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

October 19, 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

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 [31], this would affect most public key encryption and signature schemes is use today. For symmetric ciphers and hash functions, longer key and output lengths will be required due to considerable improvements in brute force attacks [23, 12]. 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 [7] 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,[1] one

---

[1] Reminder: a proof or proof system is a protocol where a prover convinces a verifier of the validity of a statement s. It is zero-knowledge if the view of the verifier can be simulated without knowing a witness

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 [39]. Watrous [37] 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 [22] is an open problem.) Unruh [32] 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 [32] 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 [22]) 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 [8, 41, 35, 10, 34], 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?

Our contribution. 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 [32] is necessary. (Theorem 16)
- In the computational setting, sigma-protocols are not necessarily quantum arguments, even if they are classical arguments. (Theorem 20)
- The Fiat-Shamir construction [19] 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 25 and 26)
- Fischlin's non-interactive proof of knowledge in the random oracle model [20] 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 28 and 29)
- Besides proof systems, we also have negative results for commitment schemes.

---

w for the statement s (i.e., the verifier learns nothing about the witness). A proof of knowledge is a proof which additionally convinces the verifier that the prover could provide a witness w (i.e., not just the mere existence of w is proven). Arguments and arguments of knowledge are like proofs and proofs of knowledge, except that they are secure only against computationally limited provers.

Underlying sigma-protocol			Sig.-pr. used directly		Fiat-Shamir		Fischlin	
zero-knowledge	special soundness	strict soundness	PoK	proof	PoK	proof	PoK	proof
stat	perf	comp	attack[16]	stat[37]	attack[25]	?	attack[28]	?
stat	comp	comp	attack[20]	attack[20]	attack[26]	attack[26]	attack[29]	attack[29]
stat	perf	perf	stat[32]	stat[37]	?	?	?	?

Figure 1: Taxonomy of proofs of knowledge. For different combinations of security properties of the underlying sigma-protocol (statistical (stat)/perfect (perf)/computational (comp)), is there an attack in the quantum setting (relative to an oracle)? Or do we get a statistically/computationally secure proof/proof of knowledge (PoK)? The superscripts refer to theorem numbers in this paper or to literature references. Note that in all cases, classically we have at least computational security.

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 12)

- 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 [32], using the computational variant of this property leads to schemes that are not even computationally secure. (Theorems 16, 20, 25, 26, 28, and 29.)

Figure 1 gives an overview of the results relating to proofs of knowledge. Our main result are the separations listed in the bullet points above. Towards that goal, we additionally develop two 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: 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 $\mathcal{O}_\Psi$ in Section 4 is an example for this.)

Related work. Van der Graaf [36] first noticed that security definitions based on rewinding might be problematic in the quantum setting. Watrous [37] showed how the problems with quantum rewinding can be solved for a large class of zero-knowledge proofs. Unruh [32] 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 [32] to hold. Protocols that are secure classically but insecure in the quantum setting were constructed before: [40] presented classically secure pseudorandom functions that become insecure when the adversary is not only quantum, but can also *query the pseudorandom function in superposition.* Similarly for secret sharing schemes [16] and one-time MACs [9]. 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). We believe that this is the first such separation. Boneh, Dagdelen, Fischlin, Lehmann, Schaffner, and Zhandry [8] 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 [8]), an impossibility result was given by Dagdelen, Fischlin, and Gagliardoni [14]. However, their impossibility only shows that security of Fiat-Shamir cannot be shown using extractors that do not perform quantum rewinding;[2] but such quantum rewinding is possible and used in the existing positive results from [37, 32] which would also not work in a model without quantum rewinding. A *variant of* Fiat-Shamir has been shown to be a quantum secure signature scheme [14]. Probably their scheme can also be shown to be a quantum zero-knowledge proof of knowledge.[3] However, their construction assumes sigma-protocols with "oblivious commitments". These are a much stronger assumption that usual sigma-protocols: as shown in [33, Appendix A], sigma-protocols with oblivious commitments are by themselves already non-interactive zero-knowledge proofs in the CRS model (albeit single-theorem, non-adaptive ones). [33] presents a non-interactive quantum zero-knowledge proof of knowledge in the random oracle model, based on arbitrary sigma-protocols (it does not even need strict soundness). That protocol uses ideas different from both Fiat-Shamir and Fischlin's scheme to avoid rewinding.

It was known for a long time that it is difficult to use classical definitions for computational binding in the quantum setting ([17] is the first reference we are aware of), but none showed so far that the computational definition was truly insufficient.

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

---

[2]They 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.

[3]The unforgeability proof from [14] is already almost a proof of the proof of knowledge property. And the techniques from [33] can probably be applied to show that the protocol form [14] is zero-knowledge.

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 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 $\mathcal{O}_F$ with $\mathcal{O}_F|\Psi\rangle = -|\Psi\rangle$ and $\mathcal{O}_F|\Psi^\perp\rangle = |\Psi^\perp\rangle$ 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 $(ch, resp)$ 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 $\mathcal{O}_\Psi$ 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 $(ch, resp) \in S$ such that $ch = ch'$. Thus sending *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 $\mathcal{O}_F$, it is hard to compute a witness. Given two accepting conversations $(com, ch, resp)$ and $(com, ch', resp')$ we can compute a witness. So we need that given $|\Psi\rangle$ and $\mathcal{O}_F$, 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 5, Corollary 8). 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 [32]. 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 $\mathcal{O}_F$.

The main technical challenge is to prove that given $|\Psi\rangle$ and $\mathcal{O}_F$, 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 [2, 3]. 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 $\mathcal{H}_A$ that contains the algorithm's quantum state and another register $\mathcal{H}_I$ that contains the information which triples $(com, ch, resp)$ 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 $\mathcal{H}_A$ and $\mathcal{H}_I$. (From a practical perspective, a quantum register $\mathcal{H}_I$ 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 $\mathcal{H}_I$ into subspaces corresponding to group representations of the symmetry group of $\mathcal{H}_I$ (the set of all permutations of triples $(com, ch, resp)$ 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 $\mathcal{H}_I$ 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 $\mathcal{H}_I$ 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 $\mathcal{O}_\Psi$ 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 $\mathcal{O}_\Psi$ 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 $x \in S$).

On the meaning of oracle separations. At this point, we should say a few words about what it implies that our impossibility results are relative to a certain oracle. Certainly, our results do not necessarily imply that the investigated schemes are insecure or unprovable in the "real world", i.e., without oracles. However, our results give a number of valuable insights. Foremost, they tell us which proof techniques cannot be used for showing security of those schemes: only non-relativizing proofs can work. This cuts down the search space for proofs considerable. Also, it shows that security proofs would need new techniques; the proof techniques from [37, 32] at least are relativizing. And even non-relativizing proof techniques such as (in the classical setting) [5] tend to use specially designed (and more complicated) protocols than their relativizing counterparts,

so our results might give some evidence that the specific protocols we investigate here have no proofs at all, whether relativizing or non-relativizing. Furthermore, oracle-based impossibilities can give ideas for non-oracle-based impossibilities. If we can find computational problems that exhibit similar properties as our oracles, we might get analogous impossibilities without resorting to oracles (using computational assumptions instead).[4] However, we should stress that even if we get rid of the oracles, our results do not state that *all* sigma-protocols lead to insecure schemes. It would not be excluded that, e.g., the graph-isomorphism sigma-protocol [22] is still a proof of knowledge. What our approach aims to show is the impossibility of *general* constructions that are secure for *all* sigma-protocols.

Finally, we mention one point that is important in general when designing oracle separations in the quantum world: even relative to an oracle, the structural properties of quantum circuits should not change. For example, any quantum algorithm (even one that involves intermediate measurements or other non-unitary operations) can be replaced by a unitary quantum circuit, and that unitary circuit can be reversed. If we choose oracles that are not reversible, then we lose this property. (E.g., oracles that perform measurements or that perform random choices are non-reversible.) So an impossibility result based on such oracles would only apply in a world where quantum circuits are not reversible. Thus for meaningful oracle separations, we need to ensure that: (a) all oracles are unitary, and (b) all oracles have inverses. This makes some of the definitions of oracles in our work (Definition 7) more involved than would be necessary if we had used non-unitary oracles.

2 Preliminaries

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

Misc. $x \xleftarrow{\$} M$ 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., [28]. A

---

[4]For example, [1] presents a construction that might allow to implement an analogue to the oracle $\mathcal{O}_F$. Essentially, if the set S (called A in [1]) is a linear code, then they give a candidate for how to obfuscate $\mathcal{O}_F$ (called V_A in [1]) such that one can apply $\mathcal{O}_F$ but does not learn A. Of course, this does not give us a candidate for how to construct the other oracles needed in this work, but it shows that the idea of actually replacing our custom made oracles by computational assumptions may not be far fetched.

quantum state is a vector of norm 1 in a Hilbert space, written $|\Psi\rangle$. Then $\langle\Psi|$ is its dual. $\mathrm{TD}(\rho, \rho')$ denotes the *trace distance* between mixed states ρ, ρ'. We write short $\mathrm{TD}(|\Psi\rangle, |\Psi'\rangle)$ for $\mathrm{TD}(|\Psi\rangle\langle\Psi|, |\Psi'\rangle\langle\Psi'|)$. $\mathrm{SD}(X; Y)$ 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 $\mathcal{O}$ is a unitary transformation on some Hilbert space $\mathcal{H}$. An oracle algorithm A with access to $\mathcal{O}$ (written $A^{\mathcal{O}}$) is then a quantum algorithm which has a special gate for applying the unitary $\mathcal{O}$. $\mathcal{O}$ may depend on the security parameter. $\mathcal{O}$ may be probabilistic in the sense that at the beginning of the execution, the unitary $\mathcal{O}$ is chosen according to some distribution (like the random oracle in cryptography). However, $\mathcal{O}$ may not be probabilistic in the sense that $\mathcal{O}$, 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 $\mathcal{O}$ in superposition. When defining $\mathcal{O}$, we use the shorthand $\mathcal{O}(x) := f(x)$ to denote $\mathcal{O}|x, y\rangle := \mathcal{O}|x, y \oplus f(x)\rangle$. 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 $(\mathcal{O}_1, \mathcal{O}_2, \dots)$ 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 $ch \xleftarrow{\$} \{0, 1\}^{\ell_{ch}}$ (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 $(s, w) \in R$, $\Pr[ok = 0 : com \leftarrow P_1(s, w), ch \xleftarrow{\$} \{0, 1\}^{\ell_{ch}}, resp \leftarrow P_2(ch), ok \leftarrow V(s, com, ch, resp)]$ is negligible.*

(Intuitively: an honestly generated proof succeeds for overwhelming probability.)

- **Perfect special soundness:** *There is a polynomial-time classical algorithm E_Σ (the extractor) such that for any $(s, com, ch, resp, ch', resp')$ with $ch \neq ch'$, we have that $\Pr[(s, w) \notin R \wedge ok = ok' = 1 : ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch', resp'), w \leftarrow E_\Sigma(s, com, ch, resp, ch', resp')] = 0$.*
  (Intuitively: given two valid interactions with the same commitment, one can efficiently extract a witness.)

- **Computational special soundness:** *There is a polynomial-time classical algorithm E_Σ (the extractor) such that for any polynomial-time quantum algorithm A (the adversary), we have that $\Pr[(s, w) \notin R \wedge ch \neq ch' \wedge ok = ok' = 1 : (s, com, ch, resp, ch', resp') \leftarrow A, ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch', resp'), w \leftarrow E_\Sigma(s, com, ch, resp, ch', resp')]$ is negligible.*
  (Intuitively: given two valid interactions with the same commitment chosen by a polynomial-time adversary, one can efficiently extract a witness with overwhelming probability.)

- **Statistical honest-verifier zero-knowledge (HVZK):**[5] *There is a polynomial-time classical algorithm S_Σ (the simulator) such that for any (possibly unlimited) quantum algorithm A and all $(s, w) \in R$, the following is negligible:*

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

  (Intuitively: An interaction between honest verifier and honest prover can be simulated in polynomial-time without knowing the witness.)

- **Strict soundness:** *For any (s, com, ch) and any $resp \neq resp'$ we have $\Pr[ok = ok' = 1 : ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch, resp')] = 0$.*
  (Intuitively: Given the commitment and the challenge, there is at most one possible accepted response.)

- **Computational strict soundness:**[6] *For any polynomial-time quantum algorithm A (the adversary), we have that $\Pr[ok = ok' = 1 \wedge resp \neq resp' : (s, com, ch, resp, resp') \leftarrow A, ok \leftarrow V(s, com, ch, resp), ok' \leftarrow V(s, com, ch, resp')]$ is negligible.*
  (Intuitively: Given the commitment and the challenge, it is computationally hard to find more than one accepting response.)

- **Commitment entropy:** *For all $(s, w) \in R$ and $com \leftarrow P_1(s, w)$, the min-entropy of com is superlogarithmic.*
  (Intuitively: the commitment produced by the prover cannot be guessed with more than negligible probability.)

In a relativized setting, all quantum algorithms additionally get access to all oracles, and all classical algorithms additionally get access to all classical oracles.

---

[5]In the context of this paper, HVZK is equivalent to zero-knowledge because our protocols have logarithmic challenge length ℓ_{ch} [37].

[6]Also known as *unique responses* in [20].

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:*

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

Here $\langle A, V(s) \rangle$ denotes the output of V in an interaction between A and $V(s)$. (Intuitively, the adversary performs a total break if the adversary manages with overwhelming probability to convince the verifier V of a statement s that is not in the language L_R.)

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: $\Pr[(s, w) \in R : s \leftarrow A, w \leftarrow E(s)]$ is negligible.*

Here E has access to the final state of A. (Intuitively, the adversary performs a total knowledge break if the adversary manages with overwhelming probability to convince the verifier V of a statement s, but the extractor E cannot extract a witness w for that statement.)

When applied to a proof system relative to an oracle $\mathcal{O}$, both A and E get access to $\mathcal{O}$. In settings where R and $\mathcal{O}$ are probabilistic, the probabilities are averaged over all values of R and $\mathcal{O}$.

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.[7]

3 State creation oracles

We first show a result that shows that having access to an oracle $\mathcal{O}_\Psi$ 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 $|\bot\rangle$). (Such an oracle is, in our setting, implemented as $\mathcal{O}_\Psi |\Psi\rangle = |\bot\rangle$, $\mathcal{O}_\Psi |\bot\rangle = |\Psi\rangle$, and is the identity on states orthogonal to $|\bot\rangle, |\Psi\rangle$.) We will need this later, because it allows

---

[7]Definitions 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.

us to assume in our proofs that the adversary has access to such a reservoir state instead of access to the oracle $\mathcal{O}_\Psi$. It turns out to be much easier to show that those reservoir states do not help the adversary in solving the Two Values problem than it is to deal directly with $\mathcal{O}_\Psi$ in the proof.

Note that the fact that $\mathcal{O}_\Psi$ is no more powerful than $|R\rangle$ is not immediate: $\mathcal{O}_\Psi$ can be queried in superposition, and its inverse applied; this might give more power than copies of the state $|\Psi\rangle$. In fact, we know of no way to generate, e.g., $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|\bot\rangle$ for a given (known) state $|\bot\rangle$ 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 $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|\bot\rangle$ can be generated with a single query to $\mathcal{O}_\Psi$.[8] 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 $|\bot\rangle$ 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 $|\bot\rangle$.) Let $\mathcal{O}_\Psi$ be an oracle with $\mathcal{O}_\Psi|\Psi\rangle = |\bot\rangle$, $\mathcal{O}_\Psi|\bot\rangle = |\Psi\rangle$, and $\mathcal{O}_\Psi|\Psi^\bot\rangle = |\Psi^\bot\rangle$ for any $|\Psi^\bot\rangle$ orthogonal to both $|\Psi\rangle$ and $|\bot\rangle$. Let $\mathcal{O}$ be an oracle, not necessarily independent of $|\Psi\rangle$. Let $|\Phi\rangle$ be a quantum state, not necessarily independent of $|\Psi\rangle$. Let $n, m \geq 0$ be integers. Let $|R\rangle := |\Psi\rangle^{\otimes m} \otimes |\alpha_1\rangle \otimes \cdots \otimes |\alpha_n\rangle$ where $|\alpha_j\rangle := (\cos\frac{j\pi}{2n})|\Psi\rangle + (\sin\frac{j\pi}{2n})|\bot\rangle$.*

Let A be an oracle algorithm that makes q_Ψ queries to $\mathcal{O}_\Psi$. Then there is an oracle algorithm B that makes the same number of queries to $\mathcal{O}$ as A such that:

$$\mathrm{TD}\big(B^{\mathcal{O}}(|R\rangle, |\Phi\rangle)), A^{\mathcal{O}_\Psi, \mathcal{O}}(|\Phi\rangle)\big)$$
$$\leq \quad \frac{\pi q_\Psi}{2\sqrt{n}} + q_\Psi\, o\big(\tfrac{1}{\sqrt{n}}\big) + \frac{2q_\Psi}{\sqrt{m+1}} \quad \leq \quad O\Big(\frac{q_\Psi}{\sqrt{n}} + \frac{q_\Psi}{\sqrt{m}}\Big).$$

The idea behind this lemma is the following: To implement $\mathcal{O}_\Psi$, we need a way to convert $|\bot\rangle$ 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 $|\bot\rangle$, 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 $\mathcal{O}_\Psi$ 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 $\mathcal{O}_\Psi$ 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 $|\bot\rangle$ to the reservoir. Such a state would be $|R^\infty\rangle := |\Psi\rangle^{\otimes\infty} \otimes |\bot\rangle^{\otimes\infty}$. If we add or remove $|\Psi\rangle$ to an infinite state $|\Psi\rangle^{\otimes\infty}$, that state will not change. Similarly for $|\bot\rangle$. (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.)

---

[8] For example, one can initialize a register with $\frac{1}{\sqrt{2}}|\bot\rangle + \frac{1}{\sqrt{2}}|0\rangle$ where $|0\rangle$ is any fixed state guaranteed to be (almost) orthogonal to $|\bot\rangle$ and $|\Psi\rangle$. Applying $\mathcal{O}_\Psi$ yields $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|0\rangle$. Finally, by applying the fixed (and thus known) unitary $U: |0\rangle \mapsto |\bot\rangle$, we get $\frac{1}{\sqrt{2}}|\Psi\rangle + \frac{1}{\sqrt{2}}|\bot\rangle$.

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 $\mathcal{O}_\Psi$? Indeed, an elementary calculation reveals that the following circuit implements $\mathcal{O}_\Psi$ on X when R, Z are initialized with $|R^\infty\rangle, |0\rangle$.
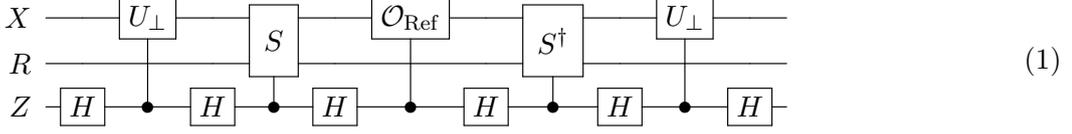
$$\text{with } U_\perp := 1 - 2|\perp\rangle\langle\perp| \tag{2}$$
$$\text{and } \mathcal{O}_{\text{Ref}} := 1 - 2|\Psi\rangle\langle\Psi| \tag{2}$$

Note that we have introduced a new oracle $\mathcal{O}_{\text{Ref}}$ 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 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 $\cdots \otimes |\perp\rangle \otimes |\perp\rangle \otimes |\perp\rangle$. 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 $\mathcal{O}_\Psi$ when R, Z are initialized with $|R\rangle, |0\rangle$.

However, we now have introduced the oracle $\mathcal{O}_{\text{Ref}}$. We need to show how to emulate that oracle: $\mathcal{O}_{\text{Ref}}$ essentially implements a measurement whether a given state $|\Phi\rangle$ is $|\Psi\rangle$ or orthogonal to $|\Psi\rangle$. Thus to implement $\mathcal{O}_{\text{Ref}}$, we need a way to test whether a given state is $|\Psi\rangle$ or not. The well-known swap test [13] is not sufficient, because for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$, 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 $|\Psi\rangle$ as reference: Let $|T\rangle := |\Psi\rangle^{\otimes m}$. Let V be the space of all $(m+1)$-partite states that are invariant under permutations. $|\Psi\rangle|T\rangle$ is such a state, while for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$, $|\Phi\rangle|T\rangle$ is almost orthogonal to V for large m (up to an error of $O(1/m)$). So by measuring whether $|\Phi\rangle|T\rangle$ is in V, we can test whether $|\Phi\rangle$ is $|\Psi\rangle$ or not (with an error $O(1/m)$), and when doing so the state $|T\rangle$ 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 $|T\rangle$.

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

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

4 The pick-one trick

In this section, we first show 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 4 (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 α 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 4, he will be able to search for one $x \in S_y$ satisfying a predicate P (shown in Theorem 6 below). But he will not be able to find two different $x, x' \in S_y$ (Theorem 5 below); we will use this to foil any attempts at extracting by rewinding.

Theorem 5 (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*

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

That is, in order to get a constant success probability in finding x_1, x_2, one would need at least $h \in \Omega(\sqrt{|Y|})$ copies of the state $|\Psi\rangle$, or make $\Omega(\min\{\sqrt{k}, \sqrt{|X|/k}\})$ queries. Or to put it differently, if $\sqrt{k}$ and $\sqrt{|X|/k}$ are both superpolynomial, a polynomial-time adversary (who necessarily has polynomially-bounded h, q_V, q_F) finds x_1, x_2 only with negligible probability.

The proof uses the adversary-method from [2, 3] 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 6 (Searching one value) *Let $S_y \subseteq X$ and $\mathcal{O}_F, \mathcal{O}_V$ be as in Definition 4.*
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

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

(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 [23]: 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 ℓ, but instead over all $x \in S_y$. When searching over S_y, the initial state of Grover's algorithm needs to be $\sum_x \frac{1}{\sqrt{|S_y|}}|x\rangle = |\Psi(y)\rangle$ instead of $\sum_x 2^{-\ell/2}|x\rangle =: |\Phi\rangle$. 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 [11].) 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 in case of failure is also a suitable initial state for Grover's algorithm. 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 7 (Oracle distribution) *Fix integers $\ell_{com}, \ell_{ch}, \ell_{resp}$ (that may depend on the security parameter) such that ℓ_{com}, ℓ_{resp} are superlogarithmic and ℓ_{ch} is logarithmic. Let $\ell_{rand} := \ell_{com} + \ell_{resp}$.*

Let $\mathcal{O}_{all} = (\mathcal{O}_E, \mathcal{O}_P, \mathcal{O}_R, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V)$ be chosen according to the following distribution:

- *Let s_0 be arbitrary but fixed (e.g., $s_0 := 0$). Pick $w_0 \xleftarrow{\$} \{0,1\}^{\ell_{rand}}$.*
- *Choose $S_y, \mathcal{O}_V, \mathcal{O}_F$ as in Definition 4 with $Y := \{0,1\}^{\ell_{com}}$ and $X := \{0,1\}^{\ell_{ch}} \times \{0,1\}^{\ell_{resp}}$ and $k := 2^{\ell_{ch} + \lfloor \ell_{resp}/3 \rfloor}$.*
- *For each $z \in \{0,1\}^{\ell_{rand}}$, pick $y \xleftarrow{\$} Y$ and $x \xleftarrow{\$} S_y$, and set $\mathcal{O}_S(z) := (y, x)$.*
- *Let $|\bot\rangle$ be a quantum state orthogonal to all $|com, ch, resp\rangle$ (i.e., we extend the dimension of the space in which $|\Sigma\Psi\rangle$ lives by one). $\mathcal{O}_\Psi|\bot\rangle := |\Sigma\Psi\rangle$, $\mathcal{O}_\Psi|\Sigma\Psi\rangle := |\bot\rangle$, and $\mathcal{O}_\Psi|\Phi\rangle := |\Phi\rangle$ for $|\Phi\rangle$ orthogonal to $|\Sigma\Psi\rangle$ and $|\bot\rangle$.*
- *Let $\mathcal{O}_E(com, ch, resp, ch', resp') := w_0$ iff $(ch, resp), (ch', resp') \in S_{com} \wedge (ch, resp) \neq (ch', resp')$ and $\mathcal{O}_E := 0$ everywhere else.*
- *Let $\mathcal{O}_R(s_0, w_0) := 1$ and $\mathcal{O}_R := 0$ everywhere else.*
- *For each $com \in \{0,1\}^{\ell_{com}}, ch \in \{0,1\}^{\ell_{ch}}, z \in \{0,1\}^{\ell_{rand}}$, let $\mathcal{O}_P(w_0, com, ch, z)$ be assigned a uniformly random $resp$ with $(ch, resp) \in S_{com}$. (Or $\bot$ if no such $resp$ exists.) Let $\mathcal{O}_P(w, \cdot, \cdot, \cdot) := 0$ for $w \neq w_0$.*

From these oracles, $\mathcal{O}_P$ is later used to implement the prover in our sigma-protocols, $\mathcal{O}_E$ for the extractor, $\mathcal{O}_R$ to test membership in the relation R, and $\mathcal{O}_S$ to implement the simulator. Notice that $\mathcal{O}_S$ and $\mathcal{O}_P$ get an additional input z that seems useless. However, z is needed to get several independent answers from the oracle given otherwise equal inputs (i.e., it emulates probabilistic behavior).

Note that both the relation R and the oracles are chosen randomly (but not independently of each other). We will assume this implicitly in all further theorems. We could also get a result relative to a fixed (i.e., non-probabilistic) relation and oracle by using the probabilistic method. We omit the details from this work.

The following corollary is a strengthening of Theorem 5 to the oracle distribution from Definition 7. 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 (50–57) in the proof.

