually works for $(N−2, 1, 1)$ as well. We have

$$\frac{\text{Tr}(I_{(N−1,1)}^{(N−1,1)} \otimes (N−2,2) |^N \Pi^{(1,s)})}{\dim(N−2,2)} \leq \frac{\text{Tr}(I_{(N−1,1)}^{(N−1,1)} \otimes (N−2,2) |^N \Pi^{(1)})}{\dim(N−1,1)} \leq O(k/N),$$

which, similarly to the condition (42) for irrep $(N−1,1)$, is sufficient to show that the main conjecture holds for irrep $(N−2,2)$ and the query $\mathcal{O}_F$.

D.6 Solution for irrep $(N−1,1)$

Recall that conditions (41) and (42) are sufficient for the main conjecture to hold for the queries $\mathcal{O}_V$ and $\mathcal{O}_F$, respectively. Hence, it suffices for us to show that

$$\frac{\text{Tr}(I_{(N−1,1)}^{(N−1,1)} \otimes (N−2,2) |^N \Pi^{(1,s)})}{\dim(N−1,1)} \geq \frac{\text{Tr}(I_{(N−1,1)}^{(N−1,1)} \otimes (N−2,2) |^N \Pi^{(1,s)})}{\dim(N−1,1)} = \frac{k−1}{k} \cdot \frac{N(N−k−1)}{(N−1)(N−2)} \geq 1 − O(\max\{k/N, 1/k\}).$$

It is easy to see that both inequalities in this expression hold, and we need to concern ourselves only with the equality in the middle.

Notice that

$$\Pi^{(N−1,1)}_{(N−1,1)} \otimes (N−2,2) |^N \Pi^{(1,s)}_{(N−1,1)} = (I \otimes \Pi_1^{(N−2,2)}) \cdot \Pi^{(1,s)}_{(N−1,1)},$$

and let us evaluate the trace of the latter. We briefly mentioned before that $\Pi_1^{(N)}, \Pi_1^{(N−1,1)}, \ldots, \Pi_1^{(N−k,k)}$ are orthogonal projectors on the eigenspaces of the Johnson scheme. Let us now use this fact.

Johnson scheme on $\mathcal{H}_I$. For any two strings $z, z′ ∈ D$, let $|z−z'|$ be the half of the Hamming distance between them (the Hamming distance between them is an even number in the range $\{0, 2, 4, \ldots, 2k\}$). For every $i ∈ \{0, 1, \ldots, k\}$, let

$$A_i^I = \sum_{z,z′∈D} |z⟩⟨z′|,$$

which is a 01-matrix in the standard basis of $\mathcal{H}_I$. Matrices $A_0^I, A_1^I, \ldots, A_k^I$ form an association scheme known as the Johnson scheme (see [13], Chapter 7).

There are matrices $C_0^I, C_1^I, \ldots, C_k^I$ of the same dimensions as $A_i$ that satisfy

$$C_j^I = \sum_{i=0}^{k−j} \binom{k−i}{j} A_i \quad \text{for all } j \quad \text{and} \quad A_i^I = \sum_{j=k−i}^{k} (-1)^{j−k+i} \binom{j}{k−i} C_j \quad \text{for all } i. \quad (43)$$

These matrices $C_j^I$ simplify the calculation of the eigenvalues of $A_i^I$, as, for all $j ∈ \{0, 1, \ldots, k\}$, we have

$$C_j^I = \sum_{h=0}^{j} \binom{j−h−i}{h} \binom{k−h}{j−h} \Pi_1^{(N−h,h)} \quad \text{for all } j. \quad (44)$$

Hence, we can express $A_i^I$ uniquely as a linear combination of orthogonal projectors $\Pi_1^{(N−h,h)}$, and the coefficients corresponding to these projectors are the eigenvalues of $A_i^I$. 

51
Here, however, we are interested in the opposite: expressing $\Pi_I^{(N-h,h)}$ as a linear combination of $A_I^i$. From (44) one can see that

$$\Pi_I^{(N-h,h)} = (N - 2h + 1) \sum_{j=0}^{h} (-1)^{j-h} \frac{(k-j)}{(N-1-j)} C_j^I$$

(45)

for $h = 0, 1, 2$. We are interested particularly in $\Pi_I^{(N-2,2)}$, and from (45) and (43) we get

$$\Pi_I^{(N-2,2)} = \frac{1}{(N-2)} \sum_{i=0}^{k} \left( \frac{k-i}{2} - \frac{(k-1)^2}{N-2} (k-i) + \frac{k^2(k-1)^2}{2(N-1)(N-2)} \right) A_I^i.$$ 

(46)

**Johnson scheme on $H^{(1,s)}$.** Recall that, for $z \in D$, we have $|\Psi(z)| = \sum_{x:z_x = 1} |x|/\sqrt{k}$, and let us define

$$A_I^{(1,s)} = \sum_{z,z' \in D} |\Psi(z), z\rangle \langle \Psi(z'), z'|$$

for all $i \in \{0, 1, \ldots, k\}$. The matrices $A_I^i$ and $A_I^{(1,s)}$ have the same eigenvalues corresponding to the same irreps. Analogously to the space $H_I$, we can define matrices $C_I^{(1,s)}$ to the space $H^{(1,s)}$. From (44) and (43) we get

$$\Pi_I^{(1,s), (N-1,1)} = \frac{1}{N} \sum_{i=0}^{k} \left( \frac{k-i}{2} - \frac{k^2}{N} \right) A_I^{(1,s)}.$$ 

(47)

**Both Johnson schemes together.** Now that we have expressions for both $\Pi_I^{(N-2,2)}$ and $\Pi_I^{(1,s)}$, we can compute $\text{Tr}(I_Q \otimes \Pi_I^{(N-2,2)}) \cdot \Pi_I^{(1,s)}$. For all $i, i' \in \{0, 1, \ldots, k\}$, we have

$$\text{Tr}(I_Q \otimes A_I^i) \cdot A_I^{(1,s)} = \delta_{i,i'} \left( \binom{N}{k} \binom{k}{i} \left( \binom{N-k}{i} \frac{k-i}{k} \right) \right).$$

(48)

Indeed, it is easy to see that this trace is 0 if $i \neq i'$, and for $i = i'$ we argue as follows. The matrix $A_I^i$ has $(N)^i$ rows, and each row has $(N)^{N-k}$ entries 1. That is, each $z \in D$ has exactly $(N)^{i}$ entries 1. For all $i', z' \in D$ such that $|z - z'| = i$. And for such $z$ and $z'$, we have $\langle \psi_z | \psi_{z'} \rangle = (k-i)/k$.

Now, if we put (40), (47), and (48) together, we get

$$\text{Tr}(I_Q \otimes \Pi_I^{(N-1,1)}) \cdot \Pi_I^{(1,s)} = \text{Tr}(I_Q \otimes \Pi_I^{(N-2,2)}) \cdot \Pi_I^{(1,s)} = \sum_{i=0}^{k} \left( \binom{k-i}{2} - \frac{(k-1)^2}{N-2} (k-i) + \frac{k^2(k-1)^2}{2(N-1)(N-2)} \right) \left( \binom{N}{k} \binom{k}{i} \left( \binom{N-k}{i} \frac{k-i}{k} \right) \right),$$

which, by using the equality

$$\sum_{i=1}^{k} \left( \binom{k}{i} \frac{(k-i)!}{(k-i-l)!} \right) = \frac{k!}{(k-l)!} \binom{N-l}{N-k},$$

can be shown to be equal to $\frac{k}{k} \frac{N(N-k-1)}{(N-2)}$. We get the desired equality by dividing this by $\text{dim}(N-1,1) = N-1$.

This concludes the proof of Theorem 6 (Hardness of the two values problem).
E Proofs for Section 4

E.1 Proof of Theorem 7

Proof of Theorem 7. Algorithm $E_1$ measures the first half of $|ΣΨ⟩$. This measurement yields a uniformly random outcome $y ∈ Y$ and leaves $|Ψ(y)⟩$ in the second half.

Let $O_F(y) := I − 2|Ψ(y)⟩⟨Ψ(y)|$. This notation is justified because $O_F(y)$ is how $O_F$ operates its the second input when the first input is $|y⟩$. In particular, given $O_F$ we can implement the unitary $O_F(y)$.

The algorithm $E_2$ is as follows:

```plaintext
1 initialize register $X$ with $|Ψ(y)⟩$ (given as input);
2 for $i = 1$ to $n + 1$ do
3     for $j = 1, \ldots, \lceil \log(π/2\sqrt{δ_{min}}) \rceil$ do
4         for $k = 1$ to $2^{j−1}$ do
5             let $U_P|x⟩ := (-1)^P(x)|x⟩$;
6             apply $O_F(y)U_P$ to register $X$
7         let $P_X := \sum_{P(x)=1} |x⟩⟨x|$;
8         measure register $X$ with projector $P_X$, outcome $b$;
9         if $b = 1$ then
10             measure register $X$ in the computational basis, outcome $x$;
11             return $x$
```

We first analyze the one iteration of the $j$-loop (i.e., lines 1–11). Let $P_y := \{x ∈ S_y : P(x) = 1\}$ and $P_y := \{x ∈ S_y : P(x) = 0\}$. Let $|yes⟩ := \sum_{x ∈ P_y} \sqrt{1/|P_y|} |x⟩$ and $|no⟩ := \sum_{x ∈ P_y} \sqrt{1/|P_y|} |x⟩$. For any $β ∈ ℝ$, let $|φ_β⟩ := \sin β|yes⟩ + \cos β|no⟩$. We check that $U_P|φ_β⟩ = |φ_−β⟩$. Let $γ := \arcsin \sqrt{|P_y|/|S_y|}$. Then $|Ψ(x)⟩ = \sin γ|yes⟩ + \cos γ|no⟩ = |φ_γ⟩$. Hence $O_F(y)|φ_β⟩ = (I − 2|Ψ(x)⟩⟨Ψ(x)|)|φ_β⟩ = |φ_β + 2γ⟩$ for all $β$. Thus $O_F(y)U_P|φ_β⟩ = |φ_β + 2γ⟩$.

