Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 31 (2018)

On the Communication Complexity of Key-Agreement Protocols

Iftach Haitner[*][†] Noam Mazor[†][‡] Rotem Oshman [§] Omer Reingold [¶]
Amir Yehudayoff [‖]

November 12, 2018

Abstract

Key-agreement protocols whose security is proven in the *random oracle model* are an important alternative to protocols based on public-key cryptography. In the random oracle model, the parties and the eavesdropper have access to a shared random function (an "oracle"), but the parties are limited in the number of queries they can make to the oracle. The random oracle serves as an abstraction for black-box access to a symmetric cryptographic primitive, such as a collision resistant hash. Unfortunately, as shown by Impagliazzo and Rudich [STOC '89] and Barak and Mahmoody [Crypto '09], such protocols can only guarantee limited secrecy: the key of any ℓ-query protocol can be revealed by an $O(\ell^2)$-query adversary. This quadratic gap between the query complexity of the honest parties and the eavesdropper matches the gap obtained by the *Merkle's Puzzles* protocol of Merkle [CACM '78].

In this work we tackle a new aspect of key-agreement protocols in the random oracle model: their *communication complexity*. In Merkle's Puzzles, to obtain secrecy against an eavesdropper that makes roughly ℓ^2 queries, the honest parties need to exchange $\Omega(\ell)$ bits. We show that for protocols with certain natural properties, ones that Merkle's Puzzle has, such high communication is unavoidable. Specifically, this is the case if the honest parties' queries are uniformly random, or alternatively if the protocol uses non-adaptive queries and has only two rounds. Our proof for the first setting uses a novel reduction from the set-disjointness problem in two-party communication complexity. For the second setting we prove the lower bound directly, using information-theoretic arguments.

Understanding the communication complexity of protocols whose security is proven in the random-oracle model is an important question in the study of practical protocols. Our results and proof techniques are a first step in this direction.

Keywords: key agreement; random oracle; communication complexity; Merkle's puzzles

[*] The Blavatnik school of computer science,, Tel Aviv University. E-mail: `iftachh@cs.tau.ac.il`. Member of the Check Point Institute for Information Security.

[†] Supported by ERC starting grant 638121.

[‡] The Blavatnik school of computer science, Tel Aviv University. E-mail: `noammaz@gmail.com`.

[§] The Blavatnik school of computer science, Tel Aviv University. E-mail: `rotem.oshman@gmail.com`. Supported by the Israeli Centers of Research Excellence program 4/11 and BSF grant 2014256.

[¶] Computer Science Department, Stanford University. E-mail: `reingold@stanford.edu`. Supported by NSF grant CCF-1749750.

[‖] Department of Mathematics, Technion-Israel Institute of Technology. E-mail: `amir.yehudayoff@gmail.com`. Supported by ISF grant 1162/15.

ISSN 1433-8092

1 Introduction

In a key-agreement protocol [5], two parties communicating over an insecure channel want to securely agree on a shared secret key, such that an eavesdropper observing their communication cannot find the key. For example, given a hash function $h : [n] \to [N]$ that is hard to invert, the players can execute the following protocol, called *Merkle's puzzles* [13]: we fix an arbitrary parameter $\ell \approx \sqrt{n}$, and the parties select uniformly random subsets $A = \{a_1, \ldots, a_\ell\}, B = \{b_1, \ldots, b_\ell\} \subseteq [n]$ (respectively) of size ℓ. We choose ℓ, n such that with constant probability there is a unique intersection, $|A \cap B| = 1$. The first party evaluates h on every element $a \in A$, and sends $h(a_1), \ldots, h(a_\ell)$ to the second party, which then looks for a unique element $b \in B$ such that $h(b) = h(a_i)$ for some $i \in [\ell]$. If found, the second party sends the index i to the first party and outputs b as the secret key; the second party outputs a_i as the secret key. Because h is a "good" hash function and $h(b) = h(a_i)$, it is likely that $b = a_i$, so the players output the same key. Moreover, since h is hard to invert, an eavesdropper that tries to find the secret key after seeing $h(a_1), \ldots, h(a_\ell), i$ must essentially compute h on the entire universe in order to invert h and find a_i. Thus, we have a quadratic gap between the work performed by the eavesdropper, which must compute $\Omega(\ell^2)$ hashes, and the work performed by the parties, which compute ℓ hashes each.