Corollary 8 (Hardness of two values 2) *Let* $\mathcal{O}_{all}$ *=* $(\mathcal{O}_E, \mathcal{O}_P, \mathcal{O}_R, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V), w_0$ *be as in Definition 7. Let A be an oracle algorithm making at most $q_E, q_P, q_R, q_S, q_F, q_\Psi, q_V$ queries to $\mathcal{O}_E, \mathcal{O}_P, \mathcal{O}_R, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V$, respectively. Assume that $q_E, q_P, q_R, q_S, q_F, q_V$ are polynomially-bounded (and ℓ_{com}, ℓ_{resp} are superlogarithmic by Definition 7). Then:*

(i) $\Pr[w = w_0 : w \leftarrow A^{\mathcal{O}_{all}}]$ *is negligible.*

(ii) $\Pr[(ch, resp) \neq (ch', resp') \wedge (ch, resp), (ch', resp') \in S_{com} : (com, ch, resp, ch', resp') \leftarrow A^{\mathcal{O}_{all}}]$ *is negligible.*

This corollary is shown by reduction to Theorem 5 (Hardness of the two values problem). Given an adversary that violates (i), we remove step by step the oracles that are not covered by Theorem 5. First, we remove the oracles $\mathcal{O}_P, \mathcal{O}_R$. Those do not help the adversary (much) to find w_0 because $\mathcal{O}_P$ and $\mathcal{O}_R$ only give non-zero output if their input already contains w_0. Next we change A to output a collision $(ch, resp) \neq (ch', resp') \wedge (ch, resp), (ch', resp') \in S_{com}$ instead of the witness w_0; since w_0 can only be found by querying $\mathcal{O}_E$ with such a collision, this adversary succeeds with non-negligible probability, too. Furthermore, A then does not need access to $\mathcal{O}_E$ any more since $\mathcal{O}_E$ only helps in finding w_0. Next we get rid of $\mathcal{O}_\Psi$: as shown in Theorem 3 (Emulating state creation oracles), $\mathcal{O}_\Psi$ can be emulated (up to an inversely polynomial error) using (suitable superpositions on) copies of the state $|\Sigma\Psi\rangle$. Finally we remove $\mathcal{O}_S$: Using the "small range distribution" theorem from [40], $\mathcal{O}_S$ 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 $\mathcal{O}_F, \mathcal{O}_V$ 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 5.

And (ii) is shown by observing that an adversary violating (i) leads to one violating (ii) using one extra $\mathcal{O}_E$-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 9 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 $\mathrm{COM}, \mathrm{COM}_{verify}$, such that $(c, u) \leftarrow \mathrm{COM}(m)$ returns a commitment c on the message m, and an opening information u. The sender then sends c to the recipient, who is not supposed to learn anything about m. Only when the sender later sends m, u, the recipients learns m. But, intuitively speaking, the sender should not be able to "change his mind" about m after sending c (binding property). We require *perfect completeness*, i.e., for any m and $(c, u) \leftarrow \mathrm{COM}(m)$, $\mathrm{COM}_{verify}(c, m, u) = 1$ with probability 1. In our setting, c, m, u are all classical.

Definition 9 (Computationally binding) *A commitment scheme* $\mathrm{COM}, \mathrm{COM}_{verify}$ *is* computationally binding *iff for any quantum polynomial-time algorithm A the following probability is negligible:*

$$\Pr[ok = ok' = 1 \ \wedge \ m \neq m' : (c, m, u, m', u') \leftarrow A,$$
$$ok \leftarrow \mathrm{COM}_{verify}(c, m, u), \ ok \leftarrow \mathrm{COM}_{verify}(c, m', u')]$$

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

[32] 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). The results from [32] show that strict binding commitments can behave better under rewinding, so perhaps strict binding commitments can avoid the problems that merely binding commitments have? We define a computational variant of this property here:

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

$$\Pr[ok = ok' = 1 \ \wedge \ (m, u) \neq (m', u') : (c, m, u, m', u') \leftarrow A,$$
$$ok \leftarrow \mathrm{COM}_{verify}(c, m, u), \ ok \leftarrow \mathrm{COM}_{verify}(c, m', u')]$$

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

Definition 11 (Statistically hiding) *A commitment scheme* $\mathrm{COM}, \mathrm{COM}_{verify}$ *is* statistically hiding *iff for all* m_1, m_2 *with* $|m_1| = |m_2|$ *and* $c_i \leftarrow \mathrm{COM}(m_i)$ *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 12 (Insecurity of binding commitments) *There is an oracle $\mathcal{O}$ and a non-interactive commitment scheme $\mathrm{COM}, \mathrm{COM}_{verify}$ 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,*

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

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_{com} from Definition 7, we will call them S_y and we refer to their members as $x \in S_y$. (Not $(ch, resp) \in S_{com}$.) In particular, oracles such as $\mathcal{O}_S$ will returns pairs (y, x), not triples $(com, ch, resp)$, etc.

We construct a commitment scheme relative to the oracle $\mathcal{O}_{all}$ from Definition 7. (Note: that oracle distribution contains more oracles than we need for Theorem 12. However, we will need in later sections that our commitment scheme is defined relative to the same oracles as the proof systems there.)

Definition 13 (Bad commitment scheme) *Let $\mathsf{bit}_i(x)$ denote the i-th bit of x. We define $\mathrm{COM}, \mathrm{COM}_{verify}$ as follows:*
- *$\mathrm{COM}(m)$: For $i = 1, \ldots, |m|$, pick $z_i \overset{\$}{\leftarrow} \{0,1\}^{\ell_{rand}}$ and let $(y_i, x_i) := \mathcal{O}_S(z_i)$. Let $p_i \overset{\$}{\leftarrow} \{1, \ldots, \ell_{ch} + \ell_{resp}\}$. Let $b_i := m_i \oplus \mathsf{bit}_{p_i}(x_i)$. 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).*
- *$\mathrm{COM}_{verify}(c, m, u)$ 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 $\mathcal{O}_V(y_i, x_i) = 1$ for $i = 1, \ldots, n$. Check whether $b_i = m_i \oplus \mathsf{bit}_{p_i}(x_i)$ 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 14 (Properties of COM) *The scheme from Definition 13 is perfectly complete, computationally binding, computationally strict binding, and statistically hiding. (Relative to $\mathcal{O}_{all}$.)*

The computational binding and computational strict binding property are a consequence of Corollary 8 (Hardness of two values 2): 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 8 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 $\mathsf{bit}_{p_i}(x_i)$ is almost independent of y_i.

The proof is given in Section F.1.

Lemma 15 (Attack on COM**)** *There is a quantum polynomial-time adversary B_1, B_2 such that for all m,*

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

is overwhelming.

Basically, the adversary B_1, B_2 commits to a random commitment. And to unveil to a message m, he needs to find values $x_i \in S_{y_i}$ with $\mathsf{bit}_{p_i}(x_i) = m_i \oplus b_i$. Since half of all x_i have this property, such x_i can be found using Theorem 6 (Searching one value).

The full proof is given in Appendix F.2.

Theorem 12 then follows immediately from Lemmas 14 and 15.

6 Attacking sigma-protocols

We will now show that in general, sigma-protocols with special soundness are not necessarily proofs of knowledge. [32] 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 [32] do not hold without strict soundness (not even with computational strict soundness), relative to an oracle.

Theorem 16 (Insecurity of sigma-protocols) *There is an oracle $\mathcal{O}_{all}$ and a relation R and a sigma-protocol relative to $\mathcal{O}_{all}$ with logarithmic ℓ_{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 $\mathcal{O}_{all}$ 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 ℓ_{ch} follows immediately by parallel repetition of the sigma-protocol.

The remainder of this section will prove Theorem 16. As a first step, we construct the sigma-protocol.

Definition 17 (Sigma-protocol) *Let* $\mathrm{COM}, \mathrm{COM}_{verify}$ *be the commitment scheme from Definition 13.*[9]

---

[9]The commitment described there has the property that it is computationally binding, but still it is possible for the adversary to open the commitment to any value, only not to several values at the same time. The commitment is defined relative to the same oracle distribution as the sigma-protocol here, which is why we can use it.

Relative to the oracle distribution from Definition 7, we define the following sigma-protocol $(\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R)$ *for the relation* $R := \{(s_0, w_0)\}$:

- $P_1(s, w)$ *picks* $com \xleftarrow{\$} \{0,1\}^{\ell_{com}}$. *For each* $ch \in \{0,1\}^{\ell_{ch}}$, *he picks* $z_{ch} \xleftarrow{\$} \{0,1\}^{\ell_{rand}}$ *and computes* $resp_{ch} := \mathcal{O}_P(w, com, ch, z_{ch})$ *and* $(c_{ch}, u_{ch}) \leftarrow \mathrm{COM}(resp_{ch})$. *Then* P_1 *outputs* $com^* := (com, (c_{ch})_{ch \in \{0,1\}^{\ell_{ch}}})$.
- $P_2(ch)$ *outputs* $resp^* := (resp_{ch}, u_{ch})$.
- *For* $com^* = (com, (c_{ch})_{ch \in \{0,1\}^{\ell_{ch}}})$ *and* $resp^* = (resp, u)$, *let* $V(s, com^*, ch, resp^*) := 1$ *iff* $\mathcal{O}_V(com, ch, resp) = 1$ *and* $s = s_0$ *and* $\mathrm{COM}_{verify}(c_{ch}, resp, u) = 1$.

The commitments c_{ch} are only needed to get computational strict soundness. A slightly weaker Theorem 16 without computational strict soundness can be achieved using the sigma-protocol from Definition 17 without the commitments c_{ch}; the proofs stay the same, except that the steps relating to the commitments are omitted.

Lemma 18 (Security of the sigma-protocol) *The sigma-protocol from Definition 17 has: completeness, perfect special soundness, computational strict soundness, statistical honest-verifier zero-knowledge, commitment entropy.*

Perfect special soundness follows from the existence of the oracle $\mathcal{O}_E$. 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 $\mathcal{O}_S$ 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 19 (Attack on the sigma-protocol) *Assume that* ℓ_{ch} *is logarithmically bounded. Then there exists a total knowledge break (Definition 2) against the sigma-protocol from Definition 17.*

To attack the sigma protocol, the malicious prover uses Theorem 6 (Searching one value) to get a com and a corresponding state $|\Psi(com)\rangle$. 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 $ch' = ch$ (ℓ_{ch} is logarithmic), this can be done with Theorem 6. 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 15 (Attack on COM).) However, an extractor that has the same information as the prover (namely, access to the oracle $\mathcal{O}_{all}$) will fail to find w_0 by Corollary 8 (Hardness of two values 2).

The full proof is given in Appendix G.2.

Now Theorem 16 follows from Lemmas 18 and 19. (The fact that the sigma-protocol is a classical proof of knowledge is shown in [15].)

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 20 (Insecurity of sigma-protocols, computational) *There is an oracle $\mathcal{O}_{all}$ and a relation R' and a sigma-protocol relative to $\mathcal{O}_{all}$ with logarithmic ℓ_{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 $\mathcal{O}_{all}$ with completeness, computational special soundness, and statistical honest-verifier zero-knowledge is a classical argument.

Note that a corresponding theorem with polynomially bounded ℓ_{ch} follows immediately by parallel repetition of the sigma-protocol. The remainder of this section is dedicated to proving Theorem 20.

Definition 21 (Sigma-protocol, computational) *We define a sigma-protocol $(\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R')$ as in Definition 17, except that the relation is $R' := \varnothing$.*

Lemma 22 (Security of the sigma-protocol, computational) *The sigma-protocol from Definition 21 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 18 (Security of the sigma-protocol). However, perfect special soundness does not hold for the sigma-protocol from Definition 21: There exist pairs of accepting conversations $(ch, resp), (ch', resp') \in S_{com}$. But these do not allow us to extract a valid witness for s_0 (because $R' = \varnothing$, so no witnesses exist). However, we have computational special soundness: by Corollary 8 (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 23 (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 21.*

In this lemma, we use the same malicious prover as in Lemma 19 (Attack on the sigma-protocol). That adversary proves the statement s_0. Since $R' = \varnothing$, 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 20 follows from Lemmas 22 and 23. (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 [15] for sigma-protocols with special soundness applies to this case, too.)

7 Attacking Fiat-Shamir

Definition 24 (Fiat-Shamir) *Fix a sigma-protocol $(\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R)$ and an integer $r > 0$. Let $H : \{0,1\}^* \to \{0,1\}^{r \cdot \ell_{ch}}$ 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 25 (Insecurity of Fiat-Shamir) *There is an oracle $\mathcal{O}_{all}$ and a relation R and a sigma-protocol relative to $\mathcal{O}_{all}$ 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 $r\ell_{ch}$ is superlogarithmic).

As the underlying sigma-protocol, we use the one from Definition 17. 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 26 (Insecurity of Fiat-Shamir, computational) *There is an oracle $\mathcal{O}_{all}$ and a relation R and a sigma-protocol relative to $\mathcal{O}_{all}$ with logarithmic ℓ_{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 25 and Lemma 23 and given in Appendix H.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 [20]. 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 27 (Fischlin's scheme) *Fix a sigma-protocol $(\ell_{com}, \ell_{ch}, \ell_{resp}, P_1, P_2, V, R)$. Fix integers $b, r, \mathbf{S}, t$ such that br and 2^{t-b} are superlogarithmic, b, r, t are logarithmic, $\mathbf{S} \in O(r)$ ($\mathbf{S} = 0$ is permitted), and $b \le t \le \ell_{ch}$.*
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 is defined as follows:
- *$P_{Fis}(s, w)$: See [20]. (Omitted here since we only need to analyze V_{Fis} for our results.)*
- *$V_{Fis}(s, \pi)$ with $\pi = (com_i, ch_i, resp_i)_{i=1,\dots,r}$: Check if $V(com_i, ch_i, resp_i) = 0$ for all $i = 1, \dots, r$. Check if $\sum_{i=1}^{r} H(x, (com_i)_i, i, ch_i, resp_i) \le \mathbf{S}$ (where $H(\dots)$ 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 $H(x, (com_i)_i, i, ch_i, resp_i)$ 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 28 (Insecurity of Fischlin's construction) *There is an oracle $\mathcal{O}_{all}$ and a relation R and a sigma-protocol relative to $\mathcal{O}_{all}$ 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 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 17. The basic idea is that the malicious prover finds conversations $(com_i^*, ch_i, resp_i^*)$ by first fixing the values com_i^*, and then using Theorem 6 to find $ch, resp^*$ where $resp_i^*$ contains $resp_i$ such that $(ch_i, resp_i) \in S_{com_i}$ and $H(x, (com_i^*)_i, i, ch_i, resp_i^*) = 0$. If $resp_i^*$ would not additionally contain commitments c_{ch} (see Definition 17), this would already suffice to break Fischlin'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 29 (Insecurity of Fischlin's construction, computational) *There is an oracle $\mathcal{O}_{all}$ and a relation R and a sigma-protocol relative to $\mathcal{O}_{all}$ with logarithmic ℓ_{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 28 and 29 even hold with strict soundness instead of computational strict soundness. We sketch our reasoning: Consider a variant of the oracle distribution from Definition 7, in which ℓ_{ch} is superlogarithmic (not logarithmic) and in which the sets S_{com} are chosen uniformly at random from all sets S which satisfy $\forall ch \exists_1 resp.(ch, resp) \in S$. Note that the results from Sections 5–7 do not hold in this setting, because ch must be polynomially-bounded to show the existence of successful adversaries. (Namely, when Theorem 6 (Searching one value) is invoked, the predicate P is true on a $2^{-\ell_{ch}}$ fraction of the all values.) But the proofs of Lemma 50 (Attack on Fischlin's construction) and Lemma 51 (Attack on Fischlin's construction, computational) do not require this. We conjecture that Corollary 8 still holds in this modified setting (the cardinality of the S_{com} satisfies the conditions of Corollary 8, but the S_{com} have additional structure). Then the sigma-protocols from Definitions 17 and 21 (without the commitments c_{ch}) will still have the properties shown in Lemmas 18 and 22, but additionally they will have strict soundness because for any com, ch, there exists only one $resp$ such that $(ch, resp) \in S_{com}$.

We leave the proof that Corollary 8 holds even for sets S_{com} with $\forall ch \exists_1 resp.(ch, resp) \in S_{com}$ as an open problem.

24

References

[1] Scott Aaronson and Paul Christiano. Quantum money from hidden subspaces. In *STOC '12*, pages 41–60. ACM, 2012.

[2] Andris Ambainis. A new quantum lower bound method, with an application to a strong direct product theorem for quantum search. *Theory of Computing*, 6(1):1–25, 2010.

[3] Andris Ambainis, Loïck Magnin, Martin Roetteler, and Jérémie Roland. Symmetry-assisted adversaries for quantum state generation. In *IEEE Conference on Computational Complexity*, pages 167–177. IEEE Computer Society, 2011.

[4] Andris Ambainis, Robert Špalek, and Ronald de Wolf. A new quantum lower bound method, with applications to direct product theorems and time-space tradeoffs. *Algorithmica*, 55(3):422–461, 2009.

[5] Boaz Barak. How to go beyond the black-box simulation barrier. In *FOCS 2001*, pages 106–115. IEEE, 2001.

[6] Daniel Berend and Aryeh Kontorovich. On the convergence of the empirical distribution. arXiv:1205.6711v2 [math.ST], 2012.

[7] Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors. *Post-Quantum Cryptography*. Springer, 2009.

[8] Dan Boneh, Özgür Dagdelen, Marc Fischlin, Anja Lehmann, Christian Schaffner, and Mark Zhandry. Random oracles in a quantum world. In *ASIACRYPT 2011*, pages 41–69, Berlin, Heidelberg, 2011. Springer-Verlag.

[9] Dan Boneh and Mark Zhandry. Quantum-secure message authentication codes. In *Eurocrypt 2013*, volume 7881 of *LNCS*, pages 592–608. Springer, 2013. Online version IACR ePrint 2012/606.

[10] Dan Boneh and Mark Zhandry. Secure signatures and chosen ciphertext security in a quantum computing world. In *Crypto 2013*, 2013. Full version at IACR ePrint 2013/088.

[11] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik*, 46(4-5):493–505, 1998. Online version at arXiv:quant-ph/9605034.

[12] Gilles Brassard, Peter Høyer, and Alain Tapp. Quantum algorithm for the collision problem. *ACM SIGACT News*, 28:14–19, 1997. Full version at arXiv:quant-ph/9705002.

[13] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87:167902, September 2001. Online version arXiv:quant-ph/0102001.

[14] Özgür Dagdelen, Marc Fischlin, and Tommaso Gagliardoni. The Fiat-Shamir transformation in a quantum world. In *Asiacrypt 2013*, volume 8270 of *LNCS*, pages 62–81. Springer, 2013. Online version IACR ePrint 2013/245.

[15] Ivan Damgård. On σ-protocols. Course notes for "Cryptologic Protocol Theory", `http://www.cs.au.dk/~ivan/Sigma.pdf`, 2010. Retrieved 2014-03-17. Archived at `http://www.webcitation.org/6O9USFecZ`.

[16] Ivan Damgård, Jakob Funder, Jesper Buus Nielsen, and Louis Salvail. Superposition attacks on cryptographic protocols. In *ICITS 2013*, volume 8317 of *LNCS*, pages 142–161. Springer, 2014. Online version IACR ePrint 2011/421.

[17] Paul Dumais, Dominic Mayers, and Louis Salvail. Perfectly concealing quantum bit commitment from any quantum one-way permutation. In *Eurocrypt '00*, volume 1807 of *LNCS*, pages 300–315, Berlin, Heidelberg, 2000. Springer.

[18] Sebastian Faust, Markulf Kohlweiss, Giorgia Azzurra Marson, and Daniele Venturi. On the non-malleability of the Fiat-Shamir transform. In Steven Galbraith and Mridul Nandi, editors, *INDOCRYPT 2012*, volume 7668 of *LNCS*, pages 60–79. Springer, 2012. Preprint on IACR ePrint 2012/704.

[19] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, *Advances in Cryptology, Proceedings of CRYPTO '86*, number 263 in Lecture Notes in Computer Science, pages 186–194. Springer-Verlag, 1987.

[20] Marc Fischlin. Communication-efficient non-interactive proofs of knowledge with online extractors. In *Crypto 2005*, volume 3621 of *LNCS*, pages 152–168. Springer, 2005.

[21] Chris Godsil. Association schemes. Lecture Notes, 2005.

[22] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(3):690–728, 1991. Online available at `http://www.wisdom.weizmann.ac.il/~oded/X/gmw1j.pdf`.

[23] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *STOC*, pages 212–219, 1996.

[24] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the Americal Statistical Association*, 58(301):13–30, 1963.

[25] Gordon James and Adalbert Kerber. *The Representation Theory of the Symmetric Group*, volume 16 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1981.

[26] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography*. Taylor & Francis, 2007.

[27] Donald E. Knuth. *Selected Papers on Discrete Mathematics*. CSLI, 2003.

[28] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 10th anniversary edition, 2010.

[29] Bruce E. Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203 of *Graduate Texts in Mathematics*. Springer, 2001.

[30] Jean-Pierre Serre. *Linear Representations of Finite Groups*, volume 42 of *Graduate Texts in Mathematics*. Springer, 1977.

[31] Peter W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *FOCS 1994*, pages 124–134. IEEE, 1994.

[32] Dominique Unruh. Quantum proofs of knowledge. In *Eurocrypt 2012*, volume 7237 of *LNCS*, pages 135–152. Springer, April 2012. Preprint on IACR ePrint 2010/212.

[33] Dominique Unruh. Non-interactive zero-knowledge proofs in the quantum random oracle model. IACR ePrint 2014/587, 2014.

[34] Dominique Unruh. Quantum position verification in the random oracle model. In *Crypto 2014*, LNCS. Springer, February 2014. To appear, preprint on IACR ePrint 2014/118.

[35] Dominique Unruh. Revocable quantum timed-release encryption. In *Eurocrypt 2014*, volume 8441 of *LNCS*, pages 129–146. Springer, 2014. Preprint on IACR ePrint 2013/606.

[36] Jeroen van de Graaf. *Towards a formal definition of security for quantum protocols.* PhD thesis, Départment d'informatique et de r.o., Université de Montréal, 1998. Online available at `http://www.cs.mcgill.ca/~crepeau/PS/these-jeroen.ps`.

[37] John Watrous. Zero-knowledge against quantum attacks. *SIAM J. Comput.*, 39(1):25–58, 2009.

[38] Eric W. Weisstein. Hypergeometric distribution. From MathWorld – A Wolfram Web Resource. `http://mathworld.wolfram.com/HypergeometricDistribution.html`. Retrieved 2014-03-19.

[39] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.

[40] Mark Zhandry. How to construct quantum random functions. In *FOCS 2013*, pages 679–687, Los Alamitos, CA, USA, 2012. IEEE Computer Society. Online version is IACR ePrint 2012/182.

[41] Mark Zhandry. Secure identity-based encryption in the quantum random oracle model. In *Crypto 2012*, volume 7417 of *LNCS*, pages 758–775. Springer, 2012. Long version on IACR ePrint 2012/076.

Symbol index

Keyword index

A   Auxiliary lemmas

Lemma 30 $\sqrt{2\big(1 - (\cos\frac{\pi}{2n})^n\big)} \in \frac{\pi}{2\sqrt{n}} + o(\frac{1}{\sqrt{n}})$.