Assume that at line 4 we have $X = |φ_β⟩$. The innermost loop (lines 3–6) thus yields $X = |φ_β + 2γ⟩$. Since $|yes⟩ ∈ \im P_X$ and $|no⟩$ is orthogonal to $\im P_X$, measuring $X$ using $P_X$ yields $b = 1$ with probability $(\sin(β + 2jγ))^2$. If $b = 1$, $X$ has state $|yes⟩$, and if $b = 0$, $X$ has state $|no⟩$. Thus, if $b = 1$, measuring $X$ in the computational basis (line 10) yields and returns $x ∈ S_y$ with $P(x) = 1$.

Summarizing so far: one iteration of the $j$-loop (i.e., lines 1–11) returns $x ∈ S_y$ with probability $(\sin(β + 2jγ))^2$ if $X$ has state $|φ_β⟩$. And if no such $x$ is returned, $X$ is in state $|no⟩ = |ψ_y⟩$.

In the first execution of the $j$-loop, $X$ contains $|Ψ(y)⟩ = |φ_γ⟩$. Thus in all further executions of the $j$-loop, $X$ contains $|no⟩ = |φ_0⟩$ and the probability of returning $x ∈ S_y$, $P(x) = 1$ in the $j$-th iteration is $(\sin 2jγ)^2 = 1 − (\sin(π/2 − 2jγ))^2 ≥ 1 − (π/2 − 2jγ)^2$.

Thus any but the first iteration of the $j$-loop (i.e., lines 3–11) fails to return $x ∈ S_y$ with probability at most:

$$χ := \min_{1 \leq j \leq \lceil \log(π/2\sqrt{δ_{min}}) \rceil} (\pi/2 − 2jγ)^2.$$ 

We distinguish two cases:

- Case $γ > \frac{π}{4}$: Since also $γ ≤ 1$, we have that $|π/2 − 2γ| ≤ 2 − π/2 < \frac{π}{2}$ and thus $χ ≤ (π/2 − 2γ)^2 ≤ (\frac{π}{2})^2 = \frac{π}{4}$.
- Case $γ ≤ \frac{π}{4}$: For at least one $1 \leq j ≤ \lceil \log(π/2\sqrt{δ_{min}}) \rceil$ we have $2jγ ≤ π/2$. And for at least one such $j$ we have

$$2jγ ≥ 2\log_{\pi/2\sqrt{δ_{min}}} γ = \frac{\pi γ}{2\sqrt{δ_{min}}} ≥ \frac{π \arcsin \sqrt{|P_y|/|S_y|}}{2\sqrt{|P_y|/|S_y|}} ≥ \frac{π}{2}.$$
Thus the minimum ranges over some $j, j+1$ such that $2^j \gamma \leq \pi/2 \leq 2^{j+1} \gamma$. For any $a \geq 0$, \(\min \{|x_2|-a, |\frac{\pi}{2} - 2a|\} \leq \frac{\pi}{6}\) if $a \leq \frac{\pi}{2} \leq 2a$. Thus $\chi \leq (\pi/6)^3 \leq \frac{1}{2}$.

Hence in all cases, $\chi \leq \frac{1}{2}$.

The algorithm executes the $j$-loop $n+1$ times, and each but the first $j$-loop fails to return $x \in S_y$, $P(x) = 1$ with probability at most $\chi \leq \frac{1}{2}$. Thus the algorithm fails to return $x \in S_y$, $P(x) = 1$ with probability at most $\chi^n \leq 2^{-n}$. \hfill $\square$

E.2 Proof of [Corollary 9]

Proof of [Corollary 9] We first show (i). Let $A$.

Proof of Corollary 9.

In the remainder of the proof, we will make the probabilistic choice of oracles explicit, as well as their use by $A$. That is, $P_A$ becomes:

$$P_A = \Pr[w = w_0 : w_0 \xleftarrow{\$} \{0,1\}^{\ell_{\text{rand}}}, (S_{\text{com}}) \leftarrow \$, O_S \leftarrow \$, O_P \leftarrow \$, w \leftarrow A^{O_E,O_P,O_R,O_S,O_F,O_Q,O_V}]$$

Here we used the following shorthands: $(S_{\text{com}}) \leftarrow \$ means that the sets $S_{\text{com}}$ are uniformly random subsets of $\{0,1\}^{\ell_{\text{ch}}} \times \{0,1\}^{\ell_{\text{resp}}}$ of size $k$. $O_S \leftarrow \$ means that the oracle $O_S$ is randomly chosen as described in Definition 8 [Oracle distribution]. $O_P \leftarrow \$ means that the oracle $O_P$ is randomly chosen as described in Definition 8. Since no random choices are involved in the definitions of $O_E, O_R, O_F, O_Q, O_V$, we do not write their definitions explicitly here, cf. Definitions 5 and 8.

Removing $O_P, O_R$: We now remove access to $O_P, O_R$. We then have

$$P_A \leq 2(q_P + q_R + 1)^\sqrt{P_1},$$

$$P_1 := \Pr[w = w_0 : w_0 \xleftarrow{\$} \{0,1\}^{\ell_{\text{rand}}}, (S_{\text{com}}) \leftarrow \$, O_S \leftarrow \$, w \leftarrow A^{O_E,O_S,O_F,O_Q,O_V}]$$

for some $A_1$ by Lemma 39 (with $O_1 := (O_P, O_R)$, $w := w_0$, $O_2 := (O_E, O_S, O_F, O_Q, O_V)$, $\forall w' : f(\cdot, w') := f(w', \cdot, \cdot, \cdot) := w'$). Here the algorithm $A_1$ makes at most as many oracle queries as $A$ to the remaining oracles. Note that we also removed $O_P \leftarrow \$ because $O_P$ is not used any more.

Removing $O_E$: We now transform $A_1$ not to output $w$, but to output the two accepting conversations $(\text{com}, ch, resp, ch', resp')$ needed for extraction. In the following, we write short Collision for $(ch, resp) \neq (ch', resp') \land (ch, resp), (ch', resp') \in S_{\text{com}}$.

$$P_1 \leq 2q_E \sqrt{P_2} + 2^{-\ell_{\text{rand}}},$$

$$P_2 := \Pr[\text{Collision} : (S_{\text{com}}) \leftarrow \$, O_S \leftarrow \$, (\text{com}, ch, resp, ch', resp') \leftarrow A^{O_S,O_F,O_Q,O_V}]$$

for some $A_2$ by Lemma 40 (with $w := w_0$, $\ell := \ell_{\text{rand}}$, $O_1 := O_E$, $O_2 := (O_S, O_F, O_Q, O_V)$, and $X := \{(\text{com}, ch, resp, ch', resp') : \text{Collision}\}$). Here $A_2$ makes at most as many oracles queries as $A_1$. We also removed the choice of $w_0$ from the formula because none of the remaining oracles depend on it.

Removing $O_Q$: Fix integers $n, m$. We determinate the actual values later. By Theorem 3 [Emulating state creation oracles], we have:

$$P_2 \leq P_3 + O\left(\frac{q_E}{\sqrt{n}} + \frac{q_E}{\sqrt{m}}\right),$$

$$P_3 := \Pr[\text{Collision} : (S_{\text{com}}) \leftarrow \$, O_S \leftarrow \$, (\text{com}, ch, resp, ch', resp') \leftarrow A^{O_S,O_F,O_V}(\{R\})]$$

for some $A_3$. Here $A_3$ makes $q_S, q_F, q_V$ queries to $O_S, O_F, O_V$. And $|R := |\Sigma\Psi \otimes m \otimes |\alpha_1 \otimes \cdots \otimes |\alpha_n\rangle$ with $|\alpha_j := (\cos \frac{j\pi}{2n})|\Sigma\Psi \rangle + (\sin \frac{j\pi}{2n})|\perp\rangle$.  

54
Removing $O_S$: For given choice of $(S_{com})_{com \in \{0,1\}^{\ell_{com}}}$, let $D_Y$ be the distribution of $O_S(z)$, i.e., $D$ picks com $\overset{\$}{\leftarrow} \{0,1\}^{\ell_{com}}$ and $(ch, resp) \overset{\cdot}{\leftarrow} S_{com}$ and returns $(com, ch, resp)$.

Fix some $\alpha \in (0, 1)$ (we determine the value of $\alpha$ later). Then, for fixed choice of $(S_{com})_{com}$ ($O_V, O_F$ are deterministic given $S_{com}$ anyway), we have by Theorem 4 (with $H \equiv O_S$):

$$\Pr[\text{Collision} : O_S \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{O_s, O_F, O_V} (|R|)] - \Pr[\text{Collision} : G \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{G, O_F, O_V} (|R|)] \leq \alpha.$$  

Here $s \in O(\frac{\alpha^2}{\log \alpha})$, and $D_s$ is as in Theorem 4. And $G \leftarrow \$ means that $G$ is chosen as: pick $(com_1, ch_1, resp_1), \ldots, (com_s, ch_s, resp_s) \leftarrow D_s$, then for all $z$, pick $i_z \overset{\$}{\leftarrow} D_s$ and set $G(z) := (com_{i_z}, ch_{i_z}, resp_{i_z})$.