Ideally we would strive for an *exponential* gap between the work required to break the security of the protocol and the work of the honest parties. There are numerous candidate constructions of such key-agreement schemes, e.g., [18, 15, 1, 12], based on assumptions implying that *public-key encryption schemes* exist. A fundamental open question is whether we can design key-agreement protocols based on the security of symmetric primitives, e.g., *private-key* encryption (e.g., collision resistant hash); the security of such primitives is believed to be more robust than public-key encryption. A very important step in this direction was made by Barak and Mahmoody [2] (following Impagliazzo and Rudich [10]): they showed that as long as the symmetric primitive is used as a black box, the quadratic gap achieved by Merkle's puzzles is the best possible.

The notion of "black box" is formalized by the *random oracle model*: instead of a concrete hash function h, we assume that the parties have access to a *random oracle* $\mathrm{F} : [n] \to [n]$, a perfectly random function. The random oracle is "the best hash function possible" (w.h.p.), so lower bounds proven in the random oracle model hold for any instantiation where the oracle is replaced by a one-way function. Thus, the lower bound of Barak and Mahmoody [2] rules out any black-box key-agreement scheme from one-way functions that achieves a better than quadratic gap between the eavesdropper's work and the honest parties.

While a quadratic gap between the ℓ-query honest parties and the ℓ^2-query eavesdropper might not seem like much, and ideally we would wish for an exponential gap, on modern architecture it can yield a good enough advantage, assuming that security is preserved when the random oracle is replaced with a fixed hash function. For example, a consumer-level CPU (Intel Core i5-6600) can compute 5 million SHA-256 hashes per second, and specialized hardware for SHA-256 computation (for example, AntMiner S9) can compute 14×10^{12} hashes per second [3]. It follows that if the honest parties spend one second of computation on standard CPU, an attacker with specialized hardware can violate the security of Merkle's puzzles in less than a second. However, if the parties spend one second on specialized hardware, an attacker with specialized hardware has to spend more than $200,000$ years to break the scheme.

So, are Merkle's puzzles a practical and realistic key-agreement scheme? The answer is probably not: even setting aside the question of replacing the random oracle by a concrete hash function, in Merkle's puzzles, the honest parties *send each other* $\widetilde{\Omega}(\ell)$ *bits* to obtain security against an eavesdropper that makes roughly ℓ^2 queries. In our example above, if we instantiate Merkle's puzzles using SHA-256 for one second on specialized hardware, the first party would need to send more than 100 terabytes to the second party. A fundamental question is whether this high communication burden is inherent to secure key-agreement, and more generally, what is the communication cost of cryptographic protocols in the random oracle model and other oracle models. In this paper we initiate the study of the communication complexity of cryptographic protocols in the random-oracle model.

1.1 Our Results

We show that for random-oracle protocols with certain natural properties, the high communication incurred by Merkle's puzzles is unavoidable: in order to achieve security against an adversary that can ask $\Theta(\ell^2)$ queries, the two parties must exchange $\Omega(\ell)$ bits of communication. Specifically, we show that the bound above holds for protocols where the parties' queries are a uniformly random set, and also for two-round protocols that make non-adaptive (but arbitrary) queries.[1]

To simplify the statements of our results, we focus here on key-agreement protocols whose *agreement* parameter, the probability that the players output the same key, is larger by some constant than their *secrecy* parameter, the probability that an eavesdropper can find the key.

Uniform-query protocols. We say that a random-oracle protocol makes *uniform queries* if each party's oracle queries are a uniformly random set. We give the following lower bound on the communication complexity of such protocols.