Proof. By Taylor's theorem, for $x \to 0$,

$$\cos x \in 1 - \tfrac{x^2}{2} + O(x^4), \tag{3}$$

$$\ln(1 - x) \in -x + O(x^2), \tag{4}$$

$$e^x \in 1 + x + O(x^2). \tag{5}$$

Hence for $n \to \infty$,

$$\ln\cos\tfrac{\pi}{2n} \overset{(3)}{\in} \ln\big(1 - \tfrac{\pi^2}{8n^2} + O(n^{-4})\big) \overset{(4)}{\subseteq} -\tfrac{\pi^2}{8n^2} + O(n^{-4}).$$

Hence

$$2n\Big(1 - \big(\cos\tfrac{\pi}{2n}\big)^n\Big) \in 2n\Big(1 - e^{n\left(-\frac{\pi^2}{8n^2} + O(n^{-4})\right)}\Big) \overset{(5)}{\subseteq} 2n\Big(\tfrac{\pi^2}{8n} + O(n^{-2})\Big) \subseteq \tfrac{\pi^2}{4} + o(1).$$

Thus

$$\sqrt{n} \cdot \sqrt{2\big(1 - (\cos\tfrac{\pi}{2n})^n\big)} \in \tfrac{\pi}{2} + o(1)$$

and

$$\sqrt{2\big(1 - (\cos\tfrac{\pi}{2n})^n\big)} \in \tfrac{\pi}{2\sqrt{n}} + o(\tfrac{1}{\sqrt{n}}). \qquad \square$$

Lemma 31 *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*

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

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 φN zeros. Note that $\mu := 1 - \varphi$ is the expected value of each X_i. Thus we get

$$\Pr[\delta < \delta_{\min}] \le \Pr[1 - \delta \ge 1 - \delta_{\min}] = \Pr\Big[\sum\tfrac{X_i}{k} \ge 1 - \delta_{\min}\Big]$$

$$= \Pr\Big[\sum\tfrac{X_i}{k} - \mu \ge \varphi - \delta_{\min}\Big] \overset{(*)}{\le} e^{-2k(\varphi - \delta_{\min})^2}.$$

Here $(*)$ uses Hoeffding's inequality [24] (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 [24, Section 6]. $\qquad \square$

Lemma 32 *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 $x \xleftarrow{\$} X$. Then*

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

(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 $f : X \to Y$, let $\mathcal{D}_f$ denote the empirical distribution of f, i.e., $\mathcal{D}_f(y) = \frac{1}{n}|\{x : f(x) = y\}|$. Let $j(f) := 2\,\mathrm{SD}(\mathcal{D}, \mathcal{D}_f)$. And let $J_n := j(\mathcal{O})$, i.e., J_n is a real-valued random variable. Then [6, Lemma 8] proves that $\mathrm{E}[J_n] \leq \frac{1}{\sqrt{n}} \sum_{y \in Y} \sqrt{\mathcal{D}(y)} =: \gamma$. Since $H_{\frac{1}{2}}(\mathcal{D}) = \frac{1}{1 - \frac{1}{2}} \log\big(\sum_{y \in Y} \mathcal{D}(y)^{\frac{1}{2}}\big)$ 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 $\mathrm{SD}(y_1, y_2|E)$ denote the statistical distance between y_1 and y_2 conditioned on an event E. We can finally compute:

$$\mathrm{SD}\big((\mathcal{O}, y_1); (\mathcal{O}, y_2)\big) = \sum_{f : X \to Y} \Pr[\mathcal{O} = f] \cdot \mathrm{SD}(y_1, y_2|\mathcal{O} = f)$$

$$= \sum_{f : X \to Y} \Pr[\mathcal{O} = f] \cdot \mathrm{SD}(\mathcal{D}, \mathcal{D}_f) = \sum_{f : X \to Y} \Pr[\mathcal{O} = f] \cdot \tfrac{1}{2} j(f) = \tfrac{1}{2} \mathrm{E}[J_n] \leq \tfrac{1}{2}\gamma. \quad \square$$

Lemma 33 *Let $\mathsf{bit}_p(x)$ denote the p-th bit of x. Let $X = \{0,1\}^\ell$ for some ℓ, and $k \geq 1$, $p \in \{1, \ldots, \ell\}$ be integers. Let $S \subseteq X$ be uniformly random with $|S| = k$. Let $x \xleftarrow{\$} S$. Let $b^* \xleftarrow{\$} \{0,1\}$. Then $\mathrm{SD}\big((S, \mathsf{bit}_p(x)); (S, b^*)\big) \leq 1/2\sqrt{k}$.*

Proof. Let $P := \{x \in S : \mathsf{bit}_p(x) = 1\}$. Let $\mathrm{SD}(X; Y|S)$ 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

$$\mathrm{SD}\big((S, \mathsf{bit}_p(x)); (S, b^*)\big)$$

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

$$= \sum_S \Pr[S] \cdot \Big| \Pr[x \in P : x \xleftarrow{\$} S] - \Pr[b^* = 1 : b^* \xleftarrow{\$} \{0, 1\}] \Big|$$

$$= \sum_S \Pr[S] \cdot \Big| \tfrac{|P|}{|S|} - \tfrac{1}{2} \Big| \overset{(*)}{\leq} \sqrt{\sum_S \Pr[S]\big(\tfrac{|P|}{|S|} - \tfrac{1}{2}\big)^2} = \sqrt{\mathrm{E}\Big[\big(\tfrac{|P|}{|S|} - \tfrac{1}{2}\big)^2\Big]}$$

$$= \sqrt{\mathrm{E}\Big[\big(\tfrac{|P|}{|S|} - E\big[\tfrac{|P|}{|S|}\big]\big)^2\Big]} = \sqrt{\mathrm{Var}\big[|P|/|S|\big]} \overset{(**)}{=} \tfrac{1}{k}\sqrt{\mathrm{Var}\big[|P|\big]}.$$

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^ℓ containing $2^{\ell-1}$ successes (the elements $x \in \{0,1\}^\ell$ with $\mathsf{bit}_p(x) = 1$). That is, $|P|$ has hypergeometric distribution with parameters $m = n = 2^{\ell-1}$ and $N := k$ (in the notation of [38]). Thus (see [38]):

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

Summarizing,

$$\mathrm{SD}\big((S, \mathsf{bit}_p(x)); (S, b^*)\big) \leq \tfrac{1}{k}\sqrt{\mathrm{Var}\big[|P|\big]} \leq \tfrac{1}{k}\sqrt{k/4} = \frac{1}{2\sqrt{k}}. \qquad \square$$

Lemma 34 *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 $c' \xleftarrow{\$} C$, and $r \xleftarrow{\$} S_{|c'} := \{r : (c', r) \in S\}$ (with $r := \bot \notin R$ iff $S_{|c'} = \varnothing$). Let $(c'', r'') \xleftarrow{\$} S$.*

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

Proof. In the following calculation, $G \overset{\varepsilon}{\approx} H$ 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 $x_1, \ldots, x_k \xleftarrow{\neq} M$ means that

the x_i are chosen uniformly but distinctly from M (drawn without replacing).

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

$$\equiv F(1),\ldots,F(k) \stackrel{\neq}{\leftarrow} C \times R, \ S := \operatorname{im} F, \ j \stackrel{\$}{\leftarrow} \{1,\ldots,k\}, \ (c,r) := F(j)$$

$$\stackrel{\varepsilon_1}{\approx} F(1),\ldots,F(k) \stackrel{\$}{\leftarrow} C \times R, \ S := \operatorname{im} F, \ j \stackrel{\$}{\leftarrow} \{1,\ldots,k\}, \ (c,r) := F(j)$$

$$\equiv F_1(1),\ldots,F_1(k) \stackrel{\$}{\leftarrow} C, \ F_2(1),\ldots,F_2(k) \stackrel{\$}{\leftarrow} R, \ S := \operatorname{im}((F_1,F_2)),$$
$$j \stackrel{\$}{\leftarrow} \{1,\ldots,k\}, \ c := F_1(j), \ r := F_2(j)$$

$$\equiv F_1(1),\ldots,F_1(k) \stackrel{\$}{\leftarrow} C, \ j \stackrel{\$}{\leftarrow} \{1,\ldots,k\}, \ c := F_1(j),$$
$$F_2(1),\ldots,F_2(k) \stackrel{\$}{\leftarrow} R, \ S := \operatorname{im}((F_1,F_2))$$

$$\stackrel{\varepsilon_2}{\approx} F_1(1),\ldots,F_1(k) \stackrel{\$}{\leftarrow} C, \ c \stackrel{\$}{\leftarrow} C, \ F_2(1),\ldots,F_2(k) \stackrel{\$}{\leftarrow} R, \ S := \operatorname{im}(F_1,F_2)$$

$$\equiv F(1),\ldots,F(k) \stackrel{\$}{\leftarrow} C \times R, \ S := \operatorname{im} F, \ c \stackrel{\$}{\leftarrow} C$$

$$\stackrel{\varepsilon_1}{\approx} F(1),\ldots,F(k) \stackrel{\neq}{\leftarrow} C \times R, \ S := \operatorname{im} F, \ c \stackrel{\$}{\leftarrow} C$$

$$\equiv S \stackrel{\$}{\leftarrow} [C \times R]_k, \ c \stackrel{\$}{\leftarrow} C$$

Here ε_1 is the probability that at least two independently chosen $F(i) \stackrel{\$}{\leftarrow} C \times R$ are equal, and $\varepsilon_2 = \operatorname{SD}\big((F_1,c);(F_1,u)\big)$ for $u \stackrel{\$}{\leftarrow} C$.

Thus $\operatorname{SD}\big((S,c');(S,c'')\big) \le 2\varepsilon_1 + \varepsilon_2$. Since r' given S, c' has the same distribution as r'' given S, c'', it follows

$$\operatorname{SD}\big((S,c',r');(S,c'',r'')\big) \le 2\varepsilon_1 + \varepsilon_2. \tag{6}$$

We have $\varepsilon_1 \le \sum_{i \ne j} \Pr[F(i) = F(j)] = \sum_{i \ne j} 1/|C \times R| \le k^2/|C \times R|$.

For a function $f : \{1,\ldots,k\} \to C$, let $\mathcal{D}_f$ denote the empirical distribution of f, i.e., $\mathcal{D}_f(c) = \frac{1}{k}\big|\{i : f(i) = c\}\big|$. Let $\mathcal{U}$ denote the uniform distribution on C. Let $j(f) := 2\operatorname{SD}(\mathcal{U}, \mathcal{D}_f)$. And let $J_k := j(F_1)$ for $F_1(1),\ldots,F_1(k) \stackrel{\$}{\leftarrow} C$, i.e., J_k is a real-valued random variable. Then [6, Lemma 8] proves that $\operatorname{E}[J_k] \le \frac{1}{\sqrt{k}} \sum_{c \in C} \sqrt{\mathcal{U}(c)} = \sqrt{|C|/k}$. Then

$$\varepsilon_2 = \operatorname{SD}\big((F_1,c);(F_1,u)\big) = \sum_f \Pr[F_1 = f] \cdot \operatorname{SD}(\mathcal{D}_f, \mathcal{U})$$

$$= \sum_f \Pr[F_1 = f] \cdot \tfrac{1}{2} j(f) = \tfrac{1}{2}\operatorname{E}[J_k] \le \tfrac{1}{2}\sqrt{|C|/k}.$$

With (6), the lemma follows. $\qquad\square$

We restate an auxiliary lemma from [35, full version, Lemma 7]:

Lemma 35 *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 $\operatorname{TD}(|\Psi_1\rangle, |\Psi_2\rangle) \le 2\||\Psi_2^*\rangle\|$.*

Lemma 36 *Let $|\Psi_1\rangle, |\Psi_2\rangle$ be quantum states. Then $\operatorname{TD}(|\Psi_1\rangle, |\Psi_2\rangle) \le \big\||\Psi_1\rangle - |\Psi_2\rangle\big\|$.*

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

$$\mathrm{TD}(|\Psi_1\rangle, |\Psi_2\rangle)^2 \overset{(*)}{\leq} 1 - |\langle\Psi_1|\Psi_2\rangle|^2 = 1 - |\alpha|^2 = |\beta|^2 \leq |1-\alpha|^2 + |\beta^2| = \big\||\Psi_1\rangle - |\Psi_2\rangle\big\|^2.$$

Here $(*)$ uses that the trace distance is bounded in terms of the fidelity (e.g., [28, (9.101)]). Thus $\mathrm{TD}(|\Psi_1\rangle, |\Psi_2\rangle) \leq \big\||\Psi_1\rangle - |\Psi_2\rangle\big\|$. $\qquad\square$

Lemma 37 (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 $F : Z \to \{0,1\}$ be the following function: For each z, $F(z) := 1$ with probability γ, 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

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

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 $O_f|a,k,v\rangle := |a,k,v \oplus f(k)\rangle$. 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 obtained by performing a projective measurement P_{final} on that final state.

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

We compute:

$$\begin{aligned}
D_i^f &:= \mathrm{TD}(|\Psi_f^i\rangle, |\Psi^i\rangle) = \mathrm{TD}(O_f|\Psi_f^{i-1}\rangle, |\Psi^{i-1}\rangle) \\
&\leq \mathrm{TD}(O_f|\Psi_f^{i-1}\rangle, O_f|\Psi^{i-1}\rangle) + \mathrm{TD}(O_f|\Psi^{i-1}\rangle, |\Psi^{i-1}\rangle) \\
&= D_{i-1}^f + \mathrm{TD}(O_f|\Psi^{i-1}\rangle, |\Psi^{i-1}\rangle).
\end{aligned}$$

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

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

$$\sum_f \alpha_f \big\|Q_f|\Psi_i\rangle\big\|^2 \overset{(*)}{=} \sum_f \alpha_f \sum_{z:f(z)=1} \big\|Q_z|\Psi_i\rangle\big\|^2 = \sum_{z\in Z} \sum_{f:f(x)=1} \alpha_f \big\|Q_z|\Psi_i\rangle\big\|^2$$

$$\overset{(**)}{=} \lambda \sum_z \big\|Q_z|\Psi_i\rangle\big\|^2 = \lambda \big\||\Psi_i\rangle\big\|^2 = \lambda. \tag{7}$$

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

Then

$$\sum_f \alpha_f \operatorname{TD}(|\Psi_f^q\rangle, |\Psi^q\rangle) = \sum_f \alpha_f D_q^f \leq \sum_{f,i} \alpha_f \operatorname{TD}(O_f|\Psi^i\rangle, |\Psi^i\rangle)$$

$$= \sum_{f,i} \alpha_f \operatorname{TD}\big(O_f Q_f|\Psi^i\rangle + (1 - Q_f)|\Psi^i\rangle, \ Q_f|\Psi^i\rangle + (1 - Q_f)|\Psi^i\rangle\big)$$

$$\overset{(*)}{\leq} \sum_{f,i} \alpha_f 2\|Q_f|\Psi^i\rangle\| \overset{(**)}{\leq} 2 \sum_i \sqrt{\sum_f \alpha_f \|Q_f|\Psi^i\rangle\|^2}$$

$$\overset{(7)}{=} 2 \sum_i \sqrt{\lambda} = 2q\sqrt{\lambda}. \tag{8}$$

Here $(*)$ uses Lemma 35. And $(**)$ uses Jensen's inequality. Finally,

$$\Big|\Pr[b = 1 : b \leftarrow A^F] - \Pr[b = 1 : b \leftarrow A^N]\Big|$$

$$\leq \sum_f \alpha_f \Big|\Pr[b = 1 : b \leftarrow A^f] - \Pr[b = 1 : b \leftarrow A^N]\Big|$$

$$\leq \sum_f \alpha_f \operatorname{TD}(|\Psi_f^q\rangle, |\Psi^q\rangle) \overset{(8)}{\leq} 2q\sqrt{\lambda}. \qquad \square$$

The following lemma formalizes that an oracle $\mathcal{O}_1$ does not help (much) in finding a value w if $\mathcal{O}_1$ only gives answers when w is already contained in its input.

Lemma 38 (Removing redundant oracles 1) *Let $w, \mathcal{O}_1, \mathcal{O}_2$ be chosen according to some joint distribution. Here w is a bitstring, and $\mathcal{O}_1, \mathcal{O}_2$ are oracles, and $\mathcal{O}_1$ is classical (i.e., $\forall x, y. \exists y'. \mathcal{O}_1|x\rangle|y\rangle = |x\rangle|y'\rangle$). Fix a function f. Assume that for all x with $f(x) \neq w$, $\mathcal{O}_1(x) = 0$. (In other words, $\mathcal{O}_1|x\rangle|y\rangle = |x\rangle|y\rangle$ for $f(x) \neq w$.)*

Let A be an oracle machine that makes at most q queries to $\mathcal{O}_1$ and q' queries to $\mathcal{O}_2$. Then there is another oracle machine $\hat{A}$ that makes at most q' queries to $\mathcal{O}_2$ such that:

$$\Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] \leq 2(q + 1)\sqrt{\Pr[w' = w : w' \leftarrow \hat{A}^{\mathcal{O}_2}]}$$

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 $|\Psi^*\rangle := (U_2\mathcal{O}_1)^q U_2|\Psi\rangle$ for some unitary U_2 depending only on $\mathcal{O}_2$, and $\mathcal{O}_1$ operating on quantum registers K, V for oracle input and output, and $|\Psi\rangle$ being some initial state independent of $\mathcal{O}_1, \mathcal{O}_2, w$. Let $|\Psi_i\rangle := (U_2\mathcal{O}_1)^{q-i} U_2^{i+1}|\Psi\rangle$. Note that $|\Psi_0\rangle = |\Psi^*\rangle$. Let $P_X := \sum_{x:f(x)=w} |x\rangle\langle x| \otimes I$ and $\bar{P}_X := 1 - P_X$. Note that since $\mathcal{O}_1|x\rangle|y\rangle = |x\rangle|y\rangle$ for $f(x) \neq w$, we have $\mathcal{O}_1 = \mathcal{O}_1 P_X + \bar{P}_X$. We have for $i = 1, \ldots, q$:

$$\operatorname{TD}(|\Psi_{i-1}\rangle, |\Psi_i\rangle) = \operatorname{TD}\big((U_2\mathcal{O}_1)^{q-i}(U_2\mathcal{O}_1)U_2^i|\Psi\rangle, (U_2\mathcal{O}_1)^{q-i}U_2 U_2^i|\Psi\rangle\big)$$

$$= \operatorname{TD}(\mathcal{O}_1 U_2^i|\Psi\rangle, U_2^i|\Psi\rangle)$$

$$= \operatorname{TD}\big(\mathcal{O}_1 P_X U_2^i|\Psi\rangle + \bar{P}_X U_2^i|\Psi\rangle, P_X U_2^i|\Psi\rangle + \bar{P}_X U_2^i|\Psi\rangle\big)$$

$$\overset{(*)}{\leq} 2\|P_X U_2^i|\Psi\rangle\|.$$

36

Here $(*)$ uses Lemma 35 (using that $|\Psi_1^*\rangle := \mathcal{O}_1 P_X U_2^i |\Psi\rangle$ and $|\Psi_2^*\rangle := P_X U_2^i |\Psi\rangle$ are both orthogonal to $|\Phi^*\rangle := \bar{P}_X U_2^i |\Psi\rangle$ because $\mathcal{O}_1$ is classical and therefore does not leave the image of P_X).

Thus $\mathrm{TD}(|\Psi^*\rangle, |\Psi_q\rangle) \leq \sum_{i=1}^q 2\|P_X U_2^i |\Psi\rangle\|$. For $i = 1, \ldots, q$, let $A_i^{\mathcal{O}_2}$ be the oracle algorithm that computes $U_2^i |\Psi\rangle$ 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 $\mathcal{O}_1$ because U_2 does not depend on $\mathcal{O}_1$.) Then $\Pr[w = w' : w' \leftarrow A_i] = \|P_X U_2^i |\Psi\rangle\|^2$. 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 $|\Psi_q\rangle$. Thus

$$\Big|\Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] - \Pr[w = w' : w' \leftarrow A_0^{\mathcal{O}_2}]\Big|$$
$$\leq \mathrm{TD}(|\Psi^*\rangle, |\Psi_q\rangle) \leq \sum_{i=1}^q 2\sqrt{\Pr[w = w' : w' \overset{\$}{\leftarrow} A_i^{\mathcal{O}_2}]}$$

Let $\hat{A}^{\mathcal{O}_2}$ be the algorithm that picks $i \overset{\$}{\leftarrow} \{0, \ldots, q\}$ and runs A_i. Then

$$\Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] \leq \sum_{i=1}^q 2\sqrt{\Pr[w = w' : w' \leftarrow A_i^{\mathcal{O}_2}]} + \Pr[w = w' : w' \leftarrow A_0^{\mathcal{O}_2}]$$
$$\leq 2(q+1) \sum_{i=0}^q \tfrac{1}{q+1} \sqrt{\Pr[w = w' : w' \leftarrow A_i^{\mathcal{O}_2}]}$$
$$\overset{(*)}{\leq} 2(q+1) \sqrt{\sum_{i=0}^q \tfrac{1}{q+1} \Pr[w = w' : w' \leftarrow A_i^{\mathcal{O}_2}]}$$
$$= 2(q+1) \sqrt{\Pr[w = w' : w' \leftarrow \hat{A}^{\mathcal{O}_2}]}.$$

Here $(*)$ uses Jensen's inequality. $\qquad\square$

The following lemma formalizes that if w is a random bitstring that can be accessed only by querying an oracle $\mathcal{O}_1$ on some input $x \in X$, then the probability of finding w using $\mathcal{O}_1$ is bounded in terms of the probability of finding some $x \in X$ without using $\mathcal{O}_1$.

Lemma 39 (Removing redundant oracles 2) *Let $w, X, \mathcal{O}_1, \mathcal{O}_2$ be chosen according to some joint distribution such that w and $\mathcal{O}_2$ are stochastically independent. Here X is a set of bitstrings, and $\mathcal{O}_1, \mathcal{O}_2$ are oracles, and $\mathcal{O}_1$ is classical (i.e., $\forall x, y. \exists y'. \mathcal{O}_1 |x\rangle |y\rangle = |x\rangle |y'\rangle$). And w is uniformly distributed on $\{0, 1\}^\ell$. Assume that for all $x \notin X$, $\mathcal{O}_1(x) = 0$. (In other words, $\mathcal{O}_1 |x\rangle |y\rangle = |x\rangle |y\rangle$ for $x \notin X$.)*

Let A be an oracle machine that makes at most q queries to $\mathcal{O}_1$ and q' queries to $\mathcal{O}_2$. Then there is another oracle machine $\hat{A}$ that makes at most q queries to $\mathcal{O}_2$ such that:

$$\Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] \leq 2q \sqrt{\Pr[x \in X : x \leftarrow \hat{A}^{\mathcal{O}_2}]} + 2^{-\ell}$$

Proof. Let $P_X := \sum_{x \in X} |x\rangle\langle x| \otimes I$ and $\bar{P}_X := 1 - P_X$. Note that since $\mathcal{O}_1|x\rangle|y\rangle = |x\rangle|y\rangle$ 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 38. (Remember that all of these only depend on $\mathcal{O}_2$, not $\mathcal{O}_1$.) Exactly as in Lemma 38, we get $\mathrm{TD}(|\Psi^*\rangle, |\Psi_q\rangle) \leq \sum_{i=1}^q 2\|P_X U_2^i|\Psi\rangle\|$. For $i = 1, \ldots, q$, let $A_i^{\mathcal{O}_2}$ be the oracle algorithm that computes $U_2^i|\Psi\rangle$ and measures register K in the computational basis and outputs the outcome. Then $\Pr[x \in X : x \leftarrow A_i^{\mathcal{O}_2}] = \|P_X U_2^i|\Psi\rangle\|^2$.

Like in the proof of Lemma 38, 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 $\hat{A}^{\mathcal{O}_2}$ pick a random $i \overset{\$}{\leftarrow} \{1, \ldots, q\}$ (not $i \overset{\$}{\leftarrow} \{0, \ldots, q\}$ as in Lemma 38!) and run $A_i^{\mathcal{O}_2}$. Then

$$\Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] - \Pr[w = w' : w' \leftarrow A_0^{\mathcal{O}_2}] \leq \mathrm{TD}(|\Psi^*\rangle, |\Psi_q\rangle)$$

$$\leq 2q \sum_{i=1}^q \tfrac{1}{q}\sqrt{\Pr[x \in X : x \overset{\$}{\leftarrow} A_i^{\mathcal{O}_2}]} \overset{(*)}{\leq} 2q\sqrt{\sum_{i=1}^q \tfrac{1}{q}\Pr[x \in X : x \overset{\$}{\leftarrow} A_i^{\mathcal{O}_2}]}$$

$$= 2q\sqrt{\Pr[x \in X : x \leftarrow \hat{A}^{\mathcal{O}_2}]}. \tag{9}$$

Here $(*)$ uses Jensen's inequality.

Since w and $\mathcal{O}_2$ are independent and w is uniform on $\{0,1\}^\ell$, $\Pr[w = w' : w' \leftarrow A_0^{\mathcal{O}_2}] \leq 2^{-\ell}$. With (9), we get $2q\sqrt{\Pr[x \in X : x \leftarrow \hat{A}^{\mathcal{O}_2}]} \geq \Pr[w = w' : w' \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}] - 2^{-\ell}$. $\qquad\square$

Theorem 40 (Small range distributions [40]) *Fix sets Z, Y and a distribution $\mathcal{D}_Y$ on Y, and integers s, q.*

Let $H : Z \to Y$ be chosen as: for each $z \in Z$, $H(z) \leftarrow \mathcal{D}_Y$.

Let $G : Z \to Y$ be chosen as: Pick $y_1, \ldots, y_s \leftarrow \mathcal{D}_Y$, then for each $z \in Z$, pick $i_z \overset{\$}{\leftarrow} \{1, \ldots, s\}$, and set $G(z) := y_{i_z}$.

Let A be an oracle algorithm making at most q queries. Then

$$\left|\Pr[b = 1 : b \leftarrow A^H] - \Pr[b = 1 : b \leftarrow A^G]\right| \leq \pi^2 (2q)^3 / 6s < 14q^3/s.$$

Proof. This is merely a reformulation of [40, Corollary VII.5]. (Note that the distance between distributions in [40] is defined to be twice the statistical, this is why in our formulation of the theorem the bound is only half as large.) $\qquad\square$

B   Proofs for Section 3

Lemma 41 *Let $|\Psi\rangle$ be a state, chosen according to some distribution. Let $|\bot\rangle$ 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 $|\bot\rangle$.)*

Let $\mathcal{O}_\Psi$ be an oracle with $\mathcal{O}_\Psi|\Psi\rangle = |\bot\rangle$, $\mathcal{O}_\Psi|\bot\rangle = |\Psi\rangle$, and $\mathcal{O}_\Psi|\Psi^\bot\rangle = |\Psi^\bot\rangle$ for any $|\Psi^\bot\rangle$ orthogonal to both $|\Psi\rangle$ and $|\bot\rangle$. Let $\mathcal{O}_{\mathrm{Ref}} := I - 2|\Psi\rangle\langle\Psi|$.

Let $\mathcal{O}$ 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 $|\alpha_j\rangle := (\cos\frac{j\pi}{2n})|\Psi\rangle + (\sin\frac{j\pi}{2n})|\perp\rangle$.

Then there is an oracle algorithm B that makes q_Ψ queries to $\mathcal{O}_{\mathrm{Ref}}$ and makes the same number of queries to $\mathcal{O}$ as A such that:

$$\mathrm{TD}\big(B^{\mathcal{O}_{\mathrm{Ref}},\mathcal{O}}(|R\rangle, |\Phi\rangle), A^{\mathcal{O}_\Psi,\mathcal{O}}(|\Phi\rangle)\big) \leq \tfrac{\pi q_\Psi}{2\sqrt{n}} + q_\Psi\, o(\tfrac{1}{\sqrt{n}}).$$

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

We first show that

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

where $S|\Phi_0\rangle|\Phi_1\rangle\ldots|\Phi_n\rangle := |\Phi_1\rangle\ldots|\Phi_n\rangle|\Phi_0\rangle$ (cyclic shift) and $|R\rangle$ is as in the statement of the lemma (the reservoir state).

We have

$$(S|\perp\rangle|R\rangle)^\dagger(|\Psi\rangle|R\rangle) = \Big(|\alpha_1\rangle|\alpha_2\rangle\ldots|\alpha_n\rangle|\perp\rangle\Big)^\dagger\Big(|\Psi\rangle|\alpha_1\rangle\ldots|\alpha_{n-1}\rangle|\alpha_n\rangle\Big)$$

$$= \langle\alpha_1|\Psi\rangle \cdot \prod_{j=1}^{n-1}\langle\alpha_{j+1}|\alpha_j\rangle \cdot \langle\perp|\alpha_n\rangle$$

$$\stackrel{(*)}{=} \cos\tfrac{\pi}{2n} \cdot \prod_{j=1}^{n-1}\cos\big(\tfrac{(j+1)\pi}{2n} - \tfrac{j\pi}{2n}\big) \cdot \sin\tfrac{n\pi}{2n} = (\cos\tfrac{\pi}{2n})^n.$$

Here $(*)$ uses that $|\Psi\rangle$ and $|\perp\rangle$ 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 $\big\||\Phi\rangle - |\Phi'\rangle\big\|^2 = (|\Phi\rangle - |\Phi'\rangle)^\dagger(|\Phi\rangle - |\Phi'\rangle) = 1 - \langle\Phi|\Phi'\rangle - \langle\Phi'|\Phi\rangle + 1 = 2(1 - \Re\langle\Phi|\Phi'\rangle)$ where $\Re$ denote the real part. Thus $\big\|S|\perp\rangle|R\rangle - |\Psi\rangle|R\rangle\big\| \leq \sqrt{2(1 - (\cos\frac{\pi}{2n})^n)} \in \tfrac{\pi}{2\sqrt{n}} + o(\tfrac{1}{\sqrt{n}}) = \varepsilon_n$. (The asymptotic bound uses Lemma 30.) This shows the lhs of (10). The rhs follows from the rhs by applying the unitary $S^\dagger$ on both sides.