By averaging over the choice of $(S_{com})$, we then get that

$$|P_3 - P_4| \leq \alpha,$$

$P_4 := \Pr[\text{Collision} : (S_{com}) \leftarrow \$, G \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{G, O_F, O_V}].$

We construct the adversary $A_4$: Let $A_4^{O_F, O_V} (com_1, ch_1, resp_1, \ldots, com_s, ch_s, resp_s, |R|)$ pick $G$ himself as: for all $z$, $i_z \overset{\$}{\leftarrow} D_s$, $G(z) := (com_{i_z}, ch_{i_z}, resp_{i_z})$.

Then

$$P_4 = \Pr[\text{Collision} : (S_{com}) \leftarrow \$, (com_1, ch_1, resp_1), \ldots, (com_s, ch_s, resp_s) \leftarrow D_Y,$$

$$(com, ch, resp, ch', resp') \leftarrow A_4^{O_F, O_V} (com_1, ch_1, resp_1, \ldots, com_s, ch_s, resp_s, |R|)].$$

(Note that the distribution $D_Y$ depends on the choice of $S_y$.)

Let $A_5^{O_F, O_V} (|\Sigma \Psi|^{\otimes s}, |R|)$ be the algorithm that does the following: For each $i$, it takes one copy of the state $|\Sigma \Psi|$ (given as input) and measures it in the computational basis to get $(com_i, ch_i, resp_i)$. Then $A_5$ runs $A_4^{O_F, O_V} (com_1, ch_1, resp_1, \ldots, com_s, ch_s, resp_s, |R|)$.

By definition of $|\Sigma \Psi|$ (Definition 5), each $(com_i, ch_i, resp_i)$ chosen by $A_5$ is independently distributed according to $D_Y$. Thus

$$P_4 = P_5,$$

$P_5 := \Pr[\text{Collision} : (S_{com}) \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_5^{O_F, O_V} (|\Sigma \Psi|^{\otimes s}, |R|)].$

Converting the $|\alpha_i|$: The adversary $A_5$ is almost an adversary as in Theorem 6 (Hardness of the two values problem), with one exception: the input to $A_5$ is a state $|R| = |\Sigma \Psi|^{\otimes m} \otimes |\alpha_1 \rangle \otimes \cdots \otimes |\alpha_n \rangle$ with $|\alpha_i \rangle := (\cos \frac{\pi i}{2^n}) |\Sigma \Psi \rangle + (\sin \frac{\pi i}{2^n}) |\bot \rangle$. Theorem 6 on the other hand assumes an adversary that takes as input states in the span of $|\Sigma \Psi|$ and $|\Sigma \Phi| := \sum_{com, ch, resp} 2^{-(\ell_{com} + \ell_{ch} + \ell_{resp}/2)} |com, ch, resp\rangle$. Let $|\alpha_i \rangle := (\cos \frac{\pi i}{2^n}) |\Sigma \Psi \rangle + (\sin \frac{\pi i}{2^n}) |\Sigma \Phi \rangle$. $|\bar{R}| = |\Sigma \Psi|^{\otimes m} \otimes |\alpha_1 \rangle \otimes \cdots \otimes |\alpha_n \rangle$ Let $U_\alpha |\Sigma \Phi \rangle := |\bot \rangle$ and $U_\alpha |\bot \rangle := |\Sigma \Phi \rangle$ and $U_\alpha |\Phi \rangle := |\Phi \rangle$ for $|\Phi \rangle$ orthogonal to $|\bot \rangle, |\Sigma \Phi \rangle$.

Let $A_6^{O_F, O_V} (|\Sigma \Psi|^{\otimes s}, |\bar{R}|)$ be the algorithm that runs $A_5^{O_F, O_V} (|\Sigma \Psi|^{\otimes s}, (I^{\otimes m} \otimes U_\alpha^{\otimes n})|\bar{R}|)$. Then

$$P_5 \leq P_5 + \text{TD}(|\Sigma \Psi|^{\otimes m} \otimes U_\alpha^{\otimes n} |\bar{R}|, |R|),$$

$P_5 := \Pr[\text{Collision} : (S_{com}) \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_6^{O_F, O_V} (|\Sigma \Psi|^{\otimes s}, |\bar{R}|)].$
Write $|\Sigma\Psi\rangle$ as $|\Sigma\Psi\rangle = \gamma|\Sigma\Phi\rangle + \delta|\Sigma\Phi^\perp\rangle$ with $|\Sigma\Phi^\perp\rangle$ a state orthogonal to $|\Sigma\Phi\rangle$. Write short $c := \cos \frac{\ell}{2m}$ and $s := \sin \frac{\ell}{2m}$. Then
\[
\chi := (\alpha_j|U_n|\tilde{\alpha}_j)
= (c|\Sigma\Psi\rangle + s|\bot\rangle)^n(U_n(c|\Sigma\Psi\rangle + s|\Sigma\Phi\rangle))
= c^2|\Sigma\Psi\rangle|U_n|\Sigma\Psi\rangle + s^2|\bot\rangle + cs|\Sigma\Psi\rangle|\bot\rangle + cs|\bot\rangle|U_n|\Sigma\Psi\rangle
\]
\[
\chi = c^2|\Sigma\Psi\rangle + s^2c + cs = c^2(1 - |\gamma|^2) + s^2 + cs|\gamma| = 1 - c^2|\gamma|^2 + cs|\gamma|.
\]
In ($*$) we use that $|\bot\rangle, |\Sigma\Phi\rangle, |\Sigma\Phi^\perp\rangle$ are orthogonal. Furthermore,
\[
\gamma = \langle \Sigma\Phi | \Sigma\Psi \rangle = \sum_{(ch,resp)\in S_{\text{com}}} 2^{-(\ell_{\text{com}} + \ell_{ch} + \ell_{\text{resp}})/2} \cdot 2^{-\ell_{\text{com}}/2} / \sqrt{k}
\]
Thus
\[
\chi = 1 - c^2|\gamma|^2 + cs|\gamma| \geq 1 - c^2|\gamma|^2 \geq 1 - \gamma^2
\]
and hence
\[
\text{TD}(|\alpha_j\rangle, |U_n|\tilde{\alpha}_j\rangle) = \sqrt{1 - \chi^2} \leq \sqrt{1 - (1 - \gamma^2)} \leq \sqrt{2} \gamma = 2^{-(\ell_{ch} + \ell_{\text{resp}} - 1)/2} / \sqrt{k}.
\]
With (54), we get
\[
P_5 \leq P_6 + \text{TD}((I_{\text{com}}^\otimes n \otimes I_{\alpha}^\otimes n)|\tilde{R}\rangle, |R\rangle) = P_6 + \sum_{i=1}^n \text{TD}(|\alpha_j\rangle, |U_n|\tilde{\alpha}_j\rangle)
\]
\[
\leq P_6 + n2^{-(\ell_{ch} + \ell_{\text{resp}} - 1)/2} / \sqrt{k}. \tag{55}
\]
Wrapping up: Note that $A_6$ is an adversary as in Theorem 6 (Hardness of the two value problem). Thus by Theorem 6 (with $h := n + m + s$), we have:
\[
P_6 \leq O \left( \frac{n + m + s}{2^{\ell_{\text{com}}/2}} + \frac{(q_v + q_F)^{1/4}k^{1/4}}{2^{(\ell_{ch} + \ell_{\text{resp}})/4}} + \frac{(q_v + q_F)^{1/2}}{k^{1/4}} \right). \tag{56}
\]
Assume that $P_A$ is non-negligible. By (39), and since $q_P, q_R$ are polynomially-bounded, $P_1$ is then non-negligible. By (50), and since $q_E$ is polynomially-bounded and $\ell_{\text{rand}}$ superlogarithmic, $P_2$ is non-negligible. Since $P_3$ is non-negligible, there is a polynomial $p_2$ such that $p_2(1/P_2)$ on an infinite set $K$. Let $n, m := q_F^2q_F^2$. Note that $n, m$ are polynomially-bounded since $q_F$ is polynomially-bounded. Then $O(q_F^2/\sqrt{m} + q_F^2/\sqrt{m}) = O(1/p_2) = o(P_2)$ on $K$. Thus by (51), $P_2 \leq P_3 + o(P_2)$ on $K$, hence $P_3$ is non-negligible. Since $P_3$ is non-negligible, there is a polynomial $p_3$ such that $1/p_3 \leq P_3$ on an infinite set $K$. Let $\alpha := 1/2p_3$. Then by (22), $|P_3 - P_4| \leq P_3/2$ on $K$. Thus $P_4$ is non-negligible. By (53), $P_4 = P_5$, hence $P_5$ is non-negligible. By (54), $P_5 \geq P_5 - n2^{-(\ell_{ch} + \ell_{\text{resp}} - 1)/2} / \sqrt{k}$. Since $\ell_{\text{resp}}$ is superlogarithmic and $k = 2^{\ell_{ch} + \ell_{\text{resp}}}/3^n$ and $n$ polynomially-bounded, this implies that $P_5$ is non-negligible. But $q_S, q_F, q_V, q_E$ are polynomially-bounded and thus $n, m$ and $s$ are $O(2\alpha^2(\log \frac{4\alpha^2}{\alpha}))$ polynomially-bounded, and $\ell_{\text{com}}, \ell_{\text{resp}}$ are superlogarithmic, by (50) we have that $P_5$ is negligible. Thus the assumption that $P_A$ is non-negligible is wrong. Hence $P_A$ is negligible. This shows part (ii) of the lemma.