Let U_Ψ denote the unitary computed by circuit (1) on page 13. We will show that for any $|\Phi\rangle$,

$$U_\Psi|\Phi\rangle|R\rangle|0\rangle = (\mathcal{O}_\Psi|\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n. \tag{11}$$

By linearity of $U_\Psi, \mathcal{O}_\Psi$ and the triangle inequality, it is sufficient to verify this for $|\Phi\rangle = |\Psi\rangle$, $|\Phi\rangle = |\perp\rangle$, and $|\Phi\rangle$ orthogonal to both $|\Psi\rangle, |\perp\rangle$. In an execution of circuit (1) 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

$$|\Phi_1\rangle = |\Psi\rangle|R\rangle|0\rangle, \qquad |\Phi_2\rangle = |\Psi\rangle|R\rangle|0\rangle,$$
$$|\Phi_3\rangle = |\Psi\rangle|R\rangle|1\rangle, \qquad |\Phi_4\rangle \overset{(10)}{=} |\bot\rangle|R\rangle|1\rangle \pm \varepsilon_n,$$
$$|\Phi'\rangle = |\bot\rangle|R\rangle|0\rangle \pm \varepsilon_n = (\mathcal{O}_\Psi|\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n.$$

For $|\Phi\rangle = |\bot\rangle$, we have

$$|\Phi_1\rangle = |\bot\rangle|R\rangle|1\rangle, \qquad\qquad |\Phi_2\rangle \overset{(10)}{=} |\Psi\rangle|R\rangle|1\rangle \pm \varepsilon_n,$$
$$|\Phi_3\rangle = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n, \qquad |\Phi_4\rangle = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n,$$
$$|\Phi'\rangle = |\Psi\rangle|R\rangle|0\rangle \pm \varepsilon_n = (\mathcal{O}_\Psi|\Phi\rangle)|R\rangle|0\rangle \pm \varepsilon_n.$$

And for $|\Phi\rangle$ orthogonal to $|\Psi\rangle$ and $|\bot\rangle$, we have

$$|\Phi_1\rangle = |\Phi\rangle|R\rangle|0\rangle, \qquad |\Phi_2\rangle = |\Phi\rangle|R\rangle|0\rangle,$$
$$|\Phi_3\rangle = |\Phi\rangle|R\rangle|0\rangle, \qquad |\Phi_4\rangle = |\Phi\rangle|R\rangle|0\rangle,$$
$$|\Phi'\rangle = |\Phi\rangle|R\rangle|0\rangle = (\mathcal{O}_\Psi|\Phi\rangle)|R\rangle|0\rangle.$$

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 (1) on X, R, Z whenever A invokes $\mathcal{O}_\Psi$ on X. (And when A performs a controlled invocation of $\mathcal{O}_\Psi$, 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 $\big\| |\Phi_A\rangle - |\Phi_B\rangle \big\| \leq q_\Psi \varepsilon_n$. By Lemma 36, $\mathrm{TD}(|\Phi_A\rangle - |\Phi_B\rangle) \leq_\Psi \varepsilon_n$. Thus

$$\mathrm{TD}\big(B^{\mathcal{O}_{\mathrm{Ref}}, \mathcal{O}}(|R\rangle, |\Phi\rangle), A^{\mathcal{O}_\Psi, \mathcal{O}}(|\Phi\rangle)\big) \quad \leq \quad q_\Psi \varepsilon_n \quad \leq \quad \frac{\pi q_\Psi}{2\sqrt{n}} + q_\Psi \, o(\tfrac{1}{\sqrt{n}}). \quad \square$$

Lemma 42 *Let $|\Psi\rangle$ be a state, chosen according to some distribution. Let $\mathcal{O}_{\mathrm{Ref}} := I - 2|\Psi\rangle\langle\Psi|$. Let $\mathcal{O}$ 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_{Ref} queries to $\mathcal{O}_{\mathrm{Ref}}$. Let $m \geq 0$ be an integer. Then there is an oracle algorithm B that makes the same number of queries to $\mathcal{O}$ as A such that:*

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

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

Let $V \subseteq \mathcal{H}^{\otimes m+1}$ 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 $\big\| |\Phi\rangle - |\Phi'\rangle \big\| \leq \varepsilon$.

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

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

$$U_V|\Phi\rangle|T\rangle = (\mathcal{O}_{\mathrm{Ref}}|\Phi\rangle)|T\rangle \pm \tfrac{2}{\sqrt{m+1}}. \tag{12}$$

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

$$
\begin{aligned}
|\alpha| = |\langle\chi|(|\Phi\rangle|T\rangle)| &= \Big| \tfrac{1}{m+1} \sum_{j=0}^{m} \langle\chi|S^j(|\Phi\rangle|T\rangle) \Big| \\
&= \tfrac{1}{m+1} \Big| \langle\chi|\Big(\sum_{j=0}^{m} S^j|\Phi\rangle|T\rangle\Big) \Big| \overset{(*)}{\leq} \tfrac{1}{m+1}\big|\sqrt{m+1}\big| = \tfrac{1}{\sqrt{m+1}}.
\end{aligned}
$$

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 $\big\|\sum_j S^j|\Phi\rangle|T\rangle\big\| = \sqrt{m+1}$. Thus

$$\big\|U_V|\Phi\rangle|T\rangle - (\mathcal{O}_{\mathrm{Ref}}|\Phi\rangle)|T\rangle\big\| = \big\| |\Phi\rangle|T\rangle - 2\alpha|\chi\rangle \ - \ |\Phi\rangle|T\rangle \big\| = |2\alpha| \leq \tfrac{2}{\sqrt{m+1}}.$$

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 $U_V|\Phi\rangle|T\rangle = -|\Phi\rangle|T\rangle = \mathcal{O}_{\mathrm{Ref}}|\Phi\rangle|T\rangle$. By linearity and the triangle inequality, (12) then holds for all $|\Phi\rangle \in \mathcal{H}$.

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 $\mathcal{O}_{\mathrm{Ref}}$ on X. (And when A performs a controlled invocation of $\mathcal{O}_{\mathrm{Ref}}$, 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 $\big\| |\Phi_A\rangle - |\Phi_B\rangle \big\| \leq \tfrac{2q_{\mathrm{Ref}}}{\sqrt{m+1}}$. By Lemma 36, $\mathrm{TD}(|\Phi_A\rangle - |\Phi_B\rangle) \leq \tfrac{2q_{\mathrm{Ref}}}{\sqrt{m+1}}$. Thus

$$\mathrm{TD}\big(B^{\mathcal{O}}(|T\rangle, |\Phi\rangle), A^{\mathcal{O}_{\mathrm{Ref}}, \mathcal{O}}(|\Phi\rangle)\big) \leq \frac{2q_{\mathrm{Ref}}}{\sqrt{m+1}}. \qquad \square$$

C   Proof of Theorem 5

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 $\mathbb{S}_L$ 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 $\mathbb{S}_n$ denote the isomorphism class of the symmetric groups $\mathbb{S}_L$ with $|L| = n$. A permutation $\sigma \in \mathbb{S}_X$ acts on $z_y \in D$ in a natural way: we define

$$\sigma(z_y) := (z_{y,\sigma^{-1}(1)}, \ldots, z_{y,\sigma^{-1}(N)}), \tag{13}$$

so that $(\sigma(z_y))_{\sigma(x)} = z_{y,x}$ holds. A permutation $\pi \in \mathbb{S}_Y$ 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 $(\boldsymbol{\sigma}, \pi)$, where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_M) \in \mathbb{S}_X^M$ and $\pi \in \mathbb{S}_Y$. Let this pair act on $z \in D^M$ by first permuting the entries of z with respect to π and then permuting entries within each $(\pi(z))_y$ with respect to σ_y. Namely, let

$$(\boldsymbol{\sigma}, \pi) : (z_1, \ldots, z_M) \mapsto (\sigma_1(z_{\pi^{-1}(1)}), \ldots, \sigma_M(z_{\pi^{-1}(M)})). \tag{14}$$

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

Definition 43 ([25, Chapter 4]) *The* wreath product $G \wr \mathbb{S}_M$ *of groups G and $\mathbb{S}_M$ is the group whose elements are $(\boldsymbol{\sigma}, \pi) \in G^M \times \mathbb{S}_M$ and whose group operation is*

$$\big((\sigma_1', \ldots, \sigma_M'), \pi'\big)\big((\sigma_1, \ldots, \sigma_M), \pi\big) := \big((\sigma_1'\sigma_{(\pi')^{-1}(1)}, \ldots, \sigma_M'\sigma_{(\pi')^{-1}(M)}), \pi'\pi\big).$$

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 $\mathbb{W}$ defined by its action on the sets $Y \times X$ and $Y \times X_2$, respectively:

$$(\boldsymbol{\sigma}, \pi) : (y, x) \qquad\qquad \mapsto (\pi(y), \sigma_{\pi(y)}(x)), \tag{15}$$
$$(\boldsymbol{\sigma}, \pi) : (y, \{x_1, x_2\}) \mapsto (\pi(y), \{\sigma_{\pi(y)}(x_1), \sigma_{\pi(y)}(x_2)\}). \tag{16}$$

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 $\tau \in \mathbb{W}$, 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

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

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

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

  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.

- $\mathcal{H}_O := \mathcal{H}_{O_Y} \otimes \mathcal{H}_{O_{X_2}}$ is the "output" register, where $\mathcal{H}_{O_Y}$ and $\mathcal{H}_{O_{X_2}}$ correspond to the sets Y and X_2, 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

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

  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 $\mathbb{I}$ 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 $\mathbb{I}_{QB}$ instead of $\mathbb{I}_Q \otimes \mathbb{I}_B$.

Let $|\xi_\emptyset(z)\rangle_A := |\xi'(z)\rangle_R \otimes |\xi''\rangle_{A-R}$ 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_Q, and U_O be unitary transformations corresponding to representations (14), (15), and (16) of $\mathbb{W}$, respectively, where the register $\mathcal{H}_{\mathcal{I}}$ is yet to be defined. (That is, $U_{\mathcal{I}}, U_Q, U_O$ are actually families of unitaries, indexed by elements $\tau \in \mathbb{W}$.) We add a subscript $\tau \in \mathbb{W}$ when we want to specify that we are considering the representation of

the element τ, for example, we may write $U_{Q,\tau}$. Since $\mathcal{H}_R$ is essentially the h-th tensor power of $\mathcal{H}_Q$, we define $U_R := U_Q^{\otimes h}$. The tensor product of two (or more) representations of $\mathbb{W}$ is also a representation of $\mathbb{W}$. Let $U_{\mathcal{I}Q} := U_{\mathcal{I}} \otimes U_Q$ and $U_{\mathcal{I}O} := U_{\mathcal{I}} \otimes U_O$, an we later use analogous notation for other "concatenations".

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

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

Then, at specific points in the algorithm, we insert unitary transformations controlled by the content τ 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,

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

2. Before each oracle call, $\mathcal{O}_V$ or $\mathcal{O}_F$, we insert the controlled transformation $U_{Q,\tau}^{-1}$ 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 $U_{O,\tau}^{-1}$ 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 $\mathcal{H}_{\mathcal{I}}$ that stores $z \in D^M$. Namely, let $\mathcal{H}_{\mathcal{I}} := \bigotimes_{y=1}^{M} \mathcal{H}_{I(y)}$ be an $\binom{N}{k}^M$-dimensional Hilbert space whose basis states correspond to possible inputs z, where we define $\mathcal{H}_{I(y)}$ to be $\binom{N}{k}$-dimensional Hilbert space whose basis states correspond to $z_y \in D$. Since all the spaces $\mathcal{H}_{I(y)}$ are essentially equivalent, we write $\mathcal{H}_I$ instead of $\mathcal{H}_{I(y)}$ when we do not care which particular $y \in Y$ we are talking about, and $\mathcal{H}_{\mathcal{I}} = \mathcal{H}_I^{\otimes M}$.

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

$$\binom{N}{k}^{-M/2} \sum_{z \in D^M} |z\rangle_{\mathcal{I}} \otimes |\mathbb{W}\rangle_S \otimes |\xi_\emptyset(z)\rangle_A.$$

We transform the symmetrised version of $\mathcal{A}$ into a sequence of transformations on a Hilbert space $\mathcal{H} = \mathcal{H}_{\mathcal{I}} \otimes \mathcal{H}_S \otimes \mathcal{H}_A$. A black-box transformation $\mathcal{O}$ (where $\mathcal{O} = \mathcal{O}_V$ or

$\mathcal{O} = \mathcal{O}_F$) is replaced by a transformation $\mathcal{O}' = \sum_{z \in D^M} |z\rangle\langle z| \otimes \mathcal{O}(z)$, where $\mathcal{O}(z)$ is the transformation $\mathcal{O}$ for the case when the input is equal to z.

At the end, the algorithm measures the input register $\mathcal{H}_{\mathcal{I}}$ and the output register $\mathcal{H}_O = \mathcal{H}_{O_Y} \otimes \mathcal{H}_{O_{X_2}}$ 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 $z_{y,x_1} = z_{y,x_2} = 1$.

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

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

where $\gamma = 1/\sqrt{M!(N!\binom{N}{k})^M}$, and, after we apply $U_{Q,\tau}^{-1}$, we have

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

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

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

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

For $q \in \{0, 1, \ldots, q_T - 1\}$, let ρ_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 ρ_q be obtained from ρ_q' by tracing out the register $\mathcal{H}_Q$. Due to (21), we have

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

Similarly, for $q = q_T$, let $|\phi_{q_T}\rangle$ be the final state of the algorithm (i.e., the state after the controlled $U_{O,\tau}^{-1}$), and it satisfies an analogous symmetry to (21): for all $\kappa \in \mathbb{W}$, we have $(U_{\mathcal{I}SO,\kappa} \otimes \mathbb{I}_{A-O})|\phi_{q_T}\rangle = |\phi_{q_T}\rangle$. Let ρ_{q_T}'' be the density matrix obtained from $|\phi_{q_T}\rangle\langle\phi_{q_T}|$ by tracing out all the registers but $\mathcal{H}_{\mathcal{I}}$ and $\mathcal{H}_O$, let ρ_{q_T} be obtained from ρ_{q_T}'' by tracing out the register $\mathcal{H}_O$. Again, we have

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

Note that, throughout the algorithm, the density matrix of the $\mathcal{H}_{\mathcal{I}}$ 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 ρ_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 $\mathbb{S}_X$

Consider a positive integer n. The representation theory of $\mathbb{S}_n$ is closely related to partitions. A partition λ 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 $\mathbb{S}_n$ 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 [30] for more background on the representation theory of finite groups and to [25, 29] for the representation theory of the symmetric group and the wreath product.)

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

$$|\mathbb{S}_N|/|\mathbb{S}_k \times \mathbb{S}_{N-k}| = |D|/|\{1^k 0^{N-k}\}|,$$

U_I is equal to the induced representation when we induce V_I from $\mathbb{S}_k \times \mathbb{S}_{N-k}$ to $\mathbb{S}_N$. For shortness, we write $U_I = V_I \uparrow \mathbb{S}_N$. The Littlewood-Richardson rule then implies

$$((k) \times (N-k)) \uparrow \mathbb{S}_N = (N) \oplus (N-1,1) \oplus (N-2,2) \oplus \ldots \oplus (N-k,k). \qquad (24)$$

Hence, we have

$$\mathcal{H}_I = \bigoplus_{i=0}^{k} \mathcal{H}_I^{(N-i,i)},$$

where U_I restricted to $\mathcal{H}_I^{(N-i,i)}$ is an irrep of $\mathbb{S}_N$ corresponding to the partition $(N-i,i)$ of N. It is also known (see [21, 27]) that $\mathcal{H}_I^{(N-i,i)} = \mathcal{T}_I^i \cap (\mathcal{T}_I^{i-1})^\perp$, where $\mathcal{T}_I^i$ is the space spanned by all $\binom{N}{i}$ states

$$|\psi_{x_1,\ldots,x_i}\rangle = \frac{1}{\sqrt{\binom{N-i}{k-i}}} \sum_{\substack{z_y \in D \\ z_{y,x_1} = \ldots = z_{y,x_i} = 1}} |z_y\rangle \qquad (25)$$

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

C.4 Framework for the proof

We use the representation-theoretic framework developed in [2] (and used in [4] and [3]). Let

$$\mathcal{H}_{I,a} := \mathcal{T}_I^1 = \mathcal{H}_I^{(N)} \oplus \mathcal{H}_I^{(N-1,1)}, \qquad \mathcal{H}_{I,b} := \mathcal{H}_I \cap (\mathcal{H}_{I,a})^\perp,$$
$$\mathcal{H}_{\mathcal{I},a} := \mathcal{H}_{I,a}^{\otimes M}, \qquad\qquad\qquad \mathcal{H}_{\mathcal{I},b} := \mathcal{H}_{\mathcal{I}} \cap (\mathcal{H}_{\mathcal{I},a})^\perp.$$

And let $\Pi_{I,a}$, $\Pi_{I,b}$, $\Pi_{\mathcal{I},a}$, and $\Pi_{\mathcal{I},b}$ denote the projections to the spaces $\mathcal{H}_{I,a}$, $\mathcal{H}_{I,b}$, $\mathcal{H}_{\mathcal{I},a}$, and $\mathcal{H}_{\mathcal{I},b}$, respectively.

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

$$p_{a,q} := \mathrm{Tr}(\rho_q \Pi_{\mathcal{I},a}) \qquad \text{and} \qquad p_{b,q} := 1 - p_{a,q} = \mathrm{Tr}(\rho_q \Pi_{\mathcal{I},b}).$$

Theorem 5 (Hardness of the two values problem) then follows from the following three lemmas.

Lemma 44 *The success probability of the algorithm is at most* $\frac{2(k-1)}{N-1} + \sqrt{2 p_{b,q_T}}$.

Lemma 45 *(At the very beginning of the algorithm) we have* $p_{b,0} < h^2/(2M)$.

Lemma 46 *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 44 and 46. The size of Y indeed does not matter for them, as in we will eventually reduce the general case for Lemmas 44 and 46 to the case when $|Y| = 1$.

C.5 Proof of Lemma 45

Let us rewrite (19) as

$$|\xi'(z)\rangle_R = \bigotimes_{\ell=1}^{h} \left(\frac{1}{\sqrt{M}} \sum_{y \in Y} |y\rangle_{R_Y(\ell)} \big(\alpha_{\ell,0} |\Psi(z_y)\rangle + \alpha_{\ell,1} |\Phi\rangle \big)_{R_X(\ell)} \right)$$

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

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

$$|\xi'(y_1, \ldots, y_h)\rangle_{R_X} := \bigotimes_{\ell=1}^{h} (\alpha_{\ell,0} |\Psi(z_{y_\ell})\rangle_{R_X(\ell)} + \alpha_{\ell,1} |\Phi\rangle_{R_X(\ell)})$$

has unit norm for $\langle \Psi(z_y) | \Phi \rangle = \langle \Sigma \Psi(z) | \Sigma \Phi \rangle = \sqrt{k/N}$. 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 $|\xi'(z)\rangle_R = |\xi'_a(z)\rangle_R + |\xi'_b(z)\rangle_R$, where the unnormalized state $|\xi'_a(z)\rangle_R$ corresponds to all $(y_1, \ldots, y_h) \in Y_h$ in the register $\mathcal{H}_{R_Y}$. Then, $\||\xi'_b(z)\rangle_{R_Y}\|^2$ 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 $h(h-1)/(2M)$ [26]. For $c \in \{a, b\}$, let

$$|\phi_c\rangle := \binom{N}{k}^{-M/2} \sum_{z \in D^m} |z\rangle_{\mathcal{I}} |\mathbb{W}\rangle_S |\xi'_c(z)\rangle_R |\xi''\rangle_{A-R},$$

47

and note that $\||\phi_c\rangle\| = \||\xi'_c(z)\rangle\|$. The initial state of the algorithm is $|\phi_a\rangle + |\phi_b\rangle$. (Note: in this proof, the subscript of ϕ does not denote the number of queries.)

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

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

$$p_{b,0} = \mathrm{Tr}(\rho_0 \Pi_{\mathcal{I},b}) = \mathrm{Tr}\Big(\mathrm{Tr}_{SA}\big((|\phi_a\rangle + |\phi_b\rangle)(\langle\phi_a| + \langle\phi_b|)\big)\Pi_{\mathcal{I},b}\Big)$$
$$= \langle\phi_b|(\Pi_{\mathcal{I},b} \otimes \mathbb{I}_{SA})|\phi_b\rangle \leq \langle\phi_b|\phi_b\rangle < h^2/(2M).$$

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

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

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

$$|\phi_{a,\beta}(y_1,\ldots,y_h)\rangle := \gamma \sum_{z\in D^m} |z\rangle_{\mathcal{I}}|\mathbb{W}\rangle_S|y_1,\ldots,y_h\rangle_{R_Y}|\Omega_{\beta_1}(z_{y_1}),\ldots,\Omega_{\beta_h}(z_{y_h})\rangle_{R_X}|\xi''\rangle_{A-R},$$

(26)

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

$$(\Pi_{\mathcal{I},a} \otimes \mathbb{I}_{SA})|\phi_{a,\beta}(y_1,\ldots,y_h)\rangle = |\phi_{a,\beta}(y_1,\ldots,y_h)\rangle$$

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

Notice that, if $\beta_\ell = 1$, then the register $\mathcal{H}_{R_X(\ell)}$ 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 $|\phi_{a,0^k}(1,\ldots,h)\rangle/\gamma$ restricted to registers $\mathcal{H}_{\mathcal{I}}$ and $\mathcal{H}_{R_X}$, for these registers are not entangled with the other registers and we have

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

We have

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

48

Recall the states $|\psi_{x_1,\dots,x_i}\rangle \in \mathcal{H}_I$ from (25). We have

$$\sum_{z_y \in D} |z_y\rangle |\Psi(z_y)\rangle \propto \sum_{z_y \in D} |z_y\rangle \sum_{\substack{x \in X \\ z_{y,x}=1}} |x\rangle = \sum_{x \in X} \Big(\sum_{\substack{z_y \in D \\ z_{y,x}=1}} |z_y\rangle \Big) |x\rangle \propto \sum_{x \in X} |\psi_x\rangle |x\rangle \in \mathcal{T}^1_{I(y)} \otimes \mathcal{H}_{R_X(y)};$$

$$\sum_{z_y \in D} |z_y\rangle \propto |\psi_\emptyset\rangle \in \mathcal{T}^0_{I(y)} = \mathcal{H}^{(N)}_{I(y)}.$$

The claim follows from the definition of $\mathcal{H}_{\mathcal{I},a}$ (Section C.4). $\qquad\square$

C.6 Proof of Lemma 44

Reduction to the $p_{q_T,b}=0$ case. Let us first reduce the lemma to its special case when $p_{q_T,b}=0$. This reduction was used in [2] for a very similar problem. Recall that the final state of the algorithm $|\phi_{q_T}\rangle$ satisfies the symmetry $(U_{\mathcal{I}SO,\tau} \otimes \mathbb{I}_{A-O})|\phi_{q_T}\rangle = |\phi_{q_T}\rangle$ for all $\tau \in \mathbb{W}$, and note that, for $c \in \{a,b\}$, the state

$$|\phi^c_{q_T}\rangle := \frac{(\Pi_{\mathcal{I},c} \otimes \mathbb{I}_{SA})|\phi_{q_T}\rangle}{\|(\Pi_{\mathcal{I},c} \otimes \mathbb{I}_{SA})|\phi_{q_T}\rangle\|} = \frac{1}{\sqrt{p_{c,q_T}}}(\Pi_{\mathcal{I},c} \otimes \mathbb{I}_{SA})|\phi_{q_T}\rangle$$

satisfies the same symmetry. We have

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

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

$$\||\phi_{q_T}\rangle - |\phi^a_{q_T}\rangle\| = \sqrt{(1 - \sqrt{1 - p_{b,q_T}})^2 + (\sqrt{p_{b,q_T}})^2} \le \sqrt{2p_{b,q_T}} \qquad (27)$$

From now on, let us assume that $p_{b,q_T}=0$ and, thus, $|\psi_{q_T}\rangle = |\psi^a_{q_T}\rangle$. Lemma 36 and (27) states that this changes the success probability by at most $\sqrt{2p_{b,q_T}}$.

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

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

The algorithm makes its final measurement of the $\mathcal{H}_{\mathcal{I}}$ and $\mathcal{H}_O$ registers, ignoring all the other registers, therefore the success probability is completely determined by ρ''_{q_T}. Let us assume that the algorithm measures (and then discards) the $\mathcal{H}_{O_Y}$ register first, before measuring $\mathcal{H}_{\mathcal{I}}$ and $\mathcal{H}_{O_{X2}}$, and that the outcome of this measurement is $y \in Y$. Due to the symmetry, we get each outcome y with the same probability $1/M$.

Now the algorithm can discard the registers $\mathcal{H}_{I(y')}$ for all $y' \ne y$, as their content do not affect the success probability. We are left with

$$\rho''_{q_T,y} = M\mathrm{Tr}_{I(y'):\, y' \ne y}\big((\mathbb{I}_{\mathcal{I}O_{X2}} \otimes \langle y|_{O_Y})\rho''_{q_T}(\mathbb{I}_{\mathcal{I}O_{X2}} \otimes |y\rangle_{O_Y})\big),$$

which is a density matrix on the registers $\mathcal{H}_{I(y)}$ and $\mathcal{H}_{O_{X2}}$, and it satisfies

$$(\Pi_{I,a} \otimes \mathbb{I}_{O_{X2}})\rho''_{q_T,y} = \rho''_{q_T,y},$$
$$\forall \sigma \in \mathbb{S}_X : (U_{I,\sigma} \otimes U_{O_{X2},\sigma})\rho''_{q_T,y}(U_{I,\sigma} \otimes U_{O_{X2},\sigma})^{-1} = \rho''_{q_T,y}$$

(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 $\rho''_{q_T,y}$ 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 44, let us write z' instead of z_y and z'_x instead of $z_{y,x}$.

Let us now assume that the algorithm measures the $\mathcal{H}_{O_{X2}}$ register, obtaining $\{x_1, x_2\} \in X_2$, and only then measures $\mathcal{H}_I$. Due to the symmetry, the measurement yields each outcome $\{x_1, x_2\}$ with the same probability $1/\binom{N}{2}$, and let

$$\hat{\rho} := \binom{N}{2}(\mathbb{I}_I \otimes \langle\{x_1, x_2\}|_{O_{X2}})\rho''_{q_T,y}(\mathbb{I}_I \otimes |\{x_1, x_2\}\rangle_{O_{X2}})$$

be the density matrix of the register $\mathcal{H}_I$ after the measurement. Without loss of generality, let $\{x_1, x_2\} = \{1, 2\}$, and let $\hat{\mathbb{S}} := \mathbb{S}_{\{1,2\}} \times \mathbb{S}_{\{3,\ldots,N\}} < \mathbb{S}_X$ be the group of all permutations $\sigma \in \mathbb{S}_X$ that map $\{1, 2\}$ to itself. Now we have

$$\Pi_{I,a}\hat{\rho} = \hat{\rho} \quad \text{and} \quad \forall \sigma \in \hat{\mathbb{S}}: U_{I,\sigma}\hat{\rho}U_{I,\sigma}^{-1} = \hat{\rho}. \tag{28}$$

Let $\hat{\Pi}$ denote the projection to the subspace of $\mathcal{H}_I$ 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 $\sigma \in \hat{\mathbb{S}}$. 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 χ_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

$$\max_{|\chi\rangle} \left(\|\hat{\Pi}|\chi\rangle\|^2/\||\chi\rangle\|^2\right)$$

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 $\mathcal{T}_I^1$, namely, $\hat{U}_I := \Pi_{I,a}U_I$. Recall that we defined $\mathcal{T}_I^1$ to be the space spanned by all

$$|\psi_x\rangle = \frac{1}{\sqrt{\binom{N-1}{k-1}}} \sum_{\substack{z' \in D \\ z'_x = 1}} |z'\rangle.$$

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 $\mathbb{S}_X$ and its subgroup $\hat{\mathbb{S}}$. We already studied U_I as a representation of $\mathbb{S}_X$ in Section C.3. Since $\mathcal{T}_I^1 = \mathcal{H}_I^{(N)} \oplus \mathcal{H}_I^{(N-1,1)}$, the representation $\hat{U}_I$ of $\mathbb{S}_X$ 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 $\hat{\mathbb{S}}$, we need to restrict (N) and $(N-1, 1)$ from $\mathbb{S}_N$ to $\mathbb{S}_2 \times \mathbb{S}_{N-2}$. The Littlewood-Richardson rule gives us the decomposition of these restrictions:

$$(N) \downarrow (\mathbb{S}_2 \times \mathbb{S}_{N-2}) \qquad = ((2) \times (N-2));$$
$$(N-1, 1) \downarrow (\mathbb{S}_2 \times \mathbb{S}_{N-2}) = ((2) \times (N-2)) \oplus ((1,1) \times (N-2)) \oplus ((2) \times (N-3, 1)).$$

Hence, Schur's lemma and (28) imply that that eigenspaces of $\hat{\rho}$ are invariant under $U_{I,\sigma}$ for all $\sigma \in \hat{\mathbb{S}}$, 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 α, β;

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,

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

Therefore,

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

We also have

$$\||\psi(\alpha, \beta)\rangle\|^2 = \langle \psi(\alpha, \beta) | \psi(\alpha, \beta) \rangle$$
$$= 2\left(1 + \frac{k-1}{N-1}\right) |\alpha|^2 + (N-2)\left(1 + (N-3)\frac{k-1}{N-1}\right) |\beta|^2 + 2(N-2)\frac{k-1}{N-1}(\alpha\beta^* + \beta\alpha^*)$$
$$\geq \frac{|2\alpha + (k-2)\beta|^2}{2}. \quad (29)$$

If $\alpha\beta^* \geq 0$, the inequality in (29) follows by showing that coefficients of $|\alpha|^2$, $|\beta|^2$, and $\alpha\beta^*$ on the left hand side are all larger than corresponding coefficients on the right hand

side. Otherwise, without loss of generality, we can assume that $\alpha = 1$ and $\beta < 0$, and the inequality follows by inspecting the extreme point of the quadratic polynomial (in β) that is obtained by subtracting the right hand side from the left hand side. Therefore,

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

In the second case, $\Pi(|\psi_1\rangle - |\psi_2\rangle) = 0$ because basis states $|1, 1, z_3', \ldots, z_N'\rangle$ have the same amplitude in $|\psi_1\rangle$ and $|\psi_2\rangle$.

In the third case, it suffices to consider a state of the form $|\psi_3\rangle - |\psi_4\rangle$, because $\{U_{I,\sigma}(|\psi_3\rangle - |\psi_4\rangle) \colon \sigma \in \hat{\mathbb{S}}\}$ spans the whole eigenspace and $\hat{\Pi}$ and $U_{I,\sigma}$ commute. Then,

$$\hat{\Pi}(|\psi_3\rangle - |\psi_4\rangle) = \frac{1}{\sqrt{\binom{N-1}{k-1}}} \sum_{\substack{z_5',\ldots,z_N' \in \{0,1\} \\ z_5'+\ldots+z_N'=k-3}} (|1,1,1,0,z_5',\ldots,z_N'\rangle - |1,1,0,1,z_5',\ldots,z_N'\rangle)$$

and

$$\|\hat{\Pi}(|\psi_3\rangle - |\psi_4\rangle)\|^2 = 2\frac{\binom{N-4}{k-3}}{\binom{N-1}{k-1}} = 2\frac{(k-1)(k-2)(N-k)}{(N-1)(N-2)(N-3)}.$$

We also have

$$\||\psi_3\rangle - |\psi_4\rangle\|^2 = 2 - \langle\psi_3|\psi_4\rangle = 2 - 2\frac{k-1}{N-1} = 2\frac{N-k}{N-1}.$$

Hence,

$$\frac{\|\hat{\Pi}(|\psi_3\rangle - |\psi_4\rangle)\|^2}{\||\psi_3\rangle - |\psi_4\rangle\|^2} = \frac{(k-2)(k-3)}{(N-2)(N-3)} = O\left(\frac{k^2}{N^2}\right).$$

$\square$

C.7 Reduction of Lemma 46 to the $|Y| = 1$ case

First, instead of the oracle $\mathcal{O}_V$ given by (18), we define

$$\mathcal{O}_V(z)|y,x,b\rangle := (-1)^{b \cdot z_{y,x}}|y,x,b\rangle.$$

Both definitions are equivalently powerful as one is obtained from another by two Hadamard gates on the register $\mathcal{H}_B$.

For all $z_y \in D$, let

$$\mathcal{O}_V''(z_y) := \mathbb{I}_{Q_X} - 2\sum_{\substack{x \in X \\ z_{y,x}=1}} |x\rangle\langle x|_{Q_X} \qquad \text{and} \qquad \mathcal{O}_F''(z_y) := \mathbb{I}_{Q_X} - 2|\Psi(z_y)\rangle\langle\Psi(z_y)|_{Q_X}$$

act on $\mathcal{H}_{Q_X}$, so that we have

$$\begin{aligned}
\mathcal{O}_V' &= \sum_{z \in D^M} |z\rangle\langle z|_{\mathcal{I}} \otimes \sum_{y \in Y} |y\rangle\langle y|_{Q_Y} \otimes \mathcal{O}_V''(z_y) \otimes |1\rangle\langle 1|_B + \mathbb{I}_{\mathcal{I}Q} \otimes |0\rangle\langle 0|_B, \\
\mathcal{O}_F' &= \sum_{z \in D^M} |z\rangle\langle z|_{\mathcal{I}} \otimes \sum_{y \in Y} |y\rangle\langle y|_{Q_Y} \otimes \mathcal{O}_F''(z_y) \otimes \mathbb{I}_B.
\end{aligned} \tag{30}$$

Let
$$\rho_q''' = \rho_{q,00}' \otimes |0\rangle\langle 1|_B + \rho_{q,01}' \otimes |0\rangle\langle 1|_B + \rho_{q,10}' \otimes |1\rangle\langle 0|_B + \rho_{q,11}' \otimes |1\rangle\langle 1|_B$$

be the state of the algorithm corresponding to the $\mathcal{H}_{\mathcal{I}}$, $\mathcal{H}_Q$, and $\mathcal{H}_B$ registers right before the $(q+1)$-th oracle call ($\mathcal{O}_V$ or $\mathcal{O}_F$). Note that $\rho_q = \mathrm{Tr}_{QB}(\rho_q''')$ and, since oracles are the only gates of the algorithm that interact with the $\mathcal{H}_{\mathcal{I}}$ register, $\rho_{q+1} = \mathrm{Tr}_{QB}(\mathcal{O}'\rho_q'''\mathcal{O}')$.

Notice that $|p_{b,q} - p_{b,q+1}| = |p_{a,q} - p_{a,q+1}|$, therefore let us deal with $p_{a,q}$ instead. We have

$$|p_{a,q} - p_{a,q+1}| = \mathrm{Tr}\big(\Pi_{\mathcal{I},a}(\rho_q - \rho_{q+1})\big) = \mathrm{Tr}\big((\Pi_{\mathcal{I},a} \otimes \mathbb{I}_{QB})(\rho_q''' - \mathcal{O}'\rho_q'''\mathcal{O}')\big), \qquad (31)$$

which for the oracle $\mathcal{O}_V$ equals

$$\mathrm{Tr}\big((\Pi_{\mathcal{I},a} \otimes \mathbb{I}_Q)(\rho_{q,11}' - \tilde{\mathcal{O}}_V'\rho_{q,11}'\tilde{\mathcal{O}}_V')\big),$$

where $\tilde{\mathcal{O}}_V' = (\mathbb{I}_{\mathcal{I}Q} \otimes \langle 1|_B)\mathcal{O}_V'(\mathbb{I}_{\mathcal{I}Q} \otimes |1\rangle_B)$. Therefore, without loss of generality, we assume that the state of $\mathcal{H}_B$ is always $|1\rangle$ throughout the execution of the algorithm. In turn, we assume that $\mathcal{O}_V'$ and $\mathcal{O}_F'$ in (30) act only on $\mathcal{H}_{\mathcal{I}} \otimes \mathcal{H}_Q$, and we take ρ_q' instead of ρ_q''' and $\mathbb{I}_Q$ instead of $\mathbb{I}_{QB}$ in (31).

Since $(\tau(z))_{\tau(y,x)} = 1$ if and only if $z_{x,y} = 1$, we have $U_{\mathcal{I}Q,\tau}\mathcal{O}'U_{\mathcal{I}Q,\tau}^{-1} = \mathcal{O}'$ for all $\tau \in \mathbb{W}$, and recall that the same symmetry holds for ρ_q', namely, (22). Hence, for all $y \in Y$,

$$\rho_{q,y}' = M(\mathbb{I}_{\mathcal{I}Q_X} \otimes \langle y|_{Q_Y})\rho_q'(\mathbb{I}_{\mathcal{I}Q_X} \otimes |y\rangle_{Q_Y})$$

has trace one and (31) equals

$$M\mathrm{Tr}\big((\mathbb{I}_{\mathcal{I}Q_X} \otimes \langle y|_{Q_Y})(\Pi_{\mathcal{I},a} \otimes \mathbb{I}_Q)(\rho_q' - \mathcal{O}'\rho_q'\mathcal{O}')(\mathbb{I}_{\mathcal{I}Q_X} \otimes |y\rangle_{Q_Y})\big)$$
$$= \mathrm{Tr}\bigg((\Pi_{\mathcal{I},a} \otimes \mathbb{I}_{Q_X})\Big(\rho_{q,y}' - \big(\sum_{z \in D^M} |z\rangle\langle z|_{\mathcal{I}} \otimes \mathcal{O}''(z_y)\big)\rho_{q,y}'\big(\sum_{z \in D^M} |z\rangle\langle z|_{\mathcal{I}} \otimes \mathcal{O}''(z_y)\big)\Big)\bigg). \qquad (32)$$

Without loss of generality, let $y = 1$, and let us write

$$\sum_{z \in D^M} |z\rangle\langle z|_{\mathcal{I}} = \sum_{z_1 \in D} |z_1\rangle\langle z_1|_{I(1)} \otimes \mathbb{I}_I^{\otimes(M-1)}.$$

Recall that $\Pi_{\mathcal{I},a} = \Pi_{I,a}^{\otimes M}$. Therefore, for

$$\hat{\rho}_{q,1}' := \mathrm{Tr}_{I(2),\dots,I(M)}\big((\mathbb{I}_{I(1)} \otimes \Pi_{I,a}^{\otimes(M-1)} \otimes \mathbb{I}_{Q_X})\rho_{q,y}'\big),$$

(31) and (32) are equal to

$$\mathrm{Tr}\bigg((\Pi_{I(1),a} \otimes \mathbb{I}_{Q_X})\Big(\hat{\rho}_{q,1}' - \big(\sum_{z_1 \in D} |z_1\rangle\langle z_1|_{I(1)} \otimes \mathcal{O}''(z_1)\big)\hat{\rho}_{q,1}'\big(\sum_{z_1 \in D} |z_1\rangle\langle z_1|_{I(1)} \otimes \mathcal{O}''(z_1)\big)\Big)\bigg). \qquad (33)$$

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

D   Proof of Lemma 46 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 $z_{y,x}$. Also, now we denote the permutations in $\mathbb{S}_N$ with π instead of σ.

We will consider the following representations of $\mathbb{S}_N$:

1. The computational basis of $\mathcal{H}_Q$ is labeled by $x \in \{1, \ldots, N\} = X$. We define the action of $\pi \in \mathbb{S}_N$ on $\mathcal{H}_Q$ via the unitary $U_{Q,\pi}|x\rangle := |\pi(x)\rangle$. U_Q is known as the *natural representation* of $\mathbb{S}_N$, 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 $\mathbb{S}_N$ 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 $\sum_{x=1}^N z_i = k$. In Section C.3 we already defined and studied the representation U_I: for $\pi \in \mathbb{S}_N$,

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

   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 $\mathbb{S}_N$ 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 $\mathbb{S}_N$.

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 $\Pi_I^{(N)}$, $\Pi_I^{(N-1,1)}$, ..., $\Pi_I^{(N-k,k)}$ denote, respectively, the projectors on $\mathcal{H}_I^{(N)}$, $\mathcal{H}_I^{(N-1,1)}$, ..., $\mathcal{H}_I^{(N-k,k)}$. 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 [21]).

Let us also denote

$$\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{\geq 2}} := \mathbb{I}_Q \otimes \sum_{j=2}^k \Pi_I^{(N-j,j)} = \left(\Pi_Q^{(N)} + \Pi_Q^{(N-1,1)} \right) \otimes \sum_{j=2}^k \Pi_I^{(N-j,j)}, \tag{34}$$

$$\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{<2}} := \mathbb{I}_{QI} - \Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{\geq 2}} = \left(\Pi_Q^{(N)} + \Pi_Q^{(N-1,1)} \right) \otimes \left(\Pi_I^{(N)} + \Pi_I^{(N-1,1)} \right), \tag{35}$$

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 (30), we have to consider

$$\mathcal{O}_V = \sum_{z \in D} \Big(\sum_{\substack{x \in X \\ z_x = 0}} |x\rangle\langle x| - \sum_{\substack{x \in X \\ z_x = 1}} |x\rangle\langle x| \Big)_Q \otimes |z\rangle\langle z|_I,$$

$$\mathcal{O}_F = \sum_{z \in D} \Big(\mathbb{I} - |\Psi(z)\rangle\langle\Psi(z)| \Big)_Q \otimes |z\rangle\langle z|_I,$$

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 $\pi \in \mathbb{S}_N$. Equivalently to (33), it suffices to prove that

$$\left| \text{Tr}\big(\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{<2}} (\rho - \mathcal{O}\rho\mathcal{O}) \big) \right| \leq O(\max\{\sqrt{k/N}, \sqrt{1/k}\})$$

for every density operator ρ on $\mathcal{H}$ that satisfies $U_\pi \rho U_\pi^{-1} = \rho$ for all $\pi \in \mathbb{S}_N$ 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_π for all $\pi \in \mathbb{S}_N$), let $U|_{\mathcal{H}'}$ be U restricted to this subspace (note: $U|_{\mathcal{H}'}$ is a representation of $\mathbb{S}_N$). 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 μ is invariant under U, and $U|_\mu$ in an irrep of $\mathbb{S}_N$. Hence, it suffices to show the following.

Lemma 47 *For every subspace $\mu \subset \mathcal{H}$ such that $U|_\mu$ is an irrep and for μ' being the subspace that μ is mapped to by $\mathcal{O}_V$ or $\mathcal{O}_F$, we have*

$$\frac{1}{\dim \mu} \left| \text{Tr}(\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{\geq 2}} (\Pi_\mu - \Pi_{\mu'})) \right| \leq O(\max\{\sqrt{k/N}, \sqrt{1/k}\}). \tag{36}$$

In order to prove Lemma 47, 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 θ present in U, let $\hat{\Pi}_\theta$ be a projector on the space corresponding to all instances of θ in U.

Approach 1: via the tensor product of irreps. We know that $U = U_Q \otimes U_I$ and we already know how U_Q and U_I decomposes into irreps. Thus, all we need to see is how, for $j \in \{0, \ldots, k\}$, $(N)_Q \otimes (N-j, j)_I$ and $(N-1, 1)_Q \otimes (N-j, j)_I$ decompose into irreps (we use subscripts Q and I here to specify which spaces these irreps act on, namely, $\mathcal{H}_Q$ and $\mathcal{H}_I$, 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*

$$
\begin{aligned}
&(N-1, 1) \otimes (N-j, j) \\
&= (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j-1, j+1) \oplus (N-j-1, j, 1),
\end{aligned}
$$

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

Proof. We use Expression 2.9.5 of [25], which, for $j \in \{2, \ldots, k\}$, gives us

$$
\begin{aligned}
(N-1, 1) \otimes (N-j, j) &= (N-j, j) \downarrow (\mathbb{S}_{N-1} \times \mathbb{S}_1) \uparrow \mathbb{S}_N \ominus (N-j, j) \downarrow \mathbb{S}_N \uparrow \mathbb{S}_N \\
&= ((N-j, j-1) \times (1)) \uparrow \mathbb{S}_N \oplus ((N-j-1, j) \times (1)) \uparrow \mathbb{S}_N \ominus (N-j, j) \\
&= (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j, j) \\
&\quad \oplus (N-j-1, j+1) \oplus (N-j-1, j, 1) \ominus (N-j, j) \\
&= (N-j+1, j-1) \oplus (N-j, j) \oplus (N-j, j-1, 1) \oplus (N-j-1, j+1) \\
&\quad \oplus (N-j-1, j, 1)
\end{aligned}
$$

and, similarly, for $j = 1$, gives us

$$
\begin{aligned}
(N-1, 1) \otimes (N-1, 1) &= (N-1, 1) \downarrow (\mathbb{S}_{N-1} \times \mathbb{S}_1) \uparrow \mathbb{S}_N \ominus (N-1, 1) \downarrow \mathbb{S}_N \uparrow \mathbb{S}_N \\
&= (N) \oplus (N-1, 1) \oplus (N-2, 2) \oplus (N-2, 1, 1). \qquad \square
\end{aligned}
$$

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 θ present in $(N-\ell, \ell)_Q \otimes (N-j, j)_I$, let

$$
\Pi_\theta^{(N-\ell, \ell)_Q \otimes (N-j, j)_I} := \hat{\Pi}_\theta \big(\Pi_Q^{(N-\ell, \ell)} \otimes \Pi_I^{(N-j, j)} \big),
$$

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

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

$$H_0 = \big\{(x,z) \in X \times D \,:\, z_x = 0\big\} \qquad \text{and} \qquad H_1 = \big\{(x,z) \in X \times D \,:\, z_x = 1\big\},$$

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

$$\mathcal{H}^{(0)} = \mathcal{H}^{(0,s)} \oplus \mathcal{H}^{(0,t)} \qquad \text{and} \qquad \mathcal{H}^{(1)} = \mathcal{H}^{(1,s)} \oplus \mathcal{H}^{(1,t)},$$

where

$$\mathcal{H}^{(0,s)} := \operatorname{span}\Big\{ \sum_{x \,:\, z_x=0} |x,z\rangle \,:\, z \in D \Big\} \qquad \text{and} \qquad \mathcal{H}^{(1,s)} := \operatorname{span}\Big\{ \sum_{x \,:\, z_x=1} |x,z\rangle \,:\, z \in D \Big\},$$

and $\mathcal{H}^{(0,t)} := \mathcal{H}^{(0)} \cap (\mathcal{H}^{(0,s)})^\perp$ and $\mathcal{H}^{(1,t)} := \mathcal{H}^{(1)} \cap (\mathcal{H}^{(1,s)})^\perp$.

Note that, for a given z, $\sum_{z \,:\, z_x=1} |x\rangle = \sqrt{k}\,|\Psi(z)\rangle$. Therefore, the query $\mathcal{O}_F$ acts on $\mathcal{H}^{(1,s)}$ as the minus identity and on $\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1,t)}$ as the identity. Meanwhile, $\mathcal{O}_V$ 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 Π^σ be the projector on the space $\mathcal{H}^\sigma$, and let U^σ be the restriction of U to $\mathcal{H}^\sigma$. Let V_π^σ be U_π^σ restricted to $\pi \in \mathbb{S}_k \times \mathbb{S}_{N-k}$ and the space

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

V^σ is a representation of $\mathbb{S}_k \times \mathbb{S}_{N-k}$. One can see that

$$|\mathbb{S}_N|/|\mathbb{S}_k \times \mathbb{S}_{N-k}| = \dim \mathcal{H}^\sigma / \dim \tilde{\mathcal{H}}^\sigma,$$

so we have $U^\sigma = V^\sigma \uparrow \mathbb{S}_N$. In order to see how U^σ decomposes into irreps, we need to see how V^σ 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 $\tilde{\mathcal{H}}^{(0,s)}$ and $\tilde{\mathcal{H}}^{(1,s)}$, respectively. That is, $V^{(0,s)} \cong V^{(1,s)} \cong (k) \times (N-k)$. Now, note that

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

The group $\mathbb{S}_k$ (in $\mathbb{S}_k \times \mathbb{S}_{N-k}$) acts trivially on $\tilde{\mathcal{H}}^{(0)}$, while and the action of $\mathbb{S}_{N-k}$ on $\tilde{\mathcal{H}}^{(0)}$ defines the natural representation of $\mathbb{S}_{N-k}$. 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 $U^{(0,s)} = V^{(0,s)} \uparrow \mathbb{S}_N$ and $U^{(1,s)} = V^{(1,s)} \uparrow \mathbb{S}_N$ into irreps are given via (24). For $U^{(0,t)} = V^{(0,t)} \uparrow \mathbb{S}_N$ and $U^{(1,t)} = V^{(1,t)} \uparrow \mathbb{S}_N$, the Littlewood-Richardson rule gives us, respectively,

$$\begin{aligned}
((k) \times (N-k-1,1)) \uparrow \mathbb{S}_N ={} & (N-1,1) \oplus (N-2,2) \oplus (N-2,1,1) \\
& \oplus (N-3,3) \oplus (N-3,2,1) \oplus (N-4,4) \oplus (N-4,3,1) \oplus \ldots \\
& \oplus (N-k,k) \oplus (N-k,k-1,1) \oplus (N-k-1,k+1) \oplus (N-k-1,k,1)
\end{aligned}$$

and

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

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

D.3 Significant irreps

We noted in Section D.2 that $\mathcal{O}_F$ acts on $\mathcal{H}^{(1,s)}$ as the minus identity and on $\mathcal{H}^{(0)} \oplus \mathcal{H}^{(1,t)}$ as the identity and $\mathcal{O}_V$ acts on $\mathcal{H}^{(1)}$ as the minus identity and on $\mathcal{H}^{(0)}$ as the identity. This means that, if μ 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

$$\left| \mathrm{Tr}(\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{\geq 2}}(\Pi_\mu - \Pi_{\mu'})) \right| = \left| \mathrm{Tr}(\Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{<2}}(\Pi_\mu - \Pi_{\mu'})) \right|. \tag{37}$$

Hence we need to consider only μ such that $U|_\mu$ is isomorphic to an irrep present in both

$$\big((N) \oplus (N-1,1)\big)_Q \otimes \big((N) \oplus (N-1,1)\big)_I \quad \text{and} \quad \big((N) \oplus (N-1,1)\big)_Q \otimes \bigoplus_{j=2}^{k} (N-j,j)_I,$$

as otherwise the expression (36) 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

$$\Pi_{(N-1,1)}^{(N)_Q \otimes (N-1,1)_I}, \ \Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N)_I}, \ \Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N-1,1)_I}, \ \Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N-2,2)_I},$$
$$\Pi_{(N-2,2)}^{(N)_Q \otimes (N-2,2)_I}, \ \Pi_{(N-2,2)}^{(N-1,1)_Q \otimes (N-1,1)_I}, \ \Pi_{(N-2,2)}^{(N-1,1)_Q \otimes (N-2,2)_I}, \ \Pi_{(N-2,2)}^{(N-1,1)_Q \otimes (N-3,3)_I}, \tag{38}$$
$$\Pi_{(N-2,1,1)}^{(N-1,1)_Q \otimes (N-1,1)_I}, \ \Pi_{(N-2,1,1)}^{(N-1,1)_Q \otimes (N-2,2)_I},$$

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

$$\Pi_{(N-1,1)}^{(0,s)}, \ \Pi_{(N-1,1)}^{(0,t)}, \ \Pi_{(N-1,1)}^{(1,s)}, \ \Pi_{(N-1,1)}^{(1,t)},$$
$$\Pi_{(N-2,2)}^{(0,s)}, \ \Pi_{(N-2,2)}^{(0,t)}, \ \Pi_{(N-2,2)}^{(1,s)}, \ \Pi_{(N-2,2)}^{(1,t)},$$
$$\Pi_{(N-2,1,1)}^{(0,t)}, \ \Pi_{(N-2,1,1)}^{(1,t)}.$$