We now show part (ii) of the lemma. For an adversary $A$ outputting $(\text{com}, ch, \text{resp}, ch', \text{resp}')$, let $B$ be the adversary that runs $(\text{com}, ch, \text{resp}, ch', \text{resp}') \leftarrow A$, then invokes $w \leftarrow O_E(\text{com}, ch, \text{resp}, ch', \text{resp}')$ and returns $w$. Note that $B$ makes $q_E + 1$ queries to $O_E$, and the same number of queries to the other oracles as $A$. By definition of $O_E$, we have
\[
\Pr[(ch, \text{resp}) \neq (ch', \text{resp}') \wedge (ch, \text{resp}), (ch', \text{resp}') \in S_{\text{com}} : (\text{com}, ch, \text{resp}, ch', \text{resp}') \leftarrow A^{O_{\text{alt}}}] \leq \Pr[w = w_0 : w \leftarrow B^{O_{\text{alt}}}] + \Pr[w \neq w_0 : w \leftarrow B^{O_{\text{alt}}}] \leq \Pr[w = w_0] + \Pr[w \neq w_0]
\]
By (i) the rhs is negligible, thus the lhs is too. This proves (ii).
F Proofs for Section 5

F.1 Proof for Lemma 15

Proof of Lemma 15: Perfect completeness: By definition of \(O_S\), we have that \(x_i \in S_{y_i}\) for all \((y_i, x_i) := O_S(z_i)\). Hence \(O\nu(y_i, x_i) = 1\) for all \(i\). Thus \(\text{COM}_{\text{verif}}(c, m, u) = 1\) for \((c, u) \leftarrow \text{COM}(m)\). Hence we have perfect completeness.

Computational strict binding: Consider an adversary \(A^\text{O-alt}\) against the computational strict binding property. Let \(\mu\) be the probability that \(A^\text{O-alt}\) outputs \((c, m, u, m', u')\) such that \((m, u) \neq (m', u')\) and \(ok = ok' = 1\) with \(ok = \text{COM}_{\text{verif}}(c, m, u)\) and \(ok' = \text{COM}_{\text{verif}}(c, m', u')\). We need to show that \(\mu\) is negligible.

Let \(c = (p_1, \ldots, p_{|m|}, y_1, \ldots, y_{|m|}, b_1, \ldots, b_{|m|})\) and \(u = (x_1, \ldots, x_{|m|})\) and \(u' = (x_1', \ldots, x_{|m'|})\). Then \((m, u) \neq (m', u')\) implies that for some \(i\), \((x_i, m_i) \neq (x_i', m_i')\). If \(x_i = x_i'\), then from \(ok = ok' = 1\) we have \(m_i = b_i \oplus \text{bit}(p_i(x_i) = b_i \oplus \text{bit}(p_i(x_i') = m_i',\in contradiction to \((x_i, m_i) \neq (x_i', m_i')\). So \(x_i \neq x_i'\). Furthermore, \(ok = ok' = 1\) implies that \(O\nu(y_i, x_i) = O\nu(y_i', x_i') = 1\), i.e., \(x_i, x_i' \in S_{y_i}\). So \(A^\text{O-alt}\) finds \(x_i \neq x_i'\) with \(x_i, x_i' \in S_{y_i}\) with probability \(\mu\). By Corollary 9 [Hardness of two values 2], this implies that \(\mu\) is negligible.

Computational binding: This is implied by computational strict binding.

Statistical hiding: Fix \(m, m' \in \{0, 1\}\). Let \((y, x) := O_S(z), z \leftarrow \{0, 1\}^\ell_{\text{rand}}, p \leftarrow \{1, \ldots, \ell_{\text{ch}} + \ell_{\text{resp}}\}, b := m \oplus \text{bit}(p(x)). \) Let \(\hat{y} \leftarrow \ell_{\text{com}}, \hat{x} \leftarrow S_{\hat{y}}\). Define analogously \(y', x', \hat{y}', \hat{x}'\).

Let \(D\) be the distribution that returns \((\hat{y}, \hat{x})\) with \(\hat{y} \leftarrow \{0, 1\}^\ell_{\text{com}}, \hat{x} \leftarrow S_{\hat{y}}\). Note that by definition of \(O_S\), \(O_S(z)\) is initialized according to \(D\) by Lemma 33 for fixed choice of the sets \(S\), \(SD((O_S, y, x); (O_S, \hat{y}, \hat{x})) \leq 2^{(\ell_{\text{com}} - \ell_{\text{rand}})/2 - \sqrt{\ell}} =: \mu_1\). (With \(X := \{0, 1\}^\ell_{\text{rand}}, Y := \{(y, x) : y \in \{0, 1\}^{\ell_{\text{com}}}, x \in S_y\}\), and \(O := O_S.\)) Thus for random \(S\) and random \(p\), \(SD((O_{\text{alt}}, p, y, \text{bit}(p(x) \oplus m); (O_{\text{alt}}, p, \hat{y}, \text{bit}(\hat{p}(x) \oplus m)) \leq \mu_1\). Let \(b^* \leftarrow \{0, 1\}\). For fixed \(\hat{y}\) and \(p\) and random sets \(S_y\) and random \(p\), \(SD((S_y, \hat{y}, \text{bit}(\hat{p}(x)); (S_y, \hat{y}, b^*) \leq 1/2\sqrt{\ell} =: \mu_2\) by Lemma 34. Thus for random \(\hat{y}\) and \(p\), \(SD((O_{\text{alt}}, p, \hat{y}, \text{bit}(\hat{p}(x) \oplus m); (O_{\text{alt}}, p, \hat{y}, b^* \oplus m)) \leq \mu_2\). And \((O_{\text{alt}}, p, \hat{y}, b^* \oplus m)\) has the same distribution as \((O_{\text{alt}}, p, \hat{y}, b^*)\) since \(b^* \in \{0, 1\}\) is uniform and independently chosen from \(O_{\text{alt}}, \hat{y}\.\) Hence \(SD((O_{\text{alt}}, p, \hat{y}, \text{bit}(\hat{p}(x) \oplus m); (O_{\text{alt}}, p, \hat{y}, b^*)) \leq \mu_1 + \mu_2\). Analogously, \(SD((O_{\text{alt}}, p, \hat{y}', \text{bit}(\hat{p}(x') \oplus m); (O_{\text{alt}}, p, \hat{y}', b^{*'})) \leq \mu_1 + \mu_2\). Then \(SD((O_{\text{alt}}, p, \hat{y}, \text{bit}(\hat{p}(x) \oplus m); (O_{\text{alt}}, p, \hat{y}', \text{bit}(\hat{p}(x') \oplus m))) \leq 2(\mu_1 + \mu_2)\).

Fix \(m_1, m_2\) with \(|m_1| = |m_2|\). Let \(z_i \leftarrow \{0, 1\}^\ell_{\text{rand}}, (y_i, x_i) := O_S(z_i), p_i \leftarrow \{1, \ldots, \ell_{\text{ch}} + \ell_{\text{resp}}\}, b_i := m_i \oplus \text{bit}(p_i(x_i))\) and analogously \(y_i', x_i', \hat{y}_i, \hat{x}_i\). By induction over \(n\), and using (57), we get for all \(1 \leq n \leq |m_1|:\)

\[
SD((O_{\text{alt}}, p_i)_{i=1, \ldots, n}, (y_i)_{i=1, \ldots, n}, (\text{bit}(p_i(x_i) \oplus m_i)_{i=1, \ldots, n}); (O_{\text{alt}}, (p_i')_{i=1, \ldots, n}, (y_i')_{i=1, \ldots, n}, (\text{bit}(p_i(x'_i) \oplus m'_i)_{i=1, \ldots, n})) \leq 2n(\mu_1 + \mu_2).
\]

For \(n = |m_1|\), this becomes

\[
SD((O_{\text{alt}}, c), (O_{\text{alt}}, c')) \leq 2|m_1|(\mu_1 + \mu_2) =: \mu\quad \text{with } c \leftarrow \text{COM}(m), c' \leftarrow \text{COM}(m').
\]

Since \(|m_1|\) is polynomially-bounded, and \(\ell_{\text{rand}} - \ell_{\text{com}} - k<br/>is superlogarithmic, and \(k\) is super-polynomial, \(\mu\) is negligible. Thus COM is statistically hiding. \(\square\)

F.2 Proof of Lemma 16

Proof of Lemma 16: Our adversary is as follows:

• \(B_1(|m|)\) invokes \(E_1\) from Theorem 7 [Searching one value] \(|m|\) times to get \((y_i, |\Psi(y_i)|)\) for \(i = 1, \ldots, |m|\).\footnote{\(E_1\) expects an input \(|\Sigma\Psi|\). \(|\Sigma\Psi|\) can be computed using the oracle \(O_{\Phi}\).} Let \(p_1, \ldots, p_{|m|} \leftarrow \{1, \ldots, \ell_{\text{ch}} + \ell_{\text{resp}}\}. \) Let \(b_1, \ldots, b_{|m|} \leftarrow \{0, 1\}. \) Output
Lemma 15), we immediately get (a). We prove (b): By definition of $\text{COM}$ and $\text{Commitment entropy}$: negligible probability. By definition of $\text{COM}$, this implies that $\text{COM}$ is overwhelming since $|m|$ is polynomial and $\ell_{\text{com}}$ and $k$ are superlogarithmic.