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 μ, and the expression (36) 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 (36) 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 $\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}$ be, up to a global phase, the unique operator of rank $\dim(N-1, 1)$ such that

$$\left(U_{(N-1,1)}^{(a_1,b_1)}\right)_\pi = \Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\left(U_{(N-1,1)}^{(a_2,b_2)}\right)_\pi\left(\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\right)^*.$$

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

$$\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\left(\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\right)^* = \Pi_{(N-1,1)}^{(a_1,b_1)}, \qquad \left(\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\right)^*\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)} = \Pi_{(N-1,1)}^{(a_2,b_2)}.$$

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

$$\left(\Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\right)^* = \Pi_{(N-1,1)}^{(a_2,b_2)\leftarrow(a_1,b_1)} \qquad \text{and} \qquad \Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}\Pi_{(N-1,1)}^{(a_2,b_2)\leftarrow(a_3,b_3)} = \Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_3,b_3)}$$

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

Fix a_3 and b_3, and note that

$$\left(\Pi_{(N-1,1)}^{(a_3,b_3)\leftarrow(a_1,b_1)}\right)^*\Pi_{(N-1,1)}^{(a_3,b_3)\leftarrow(a_2,b_2)} = \Pi_{(N-1,1)}^{(a_1,b_1)\leftarrow(a_2,b_2)}$$

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

$$\Pi_{(N-1,1)}^{\leftarrow(a_1,b_1)} := \Pi_{(N-1,1)}^{(a_3,b_3)\leftarrow(a_1,b_1)}.$$

Fact 1 *Let $\mu \subset \mathcal{H}$ be such that $U|_\mu$ is an irrep isomorphic to $(N-1, 1)$ and let Π_μ be the projector on this subspace. There exists, up to a global phase, a unique vector $\gamma = (\gamma_{0,s}, \gamma_{0,t}, \gamma_{1,s}, \gamma_{1,t})$ such that $\Pi_\mu = \bar{\Pi}_\gamma^*\bar{\Pi}_\gamma$, where*

$$\bar{\Pi}_\gamma = \left(\gamma_{0,s}\Pi_{(N-1,1)}^{\leftarrow(0,s)} + \gamma_{0,t}\Pi_{(N-1,1)}^{\leftarrow(0,t)} + \gamma_{1,s}\Pi_{(N-1,1)}^{\leftarrow(1,s)} + \gamma_{1,t}\Pi_{(N-1,1)}^{\leftarrow(1,t)}\right).$$

The norm of the vector γ is 1. The converse also holds: for any unit vector γ, $\bar{\Pi}_\gamma^*\bar{\Pi}_\gamma$ 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 $\hat{\Pi}_{(N-1,1)}$ are, respectively,

$$\mathcal{O}_V|_{(N-1,1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \mathcal{O}_F|_{(N-1,1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

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

$$\Pi_\mu = \begin{pmatrix} \gamma_{0,s}^* \\ \gamma_{0,t}^* \\ \gamma_{1,s}^* \\ \gamma_{1,t}^* \end{pmatrix} \cdot \begin{pmatrix} \gamma_{0,s} & \gamma_{0,t} & \gamma_{1,s} & \gamma_{1,t} \end{pmatrix} = \begin{pmatrix} |\gamma_{0,s}|^2 & \gamma_{0,s}^*\gamma_{0,t} & \gamma_{0,s}^*\gamma_{1,s} & \gamma_{0,s}^*\gamma_{1,t} \\ \gamma_{0,t}^*\gamma_{0,s} & |\gamma_{0,t}|^2 & \gamma_{0,t}^*\gamma_{1,s} & \gamma_{0,t}^*\gamma_{1,t} \\ \gamma_{1,s}^*\gamma_{0,s} & \gamma_{1,s}^*\gamma_{0,t} & |\gamma_{1,s}|^2 & \gamma_{1,s}^*\gamma_{1,t} \\ \gamma_{1,t}^*\gamma_{0,s} & \gamma_{1,t}^*\gamma_{0,t} & \gamma_{1,t}^*\gamma_{1,s} & |\gamma_{1,t}|^2 \end{pmatrix},$$
(39)

and note that

$$|\gamma_{a,b}|^2 = \operatorname{Tr}\big(\Pi_\mu \Pi^{(a,b)}_{(N-1,1)}\big) \Big/ \dim(N-1,1).$$

From (38), one can see that

$$\hat{\Pi}_{(N-1,1)} \Pi_{\mathcal{H}_Q \otimes \mathcal{S}_{\geq 2}} = \Pi^{(N-1,1)Q \otimes (N-2,2)I}_{(N-1,1)}.$$

Hence, for the space μ, the desired inequality (36) becomes

$$\frac{1}{\dim(N-1,1)} \Big| \operatorname{Tr}\big(\Pi^{(N-1,1)Q \otimes (N-2,2)I}_{(N-1,1)}(\Pi_\mu - \Pi_{\mu'})\big) \Big| \leq O(\max\{\sqrt{k/N}, \sqrt{1/k}\}). \quad (40)$$

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

In the same transporter basis, let

$$\Pi^{(N-1,1)Q \otimes (N-2,2)I}_{(N-1,1)} = \begin{pmatrix} |\beta_{0,s}|^2 & \beta_{0,s}^*\beta_{0,t} & \beta_{0,s}^*\beta_{1,s} & \beta_{0,s}^*\beta_{1,t} \\ \beta_{0,t}^*\beta_{0,s} & |\beta_{0,t}|^2 & \beta_{0,t}^*\beta_{1,s} & \beta_{0,t}^*\beta_{1,t} \\ \beta_{1,s}^*\beta_{0,s} & \beta_{1,s}^*\beta_{0,t} & |\beta_{1,s}|^2 & \beta_{1,s}^*\beta_{1,t} \\ \beta_{1,t}^*\beta_{0,s} & \beta_{1,t}^*\beta_{0,t} & \beta_{1,t}^*\beta_{1,s} & |\beta_{1,t}|^2 \end{pmatrix}. \quad (41)$$

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:

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

We have

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

so, for this space, the inequality (40) becomes

$$\big| e^{i\phi}\beta_{1,b_1}^*\beta_{0,b_0} + e^{-i\phi}\beta_{0,b_0}^*\beta_{1,b_1} \big| \leq O(\max\{\sqrt{k/N}, \sqrt{1/k}\}).$$

Since this has to hold for all b_0, b_1, and ϕ (in particular, consider b_0 and b_1 that maximize $|\beta_{1,b_1}^*\beta_{0,b_0}|$), we must have either

$$|\beta_{1,s}|^2 + |\beta_{1,t}|^2 \leq O(\max\{k/N, 1/k\}) \qquad \text{or} \qquad |\beta_{1,s}|^2 + |\beta_{1,t}|^2 \geq 1 - O(\max\{k/N, 1/k\}), \quad (42)$$

and note that

$$|\beta_{1,s}|^2 + |\beta_{1,t}|^2 = \operatorname{Tr}\big(\Pi^{(N-1,1)Q \otimes (N-2,2)I}_{(N-1,1)} \cdot \Pi^{(1)}\big) \Big/ \dim(N-1,1).$$

The condition (42) is necessary, but it is also sufficient for (40). Because, if it holds, then $|\beta_{1,b_1}^*\beta_{0,b_0}| \leq O(\max\{\sqrt{k/N}, \sqrt{1/k}\})$ for all $b_0, b_1 \in \{s, t\}$ and, clearly, $|\gamma_{1,b_1}^*\gamma_{0,b_0}| \in O(1)$ for all unit vectors γ. Therefore, if we plug (39) and (41) into (40), the inequality is satisfied.

Necessary and sufficient condition for the query $\mathcal{O}_F$. Almost identical analysis shows that, in order for Lemma 47 to hold when $U|_\mu$ is isomorphic to $(N-1,1)$ and we apply $\mathcal{O}_F$, it is necessary and sufficient that

$$|\beta_{1,s}|^2 \le O(\max\{k/N, 1/k\}) \qquad \text{or} \qquad |\beta_{1,s}|^2 \ge 1 - O(\max\{k/N, 1/k\}). \qquad (43)$$

Note that

$$|\beta_{1,s}|^2 = \mathrm{Tr}\big(\Pi_{(N-1,1)}^{(N-1,1)Q\otimes(N-2,2)I} \cdot \Pi_{(N-1,1)}^{(1,s)}\big)\Big/ \dim(N-1,1).$$

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 (37). Mainly, we do that because the space $\mathcal{H}_Q \otimes \mathcal{S}_{\ge 2}$ contains three instances of irrep $(N-2,2)$, while $\mathcal{H}_Q \otimes \mathcal{S}_{<2}$ contains only one. From (38) we get

$$\hat{\Pi}_{(N-2,2)}\Pi_{\mathcal{H}_Q\otimes\mathcal{S}_{<2}} = \Pi_{(N-2,2)}^{(N-1,1)Q\otimes(N-1,1)I} \qquad \text{and} \qquad \hat{\Pi}_{(N-2,1,1)}\Pi_{\mathcal{H}_Q\otimes\mathcal{S}_{<2}} = \Pi_{(N-2,1,1)}^{(N-1,1)Q\otimes(N-1,1)I}.$$

Condition for the query $\mathcal{O}_V$. An analysis analogous to that of the irrep $(N-1,1)$ shows that, in order for the desired inequality (36) to hold for query $\mathcal{O}_V$ and irreps $(N-2,2)$ and $(N-2,1,1)$, it is sufficient to have

$$\frac{\mathrm{Tr}\big(\Pi_{(N-2,2)}^{(N-1,1)Q\otimes(N-1,1)I} \cdot \Pi^{(1)}\big)}{\dim(N-2,2)} \le O(k/N) \qquad \text{and} \qquad \frac{\mathrm{Tr}\big(\Pi_{(N-2,1,1)}^{(N-1,1)Q\otimes(N-1,1)I} \cdot \Pi^{(1)}\big)}{\dim(N-2,1,1)} \le O(k/N).$$

Let us prove this. Consider irrep $(N-2,2)$ and the hook-length formula gives us $\dim(N-2,2) = N(N-3)/2$. We have

$$\mathrm{Tr}\big(\Pi_{(N-2,2)}^{(N-1,1)Q\otimes(N-1,1)I} \cdot \Pi^{(1)}\big) \le \mathrm{Tr}\big((\Pi_Q^{(N-1,1)} \otimes \Pi_I^{(N-1,1)}) \cdot \Pi^{(1)}\big),$$

and we can evaluate the right hand side of this exactly. $\Pi^{(1)}$ 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 $\Pi_Q^{(N-1,1)}$ are all the same and equal to $\frac{N-1}{N}$. The diagonal entries of $\Pi_I^{(N-1,1)}$ are also all the same, because $\Pi_I^{(N-1,1)}$ projects to an eigenspace of the Johnson scheme. More precisely, we have $\mathrm{Tr}(\Pi_I^{(N-1,1)}) = \dim(N-1,1) = N-1$, therefore the diagonal entries of $\Pi_I^{(N-1,1)}$ are $(N-1)/\binom{N}{k}$. Hence, the diagonal entries of $\Pi_Q^{(N-1,1)} \otimes \Pi_I^{(N-1,1)}$ are $(N-1)^2/(N\binom{N}{k})$, implying that

$$\mathrm{Tr}\big((\Pi_Q^{(N-1,1)} \otimes \Pi_I^{(N-1,1)})\Pi^{(1)}\big) = \frac{k(N-1)^2}{N}$$

and, in turn,

$$\frac{\mathrm{Tr}\big(\Pi_{(N-2,2)}^{(N-1,1)Q\otimes(N-1,1)I}\Pi^{(1)}\big)}{\dim(N-2,2)} \le \frac{2k(N-1)^2}{N^2(N-3)} \in O(k/N)$$

as required. The same argument works for irrep $(N-2,1,1)$ as, by the hook-length formula, $\dim(N-2,1,1) = (N-1)(N-2)/2 = \dim(N-2,2) + 1$.

Condition for the query $\mathcal{O}_F$. As we mentioned in the very end of Section D.3, $\mathcal{O}_F$ affects no space μ 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

$$\frac{\text{Tr}\big(\Pi_{(N-2,2)}^{(N-1,1)_Q \otimes (N-1,1)_I}\Pi^{(1,s)}\big)}{\dim(N-2,2)} \leq \frac{\text{Tr}\big(\Pi_{(N-2,2)}^{(N-1,1)_Q \otimes (N-1,1)_I}\Pi^{(1)}\big)}{\dim(N-2,2)} \leq O(k/N),$$

which, similarly to the condition (43) for irrep $(N-1,1)$, is sufficient to show that Lemma 47 holds for irrep $(N-2,2)$ and the query $\mathcal{O}_F$.

D.6 Solution for irrep $(N-1,1)$

Recall that conditions (42) and (43) are sufficient for Lemma 47 to hold for the queries $\mathcal{O}_V$ and $\mathcal{O}_F$, respectively. Hence, it suffices for us to show that

$$\frac{\text{Tr}\big(\Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N-2,2)_I} \cdot \Pi^{(1)}\big)}{\dim(N-1,1)} \geq \frac{\text{Tr}\big(\Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N-2,2)_I} \cdot \Pi_{(N-1,1)}^{(1,s)}\big)}{\dim(N-1,1)} =$$
$$= \frac{k-1}{k} \cdot \frac{N(N-k-1)}{(N-1)(N-2)} \geq 1 - O(\max\{k/N, 1/k\}).$$

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

$$\Pi_{(N-1,1)}^{(N-1,1)_Q \otimes (N-2,2)_I} \cdot \Pi_{(N-1,1)}^{(1,s)} = (\mathbb{I}_Q \otimes \Pi_I^{(N-2,2)}) \cdot \Pi_{(N-1,1)}^{(1,s)},$$

and let us evaluate the trace of the latter. We briefly mentioned before that $\Pi_I^{(N)}$, $\Pi_I^{(N-1,1)}$, ..., $\Pi_I^{(N-k,k)}$ 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 $z, z' \in D$, 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 $i \in \{0, 1, \ldots, k\}$, let

$$A_i^I = \sum_{\substack{z,z' \in D \\ |z-z'|=i}} |z\rangle\langle z'|,$$

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 [21, Chapter 7]).

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

$$C_j^I = \sum_{i=0}^{k-j} \binom{k-i}{j} A_i \quad \text{for all } j \quad \text{and} \quad A_i^I = \sum_{j=k-i}^{k} (-1)^{j-k+i} \binom{j}{k-i} C_j \quad \text{for all } i.$$
$$(44)$$

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

$$C_j^I = \sum_{h=0}^{j} \binom{N-j-h}{N-k-h}\binom{k-h}{j-h}\Pi_I^{(N-h,h)} \quad \text{for all } j. \tag{45}$$

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

Here, however, we are interested in the opposite: expressing $\Pi_I^{(N-h,h)}$ as a linear combination of A_i^I. From (45) one can see that

$$\Pi_I^{(N-h,h)} = (N-2h+1)\sum_{j=0}^{h}(-1)^{j-h}\frac{\binom{k-j}{h-j}}{(k-j+1)\binom{N-j-h+1}{N-k-h}}C_j^I \tag{46}$$

for $h = 0, 1, 2$. We are interested particularly in $\Pi_I^{(N-2,2)}$, and from (46) and (44) we get

$$\Pi_I^{(N-2,2)} = \frac{1}{\binom{N-4}{k-2}}\sum_{i=0}^{k}\left(\binom{k-i}{2} - \frac{(k-1)^2}{N-2}(k-i) + \frac{k^2(k-1)^2}{2(N-1)(N-2)}\right)A_i^I. \tag{47}$$

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

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

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 $\mathcal{H}_I$, we can define matrices $C_j^{(1,s)}$ to the space $\mathcal{H}^{(1,s)}$. From (46) and (44) we get

$$\Pi_{(N-1,1)}^{(1,s)} = \frac{1}{\binom{N-2}{k-1}}\sum_{i=0}^{k}\left((k-i) - \frac{k^2}{N}\right)A_i^{(1,s)}. \tag{48}$$

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

$$\text{Tr}\left((\mathbb{I}_Q \otimes A_i^I)\cdot A_{i'}^{(1,s)}\right) = \delta_{i,i'}\binom{N}{k}\binom{k}{i}\binom{N-k}{i}\frac{k-i}{k}. \tag{49}$$

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 $\binom{N}{k}$ rows, and each row has $\binom{k}{i}\binom{N-k}{i}$ entries 1. That is, each $z \in D$ has exactly $\binom{k}{i}\binom{N-k}{i}$ $z' \in D$ 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 (47), (48), and (49) together, we get

$$\mathrm{Tr}\big(\Pi^{(N-1,1)Q\otimes(N-2,2)I}_{(N-1,1)}\cdot\Pi^{(1,s)}_{(N-1,1)}\big)=\mathrm{Tr}\big((\mathbb{I}_Q\otimes\Pi^{(N-2,2)}_I)\cdot\Pi^{(1,s)}_{(N-1,1)}\big)=$$

$$=\sum_{i=0}^{k}\frac{\big((k-i)-\frac{k^2}{N}\big)}{\binom{N-2}{k-1}}\frac{\big(\frac{(k-i)(k-i-1)}{2}-\frac{(k-1)^2}{N-2}(k-i)+\frac{k^2(k-1)^2}{2(N-1)(N-2)}\big)}{\binom{N-4}{k-2}}\binom{N}{k}\binom{k}{i}\binom{N-k}{i}\frac{k-i}{k},$$

which, by using the equality

$$\sum_{i=0}^{k}\binom{k}{i}\binom{N-k}{i}\frac{(k-i)!}{(k-i-l)!}=\frac{k!}{(k-l)!}\binom{N-l}{N-k},$$

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

This concludes the proof of Theorem 5 (Hardness of the two values problem).

E   Proofs for Section 4

E.1   Proof of Theorem 6

Proof of Theorem 6.   Algorithm E_1 measures the first half of $|\Sigma\Psi\rangle$. This measurement yields a uniformly random outcome $y\in Y$ and leaves $|\Psi(y)\rangle$ in the second half.

Let $\mathcal{O}_F(y):=I-2|\Psi(y)\rangle\langle\Psi(y)|$. This notation is justified because $\mathcal{O}_F(y)$ is how $\mathcal{O}_F$ operates its the second input when the first input is $|y\rangle$. In particular, given $\mathcal{O}_F$ we can implement the unitary $\mathcal{O}_F(y)$.