\section*{G Proofs for Section 6}

\subsection*{G.1 Proof of Lemma 19}

\textbf{Proof of Lemma 19} \textbf{Completeness}: We need to show that with overwhelming probability, (a) $\text{COM}_{\text{verify}}(c_{\text{ch}}, \text{resp}_{\text{ch}}, u_{\text{ch}}) = 1$ for $(c_{\text{ch}}, u_{\text{ch}}) \leftarrow \text{COM}(\text{resp}_{\text{ch}})$ and (b) $\text{O}_V(\text{com}, c_{\text{ch}}, \text{resp}_{\text{ch}}) = 1$ for uniform $\text{com}$, $c_{\text{ch}}$ and $\text{resp}_{\text{ch}} := \text{O}_P(w, \text{com}, c_{\text{ch}})$. From the completeness of COM (Lemma 15), we immediately get (a). We prove (b): By definition of $\text{O}_P$ and $\text{O}_V$, (b) holds iff $\exists \text{resp}_{\text{ch}}, \text{resp} \in S_{\text{com}}$. We thus need to show that $p_1 := \Pr[\exists \text{resp}_{\text{ch}}, \text{resp} \in S_{\text{com}}]$ is overwhelming. $S_{\text{com}}$ is a uniformly random subset of size $k = 2^{\ell_{\text{ch}} + \ell_{\text{resp}} / 3}$ of $X = \{0, 1\}^{\ell_{\text{ch}}} \times \{0, 1\}^{\ell_{\text{resp}}}$. Thus $p_1$ is lower bounded by the probability $p_2$ that out of $k$ uniform independent samples from $\{0, 1\}^{\ell_{\text{ch}}}$, at least one is $c_{\text{ch}}$. Thus $p_1 \geq p_2 = 1 - (1 - 2^{-\ell_{\text{ch}}})^k = 1 - \left(\left(1 - 1/2^{\ell_{\text{ch}}}ight)^2\right)^{2\ell_{\text{resp}} / 3} \geq 1 - e^{-2\ell_{\text{resp}} / 3}$ where $(\ast)$ uses the fact that $(1 - 1/n)^n$ converges from below to $1/e$ for integers $n \to \infty$. Thus $p_1$ is overwhelming for superlogarithmic $\ell_{\text{resp}}$, and the sigma-protocol is complete.

\textbf{Commitment entropy}: We need to show that $\text{com}^* \leftarrow P_1(s, w)$ has superlogarithmic min-entropy. Since $\text{com}^* = (\text{com}, \ldots)$, and $\text{com}$ is uniformly distributed on $\{0, 1\}^{\ell_{\text{com}}}$, the min-entropy of $\text{com}^*$ is at least $\ell_{\text{com}}$ which is superlogarithmic.

\textbf{Perfect special soundness}: Observe that $V(s, \text{com}^*, c_{\text{ch}}, \text{resp}^* ) = V(s, \text{com}^*, c_{\text{ch}}, \text{resp}^{**}) = 1$ and $c_{\text{ch}} \neq c_{\text{ch}}'$ implies $(c_{\text{ch}}, \text{resp}), (c_{\text{ch}}', \text{resp}) \in S_{\text{com}}$ and $s = s_0$ and $c_{\text{ch}} \neq c_{\text{ch}}'$ which in turn implies $\text{O}_E(\text{com}, c_{\text{ch}}, \text{resp}, c_{\text{ch}}', \text{resp}') = w_0$ and $(s, w_0) \in R$. Thus an extractor $E$ that just outputs $\text{O}_E(\text{com}, c_{\text{ch}}, \text{resp}, c_{\text{ch}}', \text{resp}')$ achieves perfect special soundness.

\textbf{Computational strict soundness}: We need to show that a polynomial-time $A$ will only with negligible probability output $(\text{com}^*, c_{\text{ch}}, \text{resp}^*, \text{resp}^{**})$ such that $\text{resp}^* \neq \text{resp}^{**}$ and $V(s, \text{com}^*, c_{\text{ch}}, \text{resp}^* ) = V(s, \text{com}^*, c_{\text{ch}}, \text{resp}^{**}) = 1$. Assume $A$ outputs such a tuple with non-negligible probability. By definition of $V$, this implies that $\text{resp}^* = (\text{resp}, u)$, $\text{resp}^{**} = (\text{resp}', u')$, and $\text{com}^*$ contains $c_{\text{ch}}$ such that $\text{COM}_{\text{verify}}(c_{\text{ch}}, \text{resp}, u) = 1$ and $\text{COM}_{\text{verify}}(c_{\text{ch}}, \text{resp}', u') = 1$. Since $\text{resp}^* \neq \text{resp}^{**}$, this contradicts the computational strict binding property of $\text{COM}, \text{COM}_{\text{verify}}$ (Lemma 15). Thus the sigma-protocol has computational strict soundness.
Statistical HVZK: Let $S$ be the simulator that picks $z \xleftarrow{\$} \{0,1\}^\ell_{rand}$, computes $(com, ch, \text{resp}) := O_S(z)$, and $(c_c, u_c) \xleftarrow{\$} \text{COM}(0^\ell_{resp})$ for all $c \in \{0,1\}^\ell_{ch} \setminus \{ch\}$, and $(c_{ch}, u_{ch}) \xleftarrow{\$} \text{COM}(\text{resp})$, and returns $(com^*, ch, \text{resp}^*)$ with $com^* := (com, (c_{ch})_{ch \in \{0,1\}^\ell_{ch}})$ and $\text{resp}^* := (\text{resp}_{ch}, u_{ch})$. We now compute the difference between the probabilities from the definition of statistical HVZK [Definition 1] for $(s, w) \in R$, i.e., for $s = s_0$ and $w = w_0$. In the calculation, $com^*$ always stands short for $(com, (c_{ch})_{ch \in \{0,1\}^\ell_{ch}})$ and $\text{resp}^*$ for $(\text{resp}_{ch}, u_{ch})$.

$$\Pr[b = 1 : com^* \leftarrow P_1(s,w), ch \xleftarrow{\$} \{0,1\}^\ell_{ch}, \text{resp}^* \leftarrow P_2(ch), b \leftarrow A(com^*, ch, \text{resp}^*)] = \Pr[b = 1 : com \xleftarrow{\$} \{0,1\}^\ell_{com}, ch \xleftarrow{\$} \{0,1\}^\ell_{ch}, \text{[for all } c \in \{0,1\}^\ell_{ch} : z_c \xleftarrow{\$} \{0,1\}^\ell_{rand}, \text{resp}_c := O_P(w, com, c, z_c), (c_c, u_c) \leftarrow \text{COM}(\text{resp}_c)], b \leftarrow A(com^{*}, ch, \text{resp}^{*})]$$

$$\approx \Pr[b = 1 : com \xleftarrow{\$} \{0,1\}^\ell_{com}, ch \xleftarrow{\$} \{0,1\}^\ell_{ch}, \text{[for all } c \in \{0,1\}^\ell_{ch} \setminus \{ch\} : (c_c, u_c) \leftarrow \text{COM}(0^\ell_{resp}), \text{z}_ch \xleftarrow{\$} \{0,1\}^\ell_{rand}, \text{resp}_ch := O_P(w, com, ch, z_{ch}), (c_{ch}, u_{ch}) \leftarrow \text{COM}(\text{resp}_ch), b \leftarrow A(com^{*}, ch, \text{resp}^{*})]$$

Here $a \approx b$ means that $|a - b| \leq \varepsilon_0$ where $\varepsilon_0 := 2^\ell_{ch} \varepsilon_{\text{COM}}$ and $\varepsilon_{\text{COM}}$ is the statistic distance between commitments $\text{COM}(com^{*})$ and $\text{COM}(0^\ell_{resp})$. We have that $\varepsilon_{\text{COM}}$ is negligible by [Lemma 15] (statistical hiding of COM).

We abbreviate $[\text{for all } c \in \{0,1\}^\ell_{ch} \setminus \{ch\} : (c_c, u_c) \leftarrow \text{COM}(0^\ell_{resp})]$ with $[\text{COM}(0)]$ and continue our calculation:

$$\cdots \approx \Pr[b = 1 : (com, ch, \text{resp}_{ch}) \xleftarrow{\$} \text{D'}, [\text{COM}(0)], (c_c, u_c) \leftarrow \text{COM}(\text{resp}_{ch}), b \leftarrow A(com^{*}, ch, \text{resp}^{*})]$$

Here $D_{\text{com, ch}}$ is the uniform distribution on $\{\text{resp} : (ch, \text{resp}) \in \text{S}_{\text{com}}\}$. (Or, if that set is empty, $D_{\text{com, ch}}$ assigns probability 1 to $\bot$.) And $a \approx b$ means that $|a - b| \leq \varepsilon_1$ where $\varepsilon_1 := \frac{1}{2} \sqrt{2^\ell_{resp} / 2^\ell_{rand}}$. The last equation follows from [Lemma 33] with $X := \{0,1\}^\ell_{rand}$ and $Y := \{0,1\}^\ell_{resp}$ and $D := \text{D}_{\text{ch, com}}$, and using the fact that for all $z$, $O_P(w_0, com, ch, z)$ is chosen according to $D_{ch, com}$. (Note that the adversary $A$ has access to $O_P$, but that is covered since $O$ occur on both sides of the statistical distance in [Lemma 33]) We continue the computation:

$$\cdots \approx \Pr[b = 1 : (com, ch, \text{resp}_{ch}) \xleftarrow{\$} \text{D'}, [\text{COM}(0)], (c_c, u_c) \leftarrow \text{COM}(\text{resp}_{ch}), b \leftarrow A(com^{*}, ch, \text{resp}^{*})]$$

Here $D'$ is the distribution resulting from choosing $com \xleftarrow{\$} \{0,1\}^\ell_{com}$, $(ch, \text{resp}) \xleftarrow{\$} S_{\text{com}}$. By [Lemma 35] $\varepsilon_2 \leq \frac{2k^2}{2^{\ell_{ch} + \ell_{resp}} + 2^{\ell_{ch}/2}}$. We continue

$$\cdots \approx \Pr[b = 1 : z \xleftarrow{\$} \{0,1\}^\ell_{rand}, (com, ch, \text{resp}) := O_S(z), [\text{COM}(0)], (c_{ch}, u_{ch}) \leftarrow \text{COM}(\text{resp}_{ch}), b \leftarrow A(com^{*}, ch, \text{resp}^{*})]$$