The algorithm E_2 is as follows:

```
1  initialize register X with |Ψ(y)⟩ (given as input);
2  for i = 1 to n + 1 do
3      for j = 1, …, ⌈log(π/2√δ_min)⌉ 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 := ∑_{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 4–11). Let $P_y:=\{x\in S_y:P(x)=1\}$ and $\bar{P}_y:=\{x\in S_y:P(x)=0\}$. Let $|yes\rangle:=\sum_{x\in P_y}\sqrt{1/|P_y|}\,|x\rangle$ and $|no\rangle:=\sum_{x\in\bar{P}_y}\sqrt{1/|\bar{P}_y|}\,|x\rangle$. For any $\beta\in\mathbb{R}$, let $|\phi_\beta\rangle:=\sin\beta|yes\rangle+\cos\beta|no\rangle$. We check that $U_P|\phi_\beta\rangle=|\phi_{-\beta}\rangle$. Let $\gamma:=\arcsin\sqrt{|P_y|/|S_y|}$. Then $|\Psi(y)\rangle=\sin\gamma|yes\rangle+$

$\cos\gamma|no\rangle = |\phi_\delta\rangle$. Hence $\mathcal{O}_F(y)|\phi_\beta\rangle = \big(I - 2|\Psi(y)\rangle\langle\Psi(y)|\big)|\phi_\beta\rangle = |\phi_{-\beta+2\gamma}\rangle$ for all β. Thus $\mathcal{O}_F(y)U_P|\phi_\beta\rangle = |\phi_{\beta+2\gamma}\rangle$.

Assume that at line 4, we have $X = |\phi_\beta\rangle$. The innermost loop (lines 4–6) thus yields $X = |\phi_{\beta+2^j\gamma}\rangle$. Since $|yes\rangle \in \operatorname{im} P_X$ and $|no\rangle$ is orthogonal to $\operatorname{im} P_X$, measuring X using P_X (line 8) yields $b = 1$ with probability $(\sin(\beta + 2^j\gamma))^2$. If $b = 1$, X has state $|yes\rangle$, and if $b = 0$, X has state $|no\rangle$. Thus, if $b = 1$, measuring X in the computational basis (line 10) yields and returns $x \in S_y$ with $P(x) = 1$.

Summarizing so far: one iteration of the j-loop (i.e., lines 4–11) returns $x \in S_y$ with probability $(\sin(\beta + 2^j\gamma))^2$ if X has state $|\phi_\beta\rangle$ initially. And if no such x is returned, X is in state $|no\rangle = |\psi_0\rangle$.

In the first execution of the j-loop, X contains $|\Psi(y)\rangle = |\phi_\gamma\rangle$. Thus in all further executions of the j-loop, X contains $|no\rangle = |\phi_0\rangle$ and the probability of returning $x \in S_y$, $P(x) = 1$ in the j-th iteration is $(\sin 2^j\gamma)^2 = 1 - \big(\sin(\pi/2 - 2^j\gamma)\big)^2 \geq 1 - (\pi/2 - 2^j\gamma)^2$.

Thus any but the first iteration of the j-loop (i.e., lines 4–11) fails to return $x \in S_y$ with probability at most:

$$\chi := \min_{1 \leq j \leq \lceil\log(\pi/2\sqrt{\delta_{\min}})\rceil} (\pi/2 - 2^j\gamma)^2.$$

We distinguish two cases:
- Case $\gamma > \frac{\pi}{4}$: Since also $\gamma \leq 1$, we have that $|\pi/2 - 2\gamma| \leq 2 - \pi/2 < \frac{1}{2}$ and thus $\chi \leq (\pi/2 - 2\gamma)^2 \leq (\frac{1}{2})^2 \leq \frac{1}{2}$.
- Case $\gamma \leq \frac{\pi}{4}$: For at least one $1 \leq j \leq \lceil\log(\pi/2\sqrt{\delta_{\min}})\rceil$ we have $2^j\gamma \leq \pi/2$. And for at least one such j we have

$$2^j\gamma \geq 2^{\log\pi/2\sqrt{\delta_{\min}}}\gamma = \frac{\pi\gamma}{2\sqrt{\delta_{\min}}} \geq \frac{\pi\arcsin\sqrt{|P_y|/|S_y|}}{2\sqrt{|P_y|/|S_y|}} \geq \pi/2.$$

Thus the minimum ranges over some $j, j+1$ such that $2^j\gamma \leq \pi/2 \leq 2^{j+1}\gamma$. For any $a \geq 0$, $\min\{|\frac{\pi}{2} - a|, |\frac{\pi}{2} - 2a|\} \leq \frac{\pi}{6}$ if $a \leq \frac{\pi}{2} \leq 2a$. Thus $\chi \leq (\pi/6)^2 \leq \frac{1}{2}$.

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 $\chi^n \leq 2^{-n}$. $\qquad\square$

E.2 Proof of Corollary 8

Proof of Corollary 8. We first show (i). Let $P_A := \Pr[w = w_0 : w \leftarrow A^{\mathcal{O}_{all}}]$.

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:

$$P_A = \Pr[w = w_0 : w_0 \xleftarrow{\$} \{0,1\}^{\ell_{rand}}, (S_{com}) \leftarrow \$, \mathcal{O}_S \leftarrow \$, \mathcal{O}_P \leftarrow \$,$$
$$w \leftarrow A^{\mathcal{O}_E, \mathcal{O}_P, \mathcal{O}_R, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V}].$$

Here we used the following shorthands: $(S_{com}) \leftarrow \$$ means that the sets S_{com} are uniformly random subsets of $\{0,1\}^{\ell_{ch}} \times \{0,1\}^{\ell_{resp}}$ of size k. $\mathcal{O}_S \leftarrow \$$ means that the oracle $\mathcal{O}_S$ is randomly chosen as described in Definition 7 (Oracle distribution). $\mathcal{O}_P \leftarrow \$$ means that the oracle $\mathcal{O}_P$ is randomly chosen as described in Definition 7. Since no random choices are involved in the definitions of $\mathcal{O}_E, \mathcal{O}_R, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V$, we do not write their definitions explicitly here, cf. Definitions 4 and 7.

Removing $\mathcal{O}_P, \mathcal{O}_R$: We now remove access to $\mathcal{O}_P, \mathcal{O}_R$. We then have

$$P_A \leq 2(q_P + q_R + 1)\sqrt{P_1}, \tag{50}$$

$$P_1 := \Pr[w = w_0 : w_0 \xleftarrow{\$} \{0,1\}^{\ell_{rand}}, (S_{com}) \leftarrow \$, \mathcal{O}_S \leftarrow \$, w \leftarrow A_1^{\mathcal{O}_E, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V}]$$

for some A_1 by Lemma 38 (with $\mathcal{O}_1 := (\mathcal{O}_P, \mathcal{O}_R)$, $w := w_0$, $\mathcal{O}_2 := (\mathcal{O}_E, \mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V)$, $\forall w' : f(\cdot, w') := f(w', \cdot, \cdot, \cdot) := w'$). Here the algorithm A_1 makes at most as many oracle queries as A to the remaining oracles. Note that we also removed $\mathcal{O}_P \leftarrow \$$ because $\mathcal{O}_P$ is not used any more.

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

$$P_1 \leq 2q_E\sqrt{P_2} + 2^{-\ell_{rand}}, \tag{51}$$

$$P_2 := \Pr[\text{Collision} : (S_{com}) \leftarrow \$, \mathcal{O}_S \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_2^{\mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V}]$$

for some A_2 by Lemma 39 (with $w := w_0$, $\ell := \ell_{rand}$, $\mathcal{O}_1 := \mathcal{O}_E$, $\mathcal{O}_2 := (\mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_\Psi, \mathcal{O}_V)$, and $X := \{(com, ch, resp, ch', resp') : \text{Collision}\}$). 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 $\mathcal{O}_\Psi$: Fix integers n, m. We determinate the actual values later. By Theorem 3 (Emulating state creation oracles), we have:

$$P_2 \leq P_3 + O\left(\frac{q_\Psi}{\sqrt{n}} + \frac{q_\Psi}{\sqrt{m}}\right), \tag{52}$$

$$P_3 := \Pr[\text{Collision} : (S_{com}) \leftarrow \$, \mathcal{O}_S \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{\mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_V}(|R\rangle)]$$

for some A_3. Here A_3 makes q_S, q_F, q_V queries to $\mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_V$. And $|R\rangle := |\Sigma\Psi\rangle^{\otimes m} \otimes |\alpha_1\rangle \otimes \cdots \otimes |\alpha_n\rangle$ with $|\alpha_j\rangle := (\cos\frac{j\pi}{2n})|\Sigma\Psi\rangle + (\sin\frac{j\pi}{2n})|\bot\rangle$.

Removing $\mathcal{O}_S$: For given choice of $(S_{com})_{com \in \{0,1\}^{\ell_{com}}}$, let $\mathcal{D}_Y$ be the distribution of $\mathcal{O}_S(z)$, i.e., $\mathcal{D}$ picks $com \xleftarrow{\$} \{0,1\}^{\ell_{com}}$ and $(ch, resp) \xleftarrow{\$} S_{com}$ and returns $(com, ch, resp)$.

Fix some integers s (we determine the value of s later). Then, for fixed choice of $(S_{com})_{com}$ ($\mathcal{O}_V, \mathcal{O}_F$ are deterministic given S_{com} anyway), we have by Theorem 40 (Small range distributions [40]) (with $H := \mathcal{O}_S$):

$$\left| \Pr[\text{Collision} : \mathcal{O}_S \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{\mathcal{O}_S, \mathcal{O}_F, \mathcal{O}_V}(|R\rangle)] - \right.$$
$$\left. \Pr[\text{Collision} : G \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{G, \mathcal{O}_F, \mathcal{O}_V}(|R\rangle)] \right| \leq 14q^3/s.$$

Here $G \leftarrow \$$ means that G is chosen as: pick $(com_1, ch_1, resp_1), \ldots, (com_s, ch_s, resp_s) \leftarrow \mathcal{D}_Y$, then for all z, pick $i_z \xleftarrow{\$} \{1, \ldots, s\}$ 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

$$|P_3 - P_4| \leq 14q_S^3/s, \tag{53}$$

$$P_4 := \Pr[\mathsf{Collision} : (S_{com}) \leftarrow \$, G \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_3^{G, \mathcal{O}_F, \mathcal{O}_V}].$$

We construct the adversary A_4: Let $A_4^{\mathcal{O}_F, \mathcal{O}_V}(com_1, ch_1, resp_1, \ldots, com_s, ch_s, resp_s, |R\rangle)$ pick G himself as: for all z, $i_z \xleftarrow{\$} \{1, \ldots, s\}$, $G(z) := (com_{i_z}, ch_{i_z}, resp_{i_z})$. Then A_4 executes $A_3^{G, \mathcal{O}_F, \mathcal{O}_V}(|R\rangle)$. Then

$$P_4 = \Pr[\mathsf{Collision} : (S_{com}) \leftarrow \$, (com_1, ch_1, resp_1), \ldots, (com_s, ch_s, resp_s) \leftarrow \mathcal{D}_Y,$$

$$(com, ch, resp, ch', resp') \leftarrow A_4^{\mathcal{O}_F, \mathcal{O}_V}(com_1, ch_1, resp_1, \ldots, com_s, ch_s, resp_s, |R\rangle)].$$

(Note that the distribution $\mathcal{D}_Y$ depends on the choice of S_y.)

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

By definition of $|\Sigma\Psi\rangle$ (Definition 4), each $(com_i, ch_i, resp_i)$ chosen by A_5 is independently distributed according to $\mathcal{D}_Y$. Thus

$$P_4 = P_5, \tag{54}$$

$$P_5 := \Pr[\mathsf{Collision} : (S_{com}) \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_5^{\mathcal{O}_F, \mathcal{O}_V}(|\Sigma\Psi\rangle^{\otimes s}, |R\rangle)].$$

Converting the $|\alpha_i\rangle$: The adversary A_5 is almost an adversary as in Theorem 5 (Hardness of the two values problem), with one exception: the input to A_5 is a state $|R\rangle = |\Sigma\Psi\rangle^{\otimes m} \otimes |\alpha_1\rangle \otimes \cdots \otimes |\alpha_n\rangle$ with $|\alpha_j\rangle := (\cos \frac{j\pi}{2n})|\Sigma\Psi\rangle + (\sin \frac{j\pi}{2n})|\bot\rangle$. Theorem 5 on the other hand assumes an adversary that takes as input states in the span of $|\Sigma\Psi\rangle$ and $|\Sigma\Phi\rangle := \sum_{com, ch, resp} 2^{-(\ell_{com} + \ell_{ch} + \ell_{resp})/2}|com, ch, resp\rangle$. Let $|\tilde{\alpha}_j\rangle := (\cos \frac{j\pi}{2n})|\Sigma\Psi\rangle + (\sin \frac{j\pi}{2n})|\Sigma\Phi\rangle$. $|\tilde{R}\rangle = |\Sigma\Psi\rangle^{\otimes m} \otimes |\tilde{\alpha}_1\rangle \otimes \cdots \otimes |\tilde{\alpha}_n\rangle$ 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 $A_6^{\mathcal{O}_F, \mathcal{O}_V}(|\Sigma\Psi\rangle^{\otimes s}, |\tilde{R}\rangle)$ be the algorithm that runs $A_5^{\mathcal{O}_F, \mathcal{O}_V}(|\Sigma\Psi\rangle^{\otimes s}, (I^{\otimes m} \otimes U_\alpha^{\otimes n})|\tilde{R}\rangle)$. Then

$$P_5 \leq P_6 + \mathrm{TD}\big((I^{\otimes m} \otimes U_\alpha^{\otimes n})|\tilde{R}\rangle, \; |R\rangle\big), \tag{55}$$

$$P_6 := \Pr[\mathsf{Collision} : (S_{com}) \leftarrow \$, (com, ch, resp, ch', resp') \leftarrow A_6^{\mathcal{O}_F, \mathcal{O}_V}(|\Sigma\Psi\rangle^{\otimes s}, |\tilde{R}\rangle)].$$

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 $c := (\cos \frac{j\pi}{2n})$ and $s := (\sin \frac{j\pi}{2n})$. Then

$$\begin{aligned}
\chi &:= \langle\alpha_j|U_\alpha|\tilde{\alpha}_j\rangle \\
&= (c|\Sigma\Psi\rangle + s|\bot\rangle)^\dagger U_\alpha(c|\Sigma\Psi\rangle + s|\Sigma\Phi\rangle) \\
&= c^2\langle\Sigma\Psi|U_\alpha|\Sigma\Psi\rangle + s^2\langle\bot|\bot\rangle + cs\langle\Sigma\Psi|\bot\rangle + cs\langle\bot|U_\alpha|\Sigma\Psi\rangle \\
&\overset{(*)}{=} c^2|\delta^2| + s^2 + cs \cdot 0 + cs\gamma = c^2(1 - |\gamma^2|) + s^2 + cs\gamma = 1 - c^2|\gamma^2| + cs\gamma.
\end{aligned}$$

In ($*$) we use that $|\bot\rangle, |\Sigma\Phi\rangle, |\Sigma\Phi^\bot\rangle$ are orthogonal. Furthermore,

$$\gamma = \langle \Sigma\Phi | \Sigma\Psi \rangle = \sum_{\substack{com, ch, resp \\ (ch, resp) \in S_{com}}} 2^{-(\ell_{com} + \ell_{ch} + \ell_{resp})/2} \cdot 2^{-\ell_{com}/2}/\sqrt{k}$$

$$= 2^{\ell_{com}} k \cdot 2^{-(\ell_{com} + \ell_{ch} + \ell_{resp})/2} \cdot 2^{-\ell_{com}/2}/\sqrt{k} = 2^{-(\ell_{ch} + \ell_{resp})/2} \sqrt{k} \geq 0.$$

Thus

$$\chi = 1 - c^2 |\gamma^2| + cs\gamma \geq 1 - c^2\gamma^2 \geq 1 - \gamma^2$$

and hence

$$\mathrm{TD}(|\alpha_j\rangle, U_\alpha|\tilde{\alpha}_j\rangle) = \sqrt{1 - \chi^2} \leq \sqrt{1 - (1 - \gamma^2)} \leq \sqrt{2\gamma^2} = 2^{-(\ell_{ch} + \ell_{resp} - 1)/2}\sqrt{k}.$$

With (55), we get

$$P_5 \leq P_6 + \mathrm{TD}\big((I^{\otimes m} \otimes U_\alpha^{\otimes n})|\tilde{R}\rangle, \ |R\rangle\big) = P_6 + \sum_{i=1}^{n} \mathrm{TD}(|\alpha_j\rangle, U_\alpha|\tilde{\alpha}_j\rangle)$$

$$\leq P_6 + n2^{-(\ell_{ch} + \ell_{resp} - 1)/2}\sqrt{k}. \tag{56}$$

Wrapping up: Note that A_6 is an adversary as in Theorem 5 (Hardness of the two values problem). Thus by Theorem 5 (with $h := n + m + s$), we have:

$$P_6 \leq O\left(\frac{(n+m+s)}{2^{\ell_{com}/2}} + \frac{(q_V + q_F)^{1/2} k^{1/4}}{2^{(\ell_{ch} + \ell_{resp})/4}} + \frac{(q_V + q_F)^{1/2}}{k^{1/4}}\right). \tag{57}$$

Let $n, m, s := \lfloor \min\{2^{\ell_{resp}/4}, 2^{\ell_{com}/3}\}\rfloor$. Since ℓ_{resp} and ℓ_{com} are superlogarithmic, n, m, s are superpolynomial. The first summand in (57) is negligible since $n + m + s \leq 3 \cdot 2^{\ell_{com}/3}$. The second summand is negligible because q_V, q_F are polynomially-bounded and $k = 2^{\ell_{ch} + \lfloor \ell_{resp}/3 \rfloor}$ and ℓ_{resp} is superlogarithmic. The third summand is negligible because q_V, q_F are polynomially-bounded and k is superlogarithmic. Thus by (57), P_6 is negligible.

Using $n \leq 2^{\ell_{resp}/4}$ and $k \leq 2^{\ell_{ch} + \ell_{resp}/3}$, we get that the second summand in (56) is upper bounded by $2^{\ell_{resp}/4} \cdot 2^{-\ell_{ch}/2 - \ell_{resp}/2 - 1/2} \cdot 2^{\ell_{ch}/2 + \ell_{resp}/6} = 2^{-1/2 - \ell_{resp}/12}$ which is negligible. Since P_6 is negligible, (56) implies that P_5 is negligible. By (54), P_4 is negligible. Since q_S is polynomially-bounded and s is superpolynomial, $14q_S^3/s$ is negligible. Thus by (53), P_3 is negligible. Since q_Ψ is polynomially-bounded and n, m are superpolynomial, the second summand in (52) is negligible, so P_2 is negligible. Since ℓ_{rand} is superlogarithmic, q_E is polynomially-bounded, and P_2 is negligible, (51) implies that P_1 is negligible. And since q_P, q_R are polynomially-bounded and P_1 is negligible, (50) implies that P_A is negligible. This shows part (i) of the lemma.

We now show part (ii) of the lemma. For an adversary A outputting $(com, ch, resp, ch', resp')$, let B be the adversary that runs $(com, ch, resp, ch', resp') \leftarrow A$,

then invokes $w \leftarrow \mathcal{O}_E(com, ch, resp, ch', resp')$ and returns w. Note that B makes $q_E + 1$ queries to $\mathcal{O}_E$, and the same number of queries to the other oracles as A. By definition of $\mathcal{O}_E$, we have

$$\Pr[(ch, resp) \neq (ch', resp') \wedge (ch, resp), (ch', resp') \in S_{com} :$$
$$(com, ch, resp, ch', resp') \leftarrow A^{\mathcal{O}_{all}}] \leq \Pr[w = w_0 : w \leftarrow B^{\mathcal{O}_{all}}].$$

By (i) the rhs is negligible, thus the lhs is, too. This proves (ii). $\qquad\square$

F Proofs for Section 5

F.1 Proof for Lemma 14

Proof of Lemma 14. **Perfect completeness:** By definition of $\mathcal{O}_S$, we have that $x_i \in S_{y_i}$ for all $(y_i, x_i) := \mathcal{O}_S(z_i)$. Hence $\mathcal{O}_V(y_i, x_i) = 1$ for all i. Thus $\mathrm{COM}_{verify}(c, m, u) = 1$ for $(c, u) \leftarrow \mathrm{COM}(m)$. Hence we have perfect completeness.

Computational strict binding: Consider an adversary $A^{\mathcal{O}_{all}}$ against the computational strict binding property. Let μ be the probability that $A^{\mathcal{O}_{all}}$ outputs (c, m, u, m', u') such that $(m, u) \neq (m', u')$ and $ok = ok' = 1$ with $ok = \mathrm{COM}_{verify}(c, m, u)$ and $ok' = \mathrm{COM}_{verify}(c, m', u')$. We need to show that μ 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 $m_i = b_i \oplus \mathsf{bit}_{p_i}(x_i) = b_i \oplus \mathsf{bit}_{p_i}(x'_i) = m'_i$, in contradiction to $(x_i, m_i) \neq (x'_i, m'_i)$. So $x_i \neq x'_i$. Furthermore, $ok = ok' = 1$ implies that $\mathcal{O}_V(y_i, x_i) = \mathcal{O}_V(y_i, x'_i) = 1$, i.e., $x_i, x'_i \in S_{y_i}$. So $A^{\mathcal{O}_{all}}$ finds $x_i \neq x'_i$ with $x_i, x'_i \in S_{y_i}$ with probability μ. By Corollary 8 (Hardness of two values 2), this implies that μ is negligible.

Computational binding: This is implied by computational strict binding.

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

Let $\mathcal{D}$ be the distribution that returns $(\hat{y}, \hat{x})$ with $\hat{y} \xleftarrow{\$} \{0, 1\}^{\ell_{com}}$, $\hat{x} \xleftarrow{\$} S_{\hat{y}}$. Note that by definition of $\mathcal{O}_S$, $\mathcal{O}_S(z)$ is initialized according to $\mathcal{D}$. By Lemma 32, for fixed choice of the sets S_y, $\mathrm{SD}((\mathcal{O}_S, y, x); (\mathcal{O}_S, \hat{y}, \hat{x})) \leq 2^{(\ell_{com} - \ell_{rand})/2 - 1}\sqrt{k} =: \mu_1$. (With $X := \{0, 1\}^{\ell_{rand}}$, $Y := \{(y, x) : y \in \{0, 1\}^{\ell_{com}}, x \in S_y\}$, and $\mathcal{O} := \mathcal{O}_S$.) Thus for random S_y and random p, $\mathrm{SD}((\mathcal{O}_{all}, p, y, \mathsf{bit}_p(x) \oplus m); (\mathcal{O}_{all}, p, \hat{y}, \mathsf{bit}_p(\hat{x}) \oplus m)) \leq \mu_1$. Let $b^* \xleftarrow{\$} \{0, 1\}$. For fixed $\hat{y}$ and p and random sets S_y and random p, $\mathrm{SD}((S_{\hat{y}}, \hat{y}, \mathsf{bit}_p(\hat{x})); (S_{\hat{y}}, \hat{y}, b^*)) \leq 1/2\sqrt{k} =: \mu_2$ by Lemma 33. Thus for random $\hat{y}$ and p, $\mathrm{SD}((\mathcal{O}_{all}, p, \hat{y}, \mathsf{bit}_p(\hat{x}) \oplus m); (\mathcal{O}_{all}, p, \hat{y}, b^* \oplus m)) \leq \mu_2$. And $(\mathcal{O}_{all}, p, \hat{y}, b^* \oplus m)$ has the same distribution as $(\mathcal{O}_{all}, p, \hat{y}, b^*)$ since $b^* \in \{0, 1\}$ is uniform and independently chosen from $\mathcal{O}_{all}, \hat{y}$. Hence $\mathrm{SD}((\mathcal{O}_{all}, p, y, \mathsf{bit}_p(x) \oplus m); (\mathcal{O}_{all}, p, \hat{y}, b^*)) \leq \mu_1 + \mu_2$. Analogously, $\mathrm{SD}((\mathcal{O}_{all}, p', y', \mathsf{bit}_p(x') \oplus m'); (\mathcal{O}_{all}, p', \hat{y}', b^{*'})) \leq \mu_1 + \mu_2$ with $b^{*'} \xleftarrow{\$} \{0, 1\}$. Since

$(\mathcal{O}_{all}, p, \hat{y}, b^*)$ and $(\mathcal{O}_{all}, p', \hat{y}', b^{*\prime})$ have the same distribution, this implies

$$\mathrm{SD}\big((\mathcal{O}_{all}, p, y, \mathsf{bit}_p(x) \oplus m); (\mathcal{O}_{all}, p', y', \mathsf{bit}_p(x') \oplus m')\big) \leq 2(\mu_1 + \mu_2). \qquad (58)$$

Fix m_1, m_2 with $|m_1| = |m_2|$. Let $z_i \xleftarrow{\$} \{0,1\}^{\ell_{rand}}$, $(y_i, x_i) := \mathcal{O}_S(z_i)$, $p_i \xleftarrow{\$} \{1, \ldots, \ell_{ch} + \ell_{resp}\}$, $b_i := m_i \oplus \mathsf{bit}_{p_i}(x_i)$ and analogously $y_i', x_i', p_i', z_i', b_i'$. By induction over n, and using (58), we get for all $1 \leq n \leq |m_1|$:

$$\mathrm{SD}\big((\mathcal{O}_{all}, (p_i)_{i=1,\ldots,n}, (y_i)_{i=1,\ldots,n}, (\mathsf{bit}_{p_i}(x_i) \oplus m_i)_{i=1,\ldots,n});$$
$$(\mathcal{O}_{all}, (p_i')_{i=1,\ldots,n}, (y_i')_{i=1,\ldots,n}, (\mathsf{bit}_{p_i}(x_i') \oplus m_i')_{i=1,\ldots,n})\big) \leq 2n(\mu_1 + \mu_2).$$

For $n = |m_1|$, this becomes

$$\mathrm{SD}\big((\mathcal{O}_{all}, c), (\mathcal{O}_{all}, c')\big) \leq 2|m_1|(\mu_1 + \mu_2) =: \mu \qquad (\text{with } c \leftarrow \mathrm{COM}(m), \ c' \leftarrow \mathrm{COM}(m')).$$

Since $|m_1|$ is polynomially-bounded, and $\ell_{rand} - \ell_{com} - k$ is superlogarithmic, and k is superpolynomial, μ is negligible. Thus COM is statistically hiding. $\qquad\square$

F.2 Proof of Lemma 15

Proof of Lemma 15. Our adversary is as follows:
- $B_1(|m|)$ invokes E_1 from Theorem 6 (Searching one value) $|m|$ times to get $(y_i, |\Psi(y_i)\rangle)$ for $i = 1, \ldots, |m|$.[10] Let $p_1, \ldots, p_{|m|} \xleftarrow{\$} \{1, \ldots, \ell_{ch} + \ell_{resp}\}$. Let $b_1, \ldots, b_{|m|} \xleftarrow{\$} \{0,1\}$. Output $c := (p_1, \ldots, p_{|m|}, y_1, \ldots, y_{|m|}, b_1, \ldots, b_{|m|})$.
- $B_2(m)$: Let $P_i(x) := 1$ iff $\mathsf{bit}_{p_i}(x) = b_i \oplus m_i$. Then, for each $i = 1, \ldots, |m|$, B_2 invokes $E_2(n, \delta_{\min}, y_i, |\Psi(y_i)\rangle)$ from Theorem 6 with oracle access to $P := P_i$ and with $n := \ell_{com}$ and $\delta_{\min} := 1/3$ to get x_i. Then B_2 outputs $u := (x_1, \ldots, x_n)$.

By Theorem 6, the probability that the i-th invocation of E_2 fails to return x_i with $x_i \in S_y \wedge P_i(x_i) = 1$ is at most:

$$f := 2^{-\ell_{com}} + f_\delta \qquad \text{with} \qquad f_\delta := \Pr\Big[\frac{|\{x \in S_{y_i} : P_i(x) = 1\}|}{|S_{y_i}|} < \delta_{\min}\Big]$$

Let $P_0' := \{x : \mathsf{bit}_{p_i}(x) = 0\}$ and $P_1' := \{x : \mathsf{bit}_{p_i}(x) = 1\}$. Since $S_{y_i} \subseteq X$ is chosen uniformly at random, by Lemma 31 we have for $b = 0, 1$:

$$f_\delta^b := \Pr\big[|S_{y_i} \cap P_b'|/|S_{y_i}| < \delta_{\min}\big] \leq e^{-2k(\frac{1}{2} - \delta_{\min})^2} = e^{-k/18}.$$

Since $P_i = P_0'$ or $P_i = P_1'$, we have $f_\delta \leq f_\delta^0 + f_\delta^1 \leq 2e^{-k/18}$. (Note: we cannot just apply Lemma 31 to P_i because P_i might not be independent of S_{y_i}.)

The probability that B_2 fails to return u with $\mathrm{COM}_{verify}(c, m, u)$ is then $|m|f$. Hence $\varepsilon_{\mathrm{COM}} \geq 1 - |m|f \geq 1 - |m|2^{-\ell_{com}} + |m|2e^{-k/18}$ which is overwhelming since $|m|$ is polynomial and ℓ_{com} and k are superlogarithmic. $\qquad\square$

---

[10] E_1 expects an input $|\Sigma\Psi\rangle$. $|\Sigma\Psi\rangle$ can be computed using the oracle $\mathcal{O}_\Psi$.

G Proofs for Section 6

G.1 Proof of Lemma 18

Proof of Lemma 18. **Completeness**: We need to show that with overwhelming probability, (a) $\mathrm{COM}_{verify}(c_{ch}, resp_{ch}, u_{ch}) = 1$ for $(c_{ch}, u_{ch}) \leftarrow \mathrm{COM}(resp_{ch})$ and (b) $\mathcal{O}_V(com, ch, resp_{ch}) = 1$ for uniform com, ch and $resp_{ch} := \mathcal{O}_P(w, com, ch)$. From the completeness of COM (Lemma 14), we immediately get (a). We prove (b): By definition of $\mathcal{O}_P$ and $\mathcal{O}_V$, (b) holds iff $\exists resp.(ch, resp) \in S_{com}$. We thus need to show that $p_1 := \Pr[\exists resp.(ch, resp) \in S_{com}]$ is overwhelming. S_{com} is a uniformly random subset of size $k = 2^{\ell_{ch} + \lfloor \ell_{resp} \rfloor / 3}$ of $X = \{0,1\}^{\ell_{ch}} \times \{0,1\}^{\ell_{resp}}$. Thus p_1 is lower bounded by the probability p_2 that out of k uniform independent samples from $\{0,1\}^{\ell_{ch}}$, at least one is ch. Thus $p_1 \geq p_2 = 1 - (1 - 2^{-\ell_{ch}})^k = 1 - \left((1 - 1/2^{\ell_{ch}})^{2^{\ell_{ch}}} \right)^{2^{\lfloor \ell_{resp} \rfloor / 3}} \overset{(*)}{\geq} 1 - e^{-2^{\lfloor \ell_{resp} \rfloor / 3}}$ where $(*)$ 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 ℓ_{resp}, and the sigma-protocol is complete.

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

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