Here $\varepsilon_3 = \sqrt{(2^{\ell_{com}} \cdot k)/2^{\ell_{rand}}}$. This follows from [Lemma 33] with $D := D'$ and $X := \{0,1\}^\ell_{rand}$ and $Y := \{(com, ch, \text{resp}) : (ch, \text{resp}) \in S_{\text{com}}\}$. (Note that $|Y| = 2^{\ell_{com}} \cdot k$.) We continue

$$\cdots = \Pr[b = 1 : (com^{*}, ch, \text{resp}^{*}) := S(s), b \leftarrow A(com^{*}, ch, \text{resp}^{*})].$$

Thus the difference of probabilities from the definition of statistical HVZK is bounded by $\varepsilon := \varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$. And $\varepsilon$ is negligible since $\varepsilon_{\text{COM}}$ is negligible, and $k = 2^\ell_{ch} + \lfloor \ell_{resp}/3 \rfloor$, and $\ell_{ch}$ is logarithmic, and $\ell_{resp}, \ell_{com}$ are superlogarithmic, and $\ell_{rand} = \ell_{com} + \ell_{resp}$. □
Proof of Lemma 20. According to Definition 2 (specialized to the case of the sigma-protocol from Definition 18) we need to construct a polynomial-time quantum adversary $A_1, A_2, A_3$ such that:

- **Adversary success:**
  
  $$P_A := \Pr[ok = 1 : s \leftarrow A_1, com^* \leftarrow A_2, ch \leftarrow \{0,1\}^{\ell_{ch}},$$
  
  $$\text{resp}^* \leftarrow A_3(ch), ok = V(s, com^*, ch, \text{resp}^*)]$$
  
  $$= \Pr[ok_v = 1 \land ok_c = 1 \land s = s_0 : s \leftarrow A_1, (com, (c, ch) \chi \in \{0,1\}^{\ell_{ch}}) \leftarrow A_2,$$
  
  $$ch \leftarrow \{0,1\}^{\ell_{ch}}, (\text{resp}, u) \leftarrow A_3(ch), ok_v := \mathcal{O}_V(com, ch, \text{resp}),$$
  
  $$ok_c = \text{COM}_{\text{verify}}(c, ch, \text{resp}, u)]$$

  is overwhelming.

- **Extractor failure:** For any polynomial-time quantum $E$ (with access to the final state of $A_1$), $\Pr[s = s_0, w = w_0 : s \leftarrow A_1, w \leftarrow E(s)]$ is negligible.

Our adversary is as follows:

- Let $B_1, B_2$ be the adversary from Lemma 16 (Attack on COM). (That is, $B_1(|m|)$ produces a fake commitment which $B_2(m)$ then opens to $m$.)
- $A_1$ outputs $s_0$.
- $A_2$ invokes $E$ from Theorem 7 (Searching one value) to get $(\text{com}, |\Psi(\text{com})|)$. Then $A_2$ invokes $c \leftarrow B_1(\ell_{\text{resp}})$ for all $c \in \{0,1\}^{\ell_{ch}}$. $A_2$ outputs $com^* := (com, (c, ch) \chi \in \{0,1\}^{\ell_{ch}})$.
- Let $P_{ch}(ch', \text{resp}') := 1$ iff $ch' = ch$. $A_3(ch)$ invokes $E_2(n, \delta_{\min}, \text{com}, |\Psi(\text{com})|)$ from Theorem 7 with oracle access to $P := P_{ch}$ and with $n := \ell_{\text{com}}$ and $\delta_{\min} := 2^{-\ell_{ch} - 1}$ to get $\text{resp}$. Then $A_3$ invokes $u \leftarrow B_1(\text{resp})$ to get opening information for $c, ch$. $A_3$ outputs $\text{resp}^* := (\text{resp}, u)$.

**Adversary success:** By Lemma 16 $\text{COM}_{\text{verify}}(c, ch, \text{resp}, u) = 1$ with overwhelming probability. Thus $ok_c = 1$ with overwhelming probability in (58).

By Theorem 7, the probability that $E_2$ fails to return $(ch', \text{resp})$ with $(ch', \text{resp}) \in S_{\text{com}} \land P_{ch}(ch', \text{resp}) = 1$ is at most:

$$f := 2^{-\ell_{\text{com}}} + f_\delta$$

with

$$f_\delta := \Pr\left[\frac{|\{(ch', \text{resp}) \in S_{\text{com}} : P_{ch}(ch', \text{resp}) = 1\}|}{|S_{\text{com}}|} < \delta_{\min}\right]$$

Let $P' := \{x : P_{ch}(x) = 1\}$ and $X := \{0,1\}^{\ell_{ch}} \times \{0,1\}^{\ell_{\text{com}}}$. Then $|P'|/|X| = 2^{-\ell_{ch}}$. Since $S_{\text{com}} \subseteq X$ is chosen uniformly at random with $|S_{\text{com}}| = k$, by Lemma 32 we have:

$$f_\delta = \Pr[|S_{\text{com}} \cap P'|/|S_{\text{com}}| < \delta_{\min}] \leq e^{-2k(2^{-\ell_{ch}} - \delta_{\min})^2} = e^{-k2^{-2\ell_{ch}} - 1}.$$ 

Thus $f \leq 2^{-\ell_{\text{com}}} + e^{-k2^{-2\ell_{ch}} - 1}$ is negligible since $\ell_{\text{com}}$ is superpolynomial, $\ell_{\text{ch}}$ logarithmic, and $k$ superpolynomial. Thus with overwhelming probability $E_2$ returns $(ch', \text{resp}) \in S_{\text{com}}$ with $P_{ch}(ch', \text{resp}) = 1$. $P_{ch}(ch', \text{resp}) = 1$ implies $ch' = ch$. Hence $(ch, \text{resp}) \in S_{\text{com}}$, thus $\mathcal{O}_V(com, ch, \text{resp}) = 1$, thus $ok_c = 1$ with overwhelming probability. Since $s = s_0$ by construction of $A_1$, it follows that $P_A$ is overwhelming. Thus we have adversary success.

**Extractor failure:** It remains to show extractor failure. Fix some polynomial-time $E$. Since $A_1$ only returns a fixed $s_0$ and has a trivial final state, without loss of generality we can assume that $E$ does not use its input $s$ or $A_1$’s final state. Then

$$P_E := \Pr[s = s_0, w = w_0 : s \leftarrow A_1, w \leftarrow E^{\mathcal{O}_{\text{init}}}(s)] = \Pr[w = w_0 : w \leftarrow E^{\mathcal{O}_{\text{init}}}]$$

is negligible by Corollary 9 (Hardness of two values). This shows extractor failure.

---

*Using $\mathcal{O}_A$ to get the input $|\Sigma\Psi|$ for $E_1$. 

---

60
G.3 Proof of Lemma 23

Proof of Lemma 23. Completeness and statistical HVZK and commitment entropy hold trivially, because they only have to hold for \((s, w) \notin R' = \emptyset\). Computational strict soundness is shown exactly as in the proof of Lemma 19 (Security of the sigma-protocol). (The definition of computational strict soundness is independent of the relation \(R'\).

Computational special soundness: Let \(E_S\) be an algorithm that always outputs \(\perp\). By Definition 1 (Properties of sigma-protocols) we have to show that the following probability is negligible:

\[
P_S := \Pr[(s, w) \notin R' \land ch \neq ch' \land ok = ok' = 1 : (s, \text{com}^*, ch, \text{resp}^*, ch', \text{resp}'^*) \leftarrow A^{O_{alt}},
ok \leftrightarrow V(s, \text{com}^*, ch, \text{resp}^*), ok' \leftrightarrow V(s, \text{com}^*, ch', \text{resp}'^*),
(w \leftarrow E_S(s, \text{com}^*, ch, \text{resp}^*, ch', \text{resp}'^*))
\leq \Pr[ch \neq ch' \land (ch, resp), (ch', resp') \in S_{com} : (\text{com}^*, ch, \text{resp}^*, ch', \text{resp}'^*) \leftarrow A^{O_{alt}},
(ch, \ldots) := \text{com}^*, (resp, \ldots) := \text{resp}^*, (resp', \ldots) := \text{resp}'^*]
\]

The right hand side is negligible by Corollary 9 (Hardness of two values 2). Hence \(P_S\) is negligible. This shows that the sigma-protocol from Definition 22 has computational special soundness. □

G.4 Proof of Lemma 24

Proof of Lemma 24. By Definition 2 (specialized to the sigma-protocol from Definition 22), we need to construct a polynomial-time adversary \(A_1, A_2, A_3\) such that:

\[
P_A := \Pr[ok = 1 \land s \notin L_{R'} : s \leftarrow A_1, \text{com}^* \leftarrow A_2, ch \leftarrow \{0, 1\}^{L_{ch}}, \text{resp}^* \leftarrow A_3(ch),
ok := V(\text{com}^*, ch, \text{resp}^*)] \text{ is overwhelming.}
\]

We use the same adversary \((A_1, A_2, A_3)\) as in the proof of Lemma 20. Then \(P_A\) here is the same as \(P_A\) in the proof of Lemma 20 (Here we additionally have the condition \(s \notin L_{R'}\), but this condition is vacuously true since \(R' = \emptyset\) and thus \(L_{R'} = \emptyset\).
And in the proof of Lemma 20 we showed that \(P_A\) is overwhelming. □

H Proofs for Section 7

H.1 Proof of Theorem 26

Lemma 49 (Attack on Fiat-Shamir) There exists a total knowledge break (Definition 2) against the Fiat-Shamir construction based on the sigma-protocol from Definition 18. (For any \(r\).

Proof. According to Definition 2 (specialized to the case of the Fiat-Shamir construction based on the sigma-protocol from Definition 18) we need to construct a polynomial-time quantum adversary \(A_1, A_2\) such that:

- Adversary success:
  \[
  \hat{P}_A := \Pr[vi, ok_i = 1 : s \leftarrow \hat{A}_1^{H_{alt}}, ((\text{com}^*_i)_i, (\text{resp}^*_i)_i) \leftarrow \hat{A}_2^{H_{alt}},
  ch_1 \ldots \| ch_r := H(s, (\text{com}^*_i)_i), ok_i := V(\text{com}^*_i, ch_i, \text{resp}^*_i)]
  \]
  is overwhelming. Here \(V\) is the verifier of the sigma-protocol (Definition 18).
- Extractor failure: For any polynomial-time quantum \(E\) (with access to the final state of \(A_1\)), \(\Pr[s = s_0, w = w_0 : s \leftarrow \hat{A}_1^{H_{alt}}, w \leftarrow E^{H, O_{alt}}(s)]\) is negligible.
Let $A_1, A_2, A_3$ be the adversary from the proof of [Lemma 20](Attack on the sigma-protocol). Our adversary is then as follows:

- $A_1$ outputs $s_0$. (Identical to $A_1$.)
- $A_2$ invokes the adversary $A_2$ $r$ times to get $\text{com}_1^*, \ldots, \text{com}_r^*$. Then $\hat{A}_2$ computes $\text{ch}_1 \parallel \ldots \parallel \text{ch}_r := H(s_1, (\text{com}_1^*)_1, \ldots, (\text{com}_r^*)_r)$. Then $\hat{A}_2$ invokes $A_3$ $r$ times to get $\text{resp}_1^* \leftarrow A_3(\text{ch}_1), \ldots, \text{resp}_r^* \leftarrow A_3(\text{ch}_r)$. Then $\hat{A}_2$ outputs $((\text{com}_1^*)_1, (\text{resp}_1^*)_1, \ldots, (\text{com}_r^*)_r, (\text{resp}_r^*)_1)$.

**Adversary success:** We have

$$1 - P_A = \Pr[\exists i. o_k = 0 : s \leftarrow A_1^{\text{alt}}, \forall i. \text{com}_i^* \leftarrow A_2^{\text{alt}},$$

$$\text{ch}_1 \parallel \ldots \parallel \text{ch}_r := H(s, (\text{com}_1^*)_1, \ldots, (\text{com}_r^*)_r)$$

$$\forall i. o_k \leftarrow V(\text{com}_i^*, \text{ch}_i, \text{resp}_i^*)]$$

$$\leq \Pr[\exists i. o_k = 0 : s \leftarrow A_1^{\text{alt}}, \forall i. \text{com}_i^* \leftarrow A_2^{\text{alt}},$$

$$\forall i. \text{ch}_i \leftarrow \{0, 1\}^\ell_{\text{ch}}, \forall i. \text{resp}_i^* \leftarrow A_3^{\text{alt}}(\text{ch}_i),$$

$$\forall i. o_k \leftarrow V(\text{com}_i^*, \text{ch}_i, \text{resp}_i^*)]$$

$$\leq \sum_{i=1}^r \Pr[\exists i. o_k = 0 : s \leftarrow A_1^{\text{alt}}, \forall i. \text{com}_i^* \leftarrow A_2^{\text{alt}},$$

$$\forall i. \text{ch}_i \leftarrow \{0, 1\}^\ell_{\text{ch}}, \forall i. \text{resp}_i^* \leftarrow A_3^{\text{alt}}(\text{ch}_i),$$

$$\forall i. o_k \leftarrow V(\text{com}_i^*, \text{ch}_i, \text{resp}_i^*)]$$

$$= \sum_{i=1}^r (1 - P_A) = r(1 - P_A).$$

Here (*) uses the fact that $H$ is only queried once (classically), and thus $H(s, (\text{com}_i^*)_1, \ldots, (\text{com}_r^*)_r)$ is uniformly random. And (**) is a union bound. And (*** is by definition of $P_A$ in the proof of [Lemma 20](Security of the sigma-protocol). There is was also shown that $P_A$ is overwhelming. Thus $1 - P_A \leq r(1 - P_A)$ is negligible and hence $P_A$ overwhelming. Thus we have adversary success.

**Extractor failure:** Extractor failure was already shown in the proof of [Lemma 20](Security of the sigma-protocol) ($A_1$ here is defined exactly as $A_1$ in the proof of [Lemma 20](Security of the sigma-protocol) and the definition of extractor failure depends only on $A_1$, not on $A_2$ or the protocol being attacked.)

Note that we have actually even shown extractor failure in the case that the extractor is allowed to choose the random oracle $H$ before and during the execution of $A_1$, because $A_1$ does not access $H$.

Now [Theorem 26](Security of the sigma-protocol) follows from [Lemma 19](Security of the sigma-protocol) (The fact that the Fiat-Shamir protocol is a classical argument of knowledge is shown in [15](Security of the sigma-protocol)).

### H.2 Proof of [Theorem 27](Security of the sigma-protocol)

**Lemma 50 (Attack on Fiat-Shamir, computational)** Then there exists a total break in [Definition 2](Attack on the sigma-protocol) against the Fiat-Shamir construction based on the sigma-protocol from [Definition 22](Attack on the sigma-protocol) (For any $r$).

**Proof.** By [Definition 2](Attack on the sigma-protocol) (specialized to the case of the Fiat-Shamir construction based on the sigma-protocol from [Definition 22](Attack on the sigma-protocol)), we need to construct a polynomial-time adversary $A_1, A_2$ such that:

$$\hat{P}_{A} := \Pr[\forall i. o_k = 1 \wedge s \notin L_{R^c} : s \leftarrow A_1^{H, \text{alt}}, ((\text{com}_1^*)_i, (\text{resp}_1^*)_i) \leftarrow A_2^{H, \text{alt}},$$

$$\text{ch}_1 \parallel \ldots \parallel \text{ch}_r := H(s, (\text{com}_1^*)_1, \ldots, (\text{com}_r^*)_r), o_k := V(\text{com}_i^*, \text{ch}_i, \text{resp}_i^*)]$$

is overwhelming.

---

Theorem 26 requires perfect completeness instead of completeness as defined here (we allow a negligible error). However, it is straightforward to see that their proof works unmodified for completeness as defined here. Also, [13](Security of the sigma-protocol) assumes that $\ell_{\text{ch}}$ is superlogarithmic, and considers the case $r = 1$. But [15](Security of the sigma-protocol) can be applied to our formulation by first parallel composing the sigma-protocol $r$ times (yielding a protocol with challenges of length $r\ell_{\text{ch}}$), and then applying the result from [15](Security of the sigma-protocol).
Here \( V \) is the verifier of the sigma-protocol (Definition 22).

We use the same adversary \((A_1, A_2)\) as in the proof of Lemma 49 (Attack on Fiat-Shamir). Then \( \tilde{P}_A \) here is the same as \( P_A \) in the proof of Lemma 49 (Here we additionally have the condition \( s \notin L_R \), but this condition is vacuously true since \( R' = \emptyset \) and thus \( L_R' = \emptyset \).) And in the proof of Lemma 49 we showed that \( \tilde{P}_A \) is overwhelming.

Now Theorem 27 follows from Lemmas 23 and 50. (The fact that the Fiat-Shamir protocol is a classical argument of knowledge is shown in 15).

I Proofs for Section 8

I.1 Proof of Theorem 29

Lemma 51 (Attack on Fischlin’s construction) There exists a total knowledge break (Definition 2) against the Fischlin construction based on the sigma-protocol from Definition 18 (Sigma-protocol).

Proof. According to Definition 2 (Total breaks) (specialized to the case of Fischlin’s construction based on the sigma-protocol from Definition 18) we need to construct a polynomial-time quantum adversary \( A_1, A_2 \) such that:

- Adversary success:
  
  \[ P_A := \Pr[\forall i. ok_i = 1 \land \sigma \leq S \land s = s_0 : s \leftarrow A_1^{H_{O\text{all}}}, \]
  
  \[ (\text{com}_{i,\ast}, ch_i, \text{resp}_{i,\ast})_{i=1..r} \leftarrow A_2^{H_{O\text{all}}}, \]
  
  \[ ok_i := V(\text{com}_{i,\ast}, ch_i, \text{resp}_{i,\ast}), \]
  
  \[ \sigma := \sum_{i=1}^{r} H(x, \text{com}_{i,\ast}, i, ch_i, \text{resp}_{i,\ast}) \]
  
  is overwhelming. \quad (59)

- Extractor failure: For any polynomial-time quantum \( E \) (with access to the final state of \( A_1 \), \( \Pr[s = s_0, w = w_0 : s \leftarrow A_1^{H_{O\text{all}}}, w \leftarrow E^{H_{O\text{all}}}(s)] \) is negligible.

Adversary success: At the first glance, it may seem that it is immediate how to construct an adversary that has adversary success: Using Theorem 7 (Searching one value), we can for each \( i \) search \( (ch_i, \text{resp}_{i}) \in S_{\text{com}} \) such that \( H(x, \text{com}_{i,\ast}, i, ch_i, \text{resp}_{i,\ast}) = 0 \). However, there is a problem: \( \text{com}_{i,\ast} \) contains commitments \( c_{ch}^i \) to all responses. Thus, after finding \( ch_i, \text{resp}_{i} \), we need to open \( c_{ch}^i \) as \( \text{resp}_{i} \). This could be done with the adversary against \( \text{COM} \) from Lemma 16 (Attack on \( \text{COM} \)). But the problem is, the corresponding openings have to be contained in \( \text{resp}_{i,\ast} \).

So we need to know these openings already when searching for \( ch_i, \text{resp}_{i} \). But at that point we do not know yet to what value the commitments \( c_{ch}^i \) should be opened! To avoid this problem, we use a special fixpoint property of the commitment scheme \( \text{COM} \) that allows us to commit in a way such that we can use the \( (ch_i, \text{resp}_{i}) \) themselves as openings for the commitments.