Computational strict soundness: We need to show that a polynomial-time A will only with negligible probability output $(com^*, ch, resp^*, resp^{*\prime})$ such that $resp^* \neq resp^{*\prime}$ and $V(s, com^*, ch, resp^*) = V(s, com^*, ch, resp^{*\prime}) = 1$. Assume A outputs such a tuple with non-negligible probability. By definition of V, this implies that $resp^* = (resp, u)$, $resp^{*\prime} = (resp', u')$, and com^* contains c_{ch} such that $\mathrm{COM}_{verify}(c_{ch}, resp, u) = 1$ and $\mathrm{COM}_{verify}(c_{ch}, resp', u') = 1$. Since $resp^* \neq resp^{*\prime}$, this contradicts the computational strict binding property of $\mathrm{COM}, \mathrm{COM}_{verify}$ (Lemma 14). Thus the sigma-protocol has computational strict soundness.

Statistical HVZK: Let S be the simulator that picks $z \overset{\$}{\leftarrow} \{0,1\}^{\ell_{rand}}$, computes $(com, ch, resp) := \mathcal{O}_S(z)$, and $(c_c, u_c) \leftarrow \mathrm{COM}(0^{\ell_{resp}})$ for all $c \in \{0,1\}^{\ell_{ch}} \setminus \{ch\}$, and $(c_{ch}, u_{ch}) \leftarrow \mathrm{COM}(resp)$, and returns $(com^*, ch, resp^*)$ with $com^* := (com, (c_{ch})_{ch \in \{0,1\}^{\ell_{ch}}})$ and $resp^* := (resp_{ch}, u_{ch})$. 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 $resp^*$ for $(resp_{ch}, u_{ch})$.

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

$$= \Pr\big[b = 1 : com \xleftarrow{\$} \{0,1\}^{\ell_{com}}, \ ch \xleftarrow{\$} \{0,1\}^{\ell_{ch}}, \ [\text{for all } c \in \{0,1\}^{\ell_{ch}} : z_c \xleftarrow{\$} \{0,1\}^{\ell_{rand}},$$
$$resp_c := \mathcal{O}_P(w, com, c, z_c), \ (c_c, u_c) \leftarrow \text{COM}(resp_c)], \ b \leftarrow A(com^*, ch, resp^*)\big]$$

$$\overset{\varepsilon_0}{\approx} \Pr\big[b = 1 : com \xleftarrow{\$} \{0,1\}^{\ell_{com}}, \ ch \xleftarrow{\$} \{0,1\}^{\ell_{ch}}, \ [\text{for all } c \in \{0,1\}^{\ell_{ch}} \setminus \{ch\} :$$
$$(c_c, u_c) \leftarrow \text{COM}(0^{\ell_{resp}})], \ z_{ch} \xleftarrow{\$} \{0,1\}^{\ell_{rand}}, resp_{ch} := \mathcal{O}_P(w, com, ch, z_{ch}),$$
$$(c_{ch}, u_{ch}) \leftarrow \text{COM}(resp_{ch}), b \leftarrow A(com^*, ch, resp^*)\big]$$

Here $a \overset{\varepsilon_0}{\approx} b$ means that $|a - b| \leq \varepsilon_0$ where $\varepsilon_0 := 2^{\ell_{ch}} \varepsilon_{\text{COM}}$ and ε_{COM} is the statistical distance between commitments $\text{COM}(resp_c)$ and $\text{COM}(0^{\ell_{resp}})$. We have that ε_{COM} is negligible by Lemma 14 (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:

$$\cdots = \Pr\big[b = 1 : com \xleftarrow{\$} \{0,1\}^{\ell_{com}}, \ ch \xleftarrow{\$} \{0,1\}^{\ell_{ch}}, \ [\text{COM}(0)], \ z_{ch} \xleftarrow{\$} \{0,1\}^{\ell_{rand}},$$
$$resp_{ch} := \mathcal{O}_P(w, com, ch, z_{ch}), \ (c_{ch}, u_{ch}) \leftarrow \text{COM}(resp_{ch}), \ b \leftarrow A(com^*, ch, resp^*)\big]$$

$$\overset{\varepsilon_1}{\approx} \Pr\big[b = 1 : com \xleftarrow{\$} \{0,1\}^{\ell_{com}}, \ ch \xleftarrow{\$} \{0,1\}^{\ell_{ch}}, \ [\text{COM}(0)],$$
$$resp_{ch} \leftarrow \mathcal{D}_{com,ch}, \ (c_{ch}, u_{ch}) \leftarrow \text{COM}(resp_{ch}), \ b \leftarrow A(com^*, ch, resp^*)\big]$$

Here $\mathcal{D}_{com,ch}$ is the uniform distribution on $\{resp : (ch, resp) \in S_{com}\}$. (Or, if that set is empty, $\mathcal{D}_{com,ch}$ assigns probability 1 to $\bot$.) And $a \overset{\varepsilon_1}{\approx} b$ 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 32, with $X := \{0,1\}^{\ell_{rand}}$ and $Y := \{0,1\}^{\ell_{resp}}$ and $\mathcal{D} := \mathcal{D}_{ch,com}$, and using the fact that for all z, $\mathcal{O}_P(w_0, com, ch, z)$ is chosen according to $\mathcal{D}_{ch,com}$. (Note that the adversary A has access to $\mathcal{O}_P$, but that is covered since $\mathcal{O}$ occur on both sides of the statistical distance in Lemma 32.) We continue the computation:

$$\cdots \overset{\varepsilon_2}{\approx} \Pr[b = 1 : (com, ch, resp_{ch}) \xleftarrow{\$} \mathcal{D}', \ [\text{COM}(0)], \ (c_{ch}, u_{ch}) \leftarrow \text{COM}(resp_{ch}),$$
$$b \leftarrow A(com^*, ch, resp^*)]$$

Here $\mathcal{D}'$ is the distribution resulting from choosing $com \xleftarrow{\$} \{0,1\}^{\ell_{com}}$, $(ch, resp) \xleftarrow{\$} S_{com}$. By Lemma 34, $\varepsilon_2 \leq \frac{2k^2}{2^{\ell_{ch}+\ell_{resp}}} + \frac{2^{\ell_{ch}/2}}{2\sqrt{k}}$. We continue

$$\ldots \overset{\varepsilon_3}{\approx} \Pr[b = 1 : z \xleftarrow{\$} \{0,1\}^{\ell_{rand}}, (com, ch, resp) := \mathcal{O}_S(z),$$
$$[\text{COM}(0)], \ (c_{ch}, u_{ch}) \leftarrow \text{COM}(resp_{ch}), \ b \leftarrow A(com^*, ch, resp^*)]$$

Here $\varepsilon_3 = \sqrt{(2^{\ell_{com}} \cdot k)/2^{\ell_{rand}}}$. This follows from Lemma 32 with $\mathcal{D} := \mathcal{D}'$ and $X := \{0,1\}^{\ell_{rand}}$ and $Y := \{(com, ch, resp) : (ch, resp) \in S_{com}\}$. (Note that $|Y| = 2^{\ell_{com}} \cdot k$.) We continue

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

Thus the difference of probabilities from the definition of statistical HVZK is bounded by $\varepsilon := \varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$. And ε is negligible since $\varepsilon_{\mathrm{COM}}$ is negligible, and $k = 2^{\ell_{ch} + \lfloor \ell_{resp} \rfloor / 3}$, and ℓ_{ch} is logarithmic, and ℓ_{resp}, ℓ_{com} are superlogarithmic, and $\ell_{rand} = \ell_{com} + \ell_{resp}$. $\square$

G.2   Proof of Lemma 19

Proof of Lemma 19.   According to Definition 2 (specialized to the case of the sigma-protocol from Definition 17) we need to construct a polynomial-time quantum adversary A_1, A_2, A_3 such that:

- Adversary success:

$$P_A := \Pr[ok = 1 : s \leftarrow A_1, com^* \leftarrow A_2, ch \overset{\$}{\leftarrow} \{0,1\}^{\ell_{ch}},$$
$$resp^* \leftarrow A_3(ch), ok = V(s, com^*, ch, resp^*)]$$
$$= \Pr[ok_v = 1 \wedge ok_c = 1 \wedge s = s_0 : s \leftarrow A_1, \big(com, (c_{ch})_{ch \in \{0,1\}^{\ell_{ch}}}\big) \leftarrow A_2,$$
$$ch \overset{\$}{\leftarrow} \{0,1\}^{\ell_{ch}}, (resp, u) \leftarrow A_3(ch), ok_v := \mathcal{O}_V(com, ch, resp),$$
$$ok_c = \mathrm{COM}_{verify}(c_{ch}, resp, u)] \tag{59}$$

  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 15 (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_1 from Theorem 6 (Searching one value) to get $(com, |\Psi(com)\rangle)$.[11] Then A_2 invokes $c_c \leftarrow B_1(\ell_{resp})$ for all $c \in \{0,1\}^{\ell_{ch}}$. A_2 outputs $com^* := (com, (c_{ch})_{ch \in \{0,1\}^{\ell_{ch}}})$.
- Let $P_{ch}(ch', resp') := 1$ iff $ch' = ch$. $A_3(ch)$ invokes $E_2(n, \delta_{\min}, com, |\Psi(com)\rangle)$ from Theorem 6 with oracle access to $P := P_{ch}$ and with $n := \ell_{com}$ and $\delta_{\min} := 2^{-\ell_{ch}-1}$ to get $resp$. Then A_3 invokes $u \leftarrow B_1(resp)$ to get opening information for c_{ch}. A_3 outputs $resp^* := (resp, u)$.

Adversary success: By Lemma 15, $\mathrm{COM}_{verify}(c_{ch}, resp, u) = 1$ with overwhelming probability. Thus $ok_c = 1$ with overwhelming probability in (59).

By Theorem 6, the probability that E_2 fails to return $(ch', resp)$ with $(ch', resp) \in S_{com} \wedge P_{ch}(ch', resp) = 1$ is at most:

$$f := 2^{-\ell_{com}} + f_\delta \qquad \text{with} \qquad f_\delta := \Pr\Big[\frac{|\{(ch', resp) \in S_{com} : P_{ch}(ch', resp) = 1\}|}{|S_{com}|} < \delta_{\min}\Big]$$

Let $P' := \{x : P_{ch}(x) = 1\}$ and $X := \{0,1\}^{\ell_{ch}} \times \{0,1\}^{\ell_{com}}$. Then $|P'|/|X| = 2^{-\ell_{ch}}$. Since $S_{com} \subseteq X$ is chosen uniformly at random with $|S_{com}| = k$, by Lemma 31 we have:

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

---

[11]Using $\mathcal{O}_\Psi$ to get the input $|\Sigma\Psi\rangle$ for E_1.

Thus $f \leq 2^{-\ell_{com}} + e^{-k2^{-2\ell_{ch}-1}}$ is negligible since ℓ_{com} is superpolynomial, ℓ_{ch} logarithmic, and k superpolynomial. Thus with overwhelming probability E_2 returns $(ch', resp) \in S_{com}$ with $P_{ch}(ch', resp) = 1$. $P_{ch}(ch', resp) = 1$ implies $ch' = ch$. Hence $(ch, resp) \in S_{com}$, thus $\mathcal{O}_V(com, ch, resp) = 1$, thus $ok_v = 1$ 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

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

is negligible by Corollary 8 (Hardness of two values 2). This shows extractor failure. $\qquad\square$

G.3   Proof of Lemma 22

Proof of Lemma 22.   **Completeness** and **statistical HVZK** and **commitment entropy** hold trivially, because they only have to hold for $(s, w) \in R' = \varnothing$. **Computational strict soundness** is shown exactly as in the proof of Lemma 18 (Security of the sigma-protocol). (The definition of computational strict soundness is independent of the relation R'.)

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

$$\begin{aligned}
P_S := \Pr[&(s, w) \notin R' \wedge ch \neq ch' \wedge ok = ok' = 1 : (s, com^*, ch, resp^*, ch', resp^{*\prime}) \leftarrow A^{\mathcal{O}_{all}}, \\
&ok \leftarrow V(s, com^*, ch, resp^*), ok' \leftarrow V(s, com^*, ch', resp^{*\prime}), \\
&w \leftarrow E_\Sigma(s, com^*, ch, resp^*, ch', resp^{*\prime})] \\
\leq \Pr[&ch \neq ch' \wedge (ch, resp), (ch', resp') \in S_{com} : (com^*, ch, resp^*, ch', resp^{*\prime}) \leftarrow A^{\mathcal{O}_{all}}, \\
&(com, \dots) := com^*, (resp, \dots) := resp^*, (resp', \dots) := resp^{*\prime}]
\end{aligned}$$

The right hand side is negligible by Corollary 8 (Hardness of two values 2). Hence P_S is negligible. This shows that the sigma-protocol from Definition 21 has computational special soundness. $\qquad\square$

G.4   Proof of Lemma 23

Proof of Lemma 23.   By Definition 2 (specialized to the sigma-protocol from Definition 21), we need to construct a polynomial-time adversary A_1, A_2, A_3 such that:

$$\begin{aligned}
P_A := \Pr[&ok = 1 \ \wedge \ s \notin L_{R'} : s \leftarrow A_1, com^* \leftarrow A_2, ch \xleftarrow{\$} \{0,1\}^{\ell_{ch}}, resp^* \leftarrow A_3(ch), \\
&ok := V(com^*, ch, resp^*)] \text{ is overwhelming}
\end{aligned}$$

We use the same adversary (A_1, A_2, A_3) as in the proof of Lemma 19. Then P_A here is the same as P_A in the proof of Lemma 19. (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 19 we showed that P_A is overwhelming. $\qquad\square$

H   Proofs for Section 7

H.1   Proof of Theorem 25

Lemma 48 (Attack on Fiat-Shamir) *There exists a total knowledge break (Definition 2) against the Fiat-Shamir construction based on the sigma-protocol from Definition 17. (For any r.)*

Proof. According to Definition 2 (specialized to the case of the Fiat-Shamir construction based on the sigma-protocol from Definition 17) we need to construct a polynomial-time quantum adversary $\hat{A}_1, \hat{A}_2$ such that:

- Adversary success:

$$\hat{P}_A := \Pr[\forall i.ok_i = 1 : s \leftarrow \hat{A}_1^{H,\mathcal{O}_{all}}, \ \big((com_i^*)_i, (resp_i^*)_i\big) \leftarrow \hat{A}_2^{H,\mathcal{O}_{all}},$$
$$ch_1 \| \ldots \| ch_r := H(s, (com_i^*)_i), ok_i := V(com_i^*, ch_i, resp_i^*)]$$

  is overwhelming. Here V is the verifier of the sigma-protocol (Definition 17).
- Extractor failure: For any polynomial-time quantum E (with access to the final state of $\hat{A}_1$), $\Pr[s = s_0, w = w_0 : s \leftarrow \hat{A}_1^{H,\mathcal{O}_{all}}, w \leftarrow E^{H,\mathcal{O}_{all}}(s)]$ is negligible.

Let A_1, A_2, A_3 be the adversary from the proof of Lemma 19 (Attack on the sigma-protocol). Our adversary is then as follows:

- $\hat{A}_1$ outputs s_0. (Identical to A_1.)
- $\hat{A}_2$ invokes the adversary A_2 r times to get $com_1^*, \ldots, com_r^*$. Then $\hat{A}_2$ computes $ch_1 \| \ldots \| ch_r := H(s, (com_i^*)_i)$. Then $\hat{A}_2$ invokes A_3 r times to get $resp_1^* \leftarrow A_3(ch_1), \ldots, resp_r^* \leftarrow A_3(ch_r)$. Then $\hat{A}_2$ outputs $((com_i^*)_i, (resp_i^*)_i)$.

Adversary success: We have

$$1 - \hat{P}_A = \Pr[\exists i.ok_i = 0 : s \leftarrow A_1^{\mathcal{O}_{all}}, \; \forall i.com_i^* \leftarrow A_2^{\mathcal{O}_{all}},$$
$$ch_1 \| \ldots \| ch_r := H(s, (com_i^*)_i), \; \forall i.resp_i^* \leftarrow A_3^{\mathcal{O}_{all}}(ch_i),$$
$$\forall i.ok_i \leftarrow V(com_i^*, ch_i, resp_i^*)]$$
$$\overset{(*)}{=} \Pr[\exists i.ok_i = 0 : s \leftarrow A_1^{\mathcal{O}_{all}}, \; \forall i.com_i^* \leftarrow A_2^{\mathcal{O}_{all}},$$
$$\forall i.ch_i \overset{\$}{\leftarrow} \{0,1\}^{\ell_{ch}}, \; \forall i.resp_i^* \leftarrow A_3^{\mathcal{O}_{all}}(ch_i),$$
$$\forall i.ok_i \leftarrow V(com_i^*, ch_i, resp_i^*)]$$
$$\overset{(**)}{\leq} \sum_{i=1}^{r} \Pr[ok_i = 0 : s \leftarrow A_1^{\mathcal{O}_{all}}, \; com_i^* \leftarrow A_2^{\mathcal{O}_{all}},$$
$$ch_i \overset{\$}{\leftarrow} \{0,1\}^{\ell_{ch}}, \; resp_i^* \leftarrow A_3^{\mathcal{O}_{all}}(ch_i),$$
$$ok_i \leftarrow V(com_i^*, ch_i, resp_i^*)]$$
$$\overset{(***)}{=} \sum_{i=1}^{r} (1 - P_A) = r(1 - P_A).$$

Here $(*)$ uses the fact that H is only queried once (classically), and thus $H(s, (com_i^*)_i)$ is uniformly random. And $(**)$ is a union bound. And $(***)$ is by definition of P_A in the proof of Lemma 19. There is was also shown that P_A is overwhelming. Thus $1 - \hat{P}_A \leq r(1 - P_A)$ is negligible and hence $\hat{P}_A$ overwhelming. Thus we have adversary success.

Extractor failure: Extractor failure was already shown in the proof of Lemma 19. (A_1 here is defined exactly as $\hat{A}_1$ in the proof of Lemma 19, and the definition of extractor failure depends only on $\hat{A}_1$, not on $\hat{A}_2$ 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. $\qquad\square$

Now Theorem 25 follows from Lemma 18 (Security of the sigma-protocol) and Lemma 48. (The fact that the Fiat-Shamir protocol is a classical argument of knowledge is shown in [18].[12])

H.2 Proof of Theorem 26

Lemma 49 (Attack on Fiat-Shamir, computational) *Then there exists a total break (Definition 2) against the Fiat-Shamir construction based on the sigma-protocol from Definition 21. (For any r.)*

---

[12]Actually, [18] 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, [18] assumes that ℓ_{ch} is superlogarithmic, and considers the case $r = 1$. But [18] 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 [18].

Proof. By Definition 2 (specialized to the case of the Fiat-Shamir construction based on the sigma-protocol from Definition 21), we need to construct a polynomial-time adversary A_1, A_2 such that:

$$\hat{P}_A := \Pr[\forall i.ok_i = 1 \land s \notin L_{R'} : s \leftarrow \hat{A}_1^{H,\mathcal{O}_{all}}, \ ((com_i^*)_i, (resp_i^*)_i) \leftarrow \hat{A}_2^{H,\mathcal{O}_{all}},$$
$$ch_1 \| \dots \| ch_r := H(s, (com_i^*)_i), ok_i := V(com_i^*, ch_i, resp_i^*)] \text{ is overwhelming}$$

Here V is the verifier of the sigma-protocol (Definition 21).

We use the same adversary $(\hat{A}_1, \hat{A}_2)$ as in the proof of Lemma 48 (Attack on Fiat-Shamir). Then $\hat{P}_A$ here is the same as $\hat{P}_A$ in the proof of Lemma 48. (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 48 we showed that $\hat{P}_A$ is overwhelming. $\qquad\square$

Now Theorem 26 follows from Lemmas 22 and 49. (The fact that the Fiat-Shamir protocol is a classical argument of knowledge is shown in [18].[13])

I Proofs for Section 8

I.1 Proof of Theorem 28

Lemma 50 (Attack on Fischlin's construction) *There exists a total knowledge break (Definition 2) against the Fischlin construction based on the sigma-protocol from Definition 17 (Sigma-protocol).*

Proof. According to Definition 2 (Total breaks) (specialized to the case of Fischlin's construction based on the sigma-protocol from Definition 17) we need to construct a polynomial-time quantum adversary A_1, A_2 such that:

- Adversary success:

$$P_A := \Pr\Big[\forall i.ok_i = 1 \land \sigma \leq \mathbf{S} \land s = s_0 : s \leftarrow A_1^{H,\mathcal{O}_{all}},$$
$$(com_i^*, ch_i, resp_i^*)_{i=1\dots r} \leftarrow A_2^{H,\mathcal{O}_{all}}, \ ok_i := V(com_i^*, ch_i, resp_i^*),$$
$$\sigma := \sum_{i=1}^{r} H(x, (com_i^*)_i, i, ch_i, resp_i^*)\Big] \text{ is overwhelming.} \tag{60}$$

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

---

[13]Actually, [18] 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, [18] assumes that ℓ_{ch} is superlogarithmic, and considers the case $r = 1$. But [18] 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 [18].

Adversary success: At the first glance, it may seem that it is immediate how to construct an adversary that has adversary success: Using Theorem 6 (Searching one value), we can for each i search $(ch_i, resp_i) \in S_{com_i}$ such that $H(x, (com_i^*)_i, i, ch_i, resp_i^*) = 0$. However, there is a problem: com_i^* contains commitments c_{ch}^i to all responses. Thus, after finding $ch_i, resp_i$, we need to open $c_{ch_i}^i$ as $resp_i$. This could be done with the adversary against COM from Lemma 15 (Attack on COM). But the problem is, the corresponding openings have to be contained in $resp_i^*$. So we need to know these openings already when searching for $ch_i, resp_i$. But at that point we do not know yet to what value the commitments $c_{ch_i}^i$ should be opened! To avoid this problem, we use a special fixpoint property of the commitment scheme COM that allows us to commit in a way such that we can use the $(ch_i, resp_i)$ themselves as openings for the commitments.

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

$$\mathrm{COM}_{verify}(c, resp, u) = 1 \text{ for } c := \mathsf{COM}^*(com) \text{ and } u := \mathsf{COMopen}^*(ch, resp). \quad (61)$$

These functions are defined as follows: $\mathsf{COM}^*(com) = (p_1, \ldots, p_{\ell_{resp}}, y_1, \ldots, y_{\ell_{resp}}, b_1, \ldots, b_{\ell_{resp}})$ with $p_i := \ell_{ch} + i$, $y_i := com$, $b_i := 0$. And $\mathsf{COMopen}^*(ch, resp) := (x_1, \ldots, x_{\ell_{resp}})$ with $x_i := (ch, resp)$ for all i. It is easy to verify from the definition of COM_{verify} (Definition 13) that (61) holds if $(ch, resp) \in S_{com}$.

Our adversary is as follows:

- A_1 outputs s_0.
- A_2 invokes E_1 from Theorem 6 (Searching one value) r times to get $(com_i, |\Psi(com_i)\rangle)$ for $i = 1, \ldots, r$. A_2 sets $c_{ch}^i := \mathsf{COM}^*(com_i)$ for all i and all $ch \in \{0,1\}^{\ell_{ch}}$. And $com_i^* := (com_i, (c_{ch}^i)_{ch})$.
  Let $P_i(ch', resp') := 1$ iff $H(s, (com_i^*)_i, i, ch', (resp', \mathsf{COMopen}^*(ch', resp'))) = 0$. Then, for each $i = 1, \ldots, r$, A_2 invokes $E_2(n, \delta_{\min}, com_i, |\Psi(com_i)\rangle)$ from Theorem 6 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 $resp_i^* := (resp_i, \mathsf{COMopen}^*(ch_i, resp_i))$. Then A_2 outputs $\pi := (com_i^*, ch_i, resp_i^*)_{i=1,\ldots,r}$.

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

$$\Pr[\mathsf{Succ}_i] = \Pr[(ch, resp) \in S_{com_i} \wedge P(ch, resp) = 1 : \forall j. (com_j, |\Psi(com_j)\rangle) \leftarrow E_1,$$
$$\forall j. com_j^* := (com_j, (\mathsf{COM}^*(com_j))_{ch}), H \xleftarrow{\$} (\{0,1\}^* \to \{0,1\}^b),$$
$$\forall ch' resp'. P(ch', resp') := 1 \text{ iff } H(s, (com_j^*)_j, i, ch', (resp', \mathsf{COMopen}^*(ch', resp'))) = 0,$$
$$(ch, resp) \leftarrow E_2(n, \delta_{\min}, com_i, |\Psi(com_i)\rangle)]. \quad (62)$$

Hence by Theorem 6 (Searching one value),

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

Here P and com are chosen as in the rhs of (62).

In the rhs of (62), 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 $\dot{P}(ch, resp) = 1$ or $P(ch, resp) = 0$, where $\Pr[P(ch, resp) = 1] = 2^{-b}$. Thus

$$p_\delta = \Pr\Big[\sum\nolimits_{i \in S} \tfrac{X_i}{|S|} \geq 1 - \delta_{\min}\Big] = \Pr\Big[\sum\nolimits_{i \in S} \tfrac{X_i}{|S|} - (1 - 2^{-b}) \geq 1 - \delta_{\min} - (1 - 2^{-b})\Big]$$

$$\overset{(*)}{\leq} e^{-2|S|(1-\delta_{\min}-(1-2^{-b}))^2} = e^{-2k(2^{-b}-\delta_{\min})^2} = e^{-k(2^{-2b-1})}$$

where $X_{ch,resp} := 1 - P(ch, resp)$ and $S := S_{com_i}$. And $(*)$ follows from Hoeffding's inequality [24].

We thus have

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

Since r is polynomially bounded and b is logarithmic and ℓ_{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 $\forall i.\mathsf{Succ}_i$ implies $\forall i.ok_i = 1 \wedge \sigma \leq \mathbf{S} \wedge s = s_0$. First, note that $s = s_0$ always holds by definition of A_1. Furthermore, $\forall i.\mathsf{Succ}_i$ implies (by definition of P_i) that

$$\sigma = \sum\nolimits_i H(s, (com_i^*)_i, i, ch_i, resp_i^*)$$
$$= \sum\nolimits_i H\Big(s, (com_i^*)_i, i, ch_i, \big(resp_i, \mathsf{COMopen}^*(ch_i, resp_i)\big)\Big) = \sum\nolimits_i 0 \leq \mathbf{S}.$$

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

$$\mathrm{COM}_{verify}(c^i_{ch_i}, resp_i, \mathsf{COMopen}^*(ch_i, resp_i))$$
$$= \mathrm{COM}_{verify}(\mathsf{COM}^*(com_i), resp_i, \mathsf{COMopen}^*(ch_i, resp_i)) \overset{(61)}{=} 1.$$

And $\mathcal{O}_V(com_i, ch_i, resp_i) = 1$. Thus $ok_i = V(com_i^*, ch_i, resp_i^*) = 1$. Summarizing, $\forall i.\mathsf{Succ}_i$ implies $\forall i.ok_i = 1 \wedge \sigma \leq \mathbf{S} \wedge s = s_0$ 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 19. (A_1 here is defined exactly as in the proof of Lemma 19, 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. $\qquad\square$

Now Theorem 28 follows from Lemma 18 (Security of the sigma-protocol) and Lemma 50. (The fact that Fischlin's construction is a classical argument of knowledge is shown in [20].[14])

---

[14]Actually, [20] 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.

I.2 Proofs for Theorem 29

Lemma 51 (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 21 (Sigma-protocol, computational).*

Proof. By Definition 2 (specialized to the case of Fischlin's construction based on the sigma-protocol from Definition 21), we need to construct a polynomial-time adversary A_1, A_2 such that:

$$P_A := \Pr\Big[\forall i.ok_i = 1 \land \sigma \leq \mathbf{S} \land s = s_0 \land s \notin L_{R'} : s \leftarrow A_1^{H, \mathcal{O}_{all}},$$
$$(com_i^*, ch_i, resp_i^*)_{i=1\ldots r} \leftarrow A_2^{H, \mathcal{O}_{all}}, \ ok_i := V(com_i^*, ch_i, resp_i^*),$$
$$\sigma := \sum_{i=1}^{r} H(x, (com_i^*)_i, i, ch_i, resp_i^*)\Big] \text{ is overwhelming.}$$

Here V is the verifier of the sigma-protocol (Definition 21).

We use the same adversary (A_1, A_2) as in the proof of Lemma 50 (Attack on Fischlin's construction). Then P_A here is the same as P_A in the proof of Lemma 19. (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 50 we showed that P_A is overwhelming. $\square$

Now Theorem 29 follows from Lemma 22 (Security of the sigma-protocol, computational) and Lemma 51. (The fact that Fischlin's construction is a classical argument of knowledge is shown in [20].[15])

---

[15]Actually, [20] 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.