The fixpoint property is the following: There are functions \( \text{COM}^\ast, \text{COMopen}^\ast \) such that for any \( \text{com} \in \{0, 1\}^{\ell_{\text{com}}} \) and any \((ch, \text{resp}) \in S_{\text{com}}\), we have

\[ \text{COM}_{\text{verify}}(c, \text{resp}, u) = 1 \quad \text{for } c := \text{COM}^\ast(\text{com}) \text{ and } u := \text{COMopen}^\ast(ch, \text{resp}). \quad (60) \]

These functions are defined as follows: \( \text{COM}^\ast(\text{com}) = (p_1, \ldots, p_{\ell_{\text{resp}}}, y_1, \ldots, y_{\ell_{\text{resp}}}, b_1, \ldots, b_{\ell_{\text{resp}}}) \) with \( p_i := \ell_{ch} + i, y_i := \text{com}, b_i := 0 \). And \( \text{COMopen}^\ast(ch, \text{resp}) := (x_1, \ldots, x_{\ell_{\text{resp}}}) \) with

\footnote{Actually, \[15\] requires perfect special soundness instead of computational special soundness, as well as perfect completeness instead of completeness as defined here (we allow a negligible error). However, it is straightforward to see that their proof works unmodified for computational special soundness and completeness as defined here.

Also, \[15\] assumes that \( \ell_{ch} \) is superlogarithmic, and considers the case \( r = 1 \). But \[15\] can be applied to our formulation by first parallel composing the sigma-protocol \( r \) times (yielding a protocol with challenges of length \( r\ell_{ch} \)), and then applying the result from \[15\].}
$x_i := (ch, resp)$ for all $i$. It is easy to verify from the definition of COM\_verify (Definition 14) that (60) holds if $(ch, resp) \in S_{com}$.

Our adversary is as follows:

- $A_1$ outputs $s_0$.
- $A_2$ invokes $E_1$ from Theorem 7 (Searching one value) $r$ times to get $(com_i, |\Psi(com_i)|)$ for $i = 1, \ldots, r$. $A_2$ sets $e^i := \text{COM}^*(com_i)$ for all $i$ and all $ch \in \{0, 1\}^{\ell_{com}}$. And $com^i := (com_i, (e^i_{ch}), ch)$. Let $P_i(ch', resp') := 1$ iff $H(s, (com^i)_s, i, ch', (resp', \text{COMopen}^*(ch', resp'))) = 0$. Then, for each $i = 1, \ldots, r$, $A_2$ invokes $E_2(n, \delta_{min}, com_i, |\Psi(com_i)|)$ from Theorem 7 with oracle access to $P := P_i$ and with $n := \ell_{com}$ and $\delta_{min} := 2^{-b-1}$ to get $ch_i, resp_i$. Let $resp^i := (resp_i, \text{COMopen}^*(ch_i, resp_i))$. Then $A_2$ outputs $\pi := (com^i_s, ch_i, resp^i)_{i=1,...,r}$.

Consider an execution of $A_1, A_2$ as in (60). Let Succ$_i$ denote the event that $(ch_i, resp_i) \in S_{com_i} \land P_i(ch_i, resp_i) = 1$ in that execution. We have

$$\Pr[\text{Succ}_i] = \Pr[(ch, resp) \in S_{com_i} \land P(ch, resp) = 1 : \forall j, (com_j, |\Psi(com_j)|) \leftarrow E_1, \forall j, com^*_j := (com_j, (\text{COM}^*(com_j), ch), H \xleftarrow{\$} (\{0, 1\}^* \rightarrow \{0, 1\}^b), \forall ch', resp'. P(ch', resp') := 1 \text{ iff } H(s, (com^*_j)_s, i, ch', (resp', \text{COMopen}^*(ch', resp'))) = 0, (ch, resp) \leftarrow E_2(n, \delta_{min}, com_i, |\Psi(com_i)|)].$$

Hence by Theorem 7 (Searching one value),

$$\Pr[\text{Succ}_i] \geq 1 - 2^{-\ell_{com}} - \Pr[\left\{\left.\left[\frac{(ch, resp) \in S_{com_i} : P(ch, resp) = 1}\right| S_{com_i}\right]\right\} < \delta_{min}] = p_\delta.$$

Here $P$ and $com$ are chosen as in the rhs of (61).

In the rhs of (61), $H$ is chosen after $S_{com_j}, s, com^*_j$, and $i$ are fixed. Thus for every $(ch, resp) \in S_{com_i}$ it is independently chosen whether $P(ch, resp) = 1$ or $P(ch, resp) = 0$, where $\Pr[P(ch, resp) = 1] = 2^{-b}$. Thus

$$p_\delta = \Pr[\sum_{i \in S} X_i \geq 1 - \delta_{min}] = \Pr[\sum_{i \in S} X_i - (1 - 2^{-b}) \geq 1 - \delta_{min} - (1 - 2^{-b})] \leq e^{-2b(1-2^{-b} - \delta_{min})} = e^{-k(2^{-2b-1})}$$

where $X_{ch, resp} := 1 - P(ch, resp)$ and $S := S_{com_i}$. And (*) follows from Hoeffding’s inequality [21].

We thus have

$$\Pr[\forall i = 1 \ldots r, \text{Succ}_i] \geq 1 - 2^{-\ell_{com} r} - r e^{-k(2^{-2b-1})} =: p_s$$

Since $r$ is polynomially bounded and $b$ is logarithmic and $\ell_{com}, k$ are superpolynomial, $p_s$ is overwhelming.

For adversary success, it remains to show that $P_A \geq p_s$ where $P_A$ is as in (60). For this, we show that $\forall i. \text{Succ}_i$ implies $\forall i. ak_i = 1$ and $\sigma \leq S \land s = s_0$. First, note that $s = s_0$ always holds by definition of $A_1$. Furthermore, $\forall i. \text{Succ}_i$ implies (by definition of $P_i$) that

$$\sigma = \sum_i H(s, (com^*_i)_s, i, ch_i, resp^i)$$

$$= \sum_i H\left(s, (com^*_i)_s, i, ch_i, (resp_i, \text{COMopen}^*(ch_i, resp_i))\right) = \sum_i 0 \leq S.$$

Finally, if $\text{Succ}_i$ holds, then $(ch_i, resp_i) \in S_{com_i}$, thus

$$\text{COM\_verify}(e^i_{ch_i}, resp_i, \text{COMopen}^*(ch_i, resp_i)) = \text{COM\_verify}(\text{COM}^*(com_i), resp_i, \text{COMopen}^*(ch_i, resp_i)) = 1.$$
And \( O_V(com_i, ch_i, resp_i) = 1 \). Thus \( ok_i = V(com^*_i, ch_i, resp^*_i) = 1 \). Summarizing, \( \forall i. \text{Succ}_i \) implies \( \forall i. ok_i = 1 \land \sigma \leq S \land s = s_0 \) and thus \( P_A \geq p_s \). Since \( p_s \) is overwhelming, so is \( P_A \), thus we have adversary success.

**Extractor failure:** Extractor failure was already shown in the proof of Lemma 20 (\( A_1 \) here is defined exactly as in the proof of Lemma 20 and the definition of extractor failure depends only on \( A_1 \), not on \( A_2 \) or the protocol being attacked.)

Note that we have actually even shown extractor failure in the case that the extractor is allowed to choose the random oracle \( H \) before and during the execution of \( A_1 \), because \( A_1 \) does not access \( H \). \( \square \)

Now Theorem 29 follows from Lemma 19 (Security of the sigma-protocol) and Lemma 51. (The fact that Fischlin’s construction is a classical argument of knowledge is shown in [17].)

### I.2 Proofs for Theorem 30

**Lemma 52 (Attack on Fischlin’s construction, computational)** Then there exists a total break (Definition 2) against Fischlin’s construction based on the sigma-protocol from Definition 22 (Sigma-protocol, computational).

**Proof.** By Definition 2 (specialized to the case of Fischlin’s construction based on the sigma-protocol from Definition 22), we need to construct a polynomial-time adversary \( A_1, A_2 \) such that:

\[
P_A := \Pr \left[ \forall i. \text{ok}_i = 1 \land \sigma \leq S \land s = s_0 \land s \notin L_{R'} : s \leftarrow A_1^{H,O_{all}}, \right.

\( (com^*_i, ch_i, resp^*_i)_{i=1}^{r} \leftarrow A_2^{H,O_{all}}, \right. \text{ok}_i := V(com^*_i, ch_i, resp^*_i), \right.

\( \sigma := \sum_{i=1}^{r} H(x, (com^*_i)_i, i, ch_i, resp^*_i) \right] \text{ is overwhelming.}
\]

Here \( V \) is the verifier of the sigma-protocol (Definition 22).

We use the same adversary \( (A_1, A_2) \) as in the proof of Lemma 51 (Attack on Fischlin’s construction). Then \( P_A \) here is the same as \( P_A \) in the proof of Lemma 20 (Here we additionally have the condition \( s \notin L_{R'} \), but this condition is vacuously true since \( R' = \varnothing \) and thus \( L_{R'} = \varnothing \).) And in the proof of Lemma 51 we showed that \( P_A \) is overwhelming. \( \square \)

Now Theorem 30 follows from Lemma 23 (Security of the sigma-protocol, computational) and Lemma 52. (The fact that Fischlin’s construction is a classical argument of knowledge is shown in [17].)

---

11 Actually, [17] requires perfect completeness instead of completeness as defined here (we allow a negligible error). However, it is straightforward to see that their proof works unmodified for completeness as defined here.

12 Actually, [17] requires perfect special soundness instead of computational special soundness, as well as perfect completeness instead of completeness as defined here (we allow a negligible error). However, it is straightforward to see that their proof works unmodified for computational special soundness and completeness as defined here.