Theorem 1.1 (lower bound on uniform-queries protocols, informal)**.** *Any ℓ-uniform-query key-agreement protocol achieving non-trivial secrecy against $o(\ell^2)$-query adversaries has communication complexity $\Omega(\ell)$.*

Theorem 1.1 is proved by a reduction from *set-disjointness*, a problem in communication complexity that is known to require high communication.

Two-round non-adaptive protocols. An oracle protocol is said to make *non-adaptive* queries if the distribution of queries made by the players is fixed in advance, i.e., it is determined before the parties communicate with each other and does not depend on the oracle's answers. We give the following lower bound on the communication complexity of such protocols.

Theorem 1.2 (lower bound on two-message non-adaptive protocols, informal)**.** *Any two-message ℓ-query non-adaptive key-agreement protocol of non-trivial secrecy against q-query adversaries has communication complexity $\Omega(q/\ell)$.*

Once again this lower bound is nearly-tight with Merkle's puzzles, where $q = \Theta(\ell^2)$, and the communication cost is $\tilde{\Theta}(\ell)$.[2]

Following Barak and Mahmoody [2] and Impagliazzo and Rudich [10], we prove this lower bound by presenting an eavesdropper that makes q queries and prevents the parties from exploiting the advantage they gain by their joint random oracle calls.

[1] These are both properties of Merkle's puzzles.

[2] This theorem is also nearly-tight for any q, with a version of Merkle's puzzle, in which Alice is sending $\Theta(q/l)$ answers, from a universe of size $\Theta(q)$.

In [2], the communication cost of the protocol is not taken into account: their eavesdropper makes $O(\ell^2)$ queries and has high probability of finding *all* intersection queries (i.e., all queries that were asked by both players). In our case, if the protocol has communication cost C, then to prove Theorem 1.2, our eavesdropper must make only $O(C \cdot \ell)$ queries (to show the trade-off that $C = \Omega(q/\ell)$). If $C \ll \ell$, our eavesdropper makes much fewer queries than the eavesdropper in [2, 10], and in particular it cannot discover all the intersection queries. Instead, our eavesdropper asks only queries that the players *were able to learn* are in their intersection. If a query is in the intersection, but the players have not communicated this fact to each other, then the eavesdropper will not necessarily ask this query (unlike [2, 10]). Finding the correct definition for what it means to "learn" that a given query is in the intersection, and constructing an eavesdropper that makes only $O(C \cdot \ell)$ queries, are the main difficulty in our proof. [3]

1.2 Related Work

Impagliazzo and Rudich [10] showed that the key of any ℓ-query key-agreement protocol in the random-oracle model can be revealed by an $\tilde{O}(\ell^6)$ query eavesdropper. Barak and Mahmoody [2] improve this bound and present an $O(\ell^2)$ query eavesdropper for this task, which shows that Merkle puzzles is optimal in this respect. Haitner, Omri, and Zarosim [8] used the machinery of [2] to relate the security of protocols that do not use a random oracle and solve tasks with no input, to the security of no-input protocols in the random-oracle model against an $O(\ell^2)$-query adversary. Finding limitation on the usefulness of random oracles for protocols that do take input seems to be a more difficult question. Chor and Kushilevitz [4] and Mahmoody et al. [11] made some progress in this direction. Finally, Haitner, Hoch, Reingold, and Segev [7] gave lower bounds on the communication complexity of statistically hiding commitments and single-server private information retrieval in a weaker oracle model that captures the hardness of one-way functions/permutation more closely than the random-oracle model.

1.3 Organization

We begin by giving a high-level overview of our proof techniques in Sections 2 and 3. Formal definitions and notation used throughout the paper are given in Section 4. The bound for uniform-query protocols is formally stated and proved in Section 5, and the bound for two-message non-adaptive protocols is stated and proved in Section 6.

2 Uniform-Query Protocols: Proof Outline

Our lower bound for uniform-query key-agreement protocols is proved via a reduction to *set disjointness*, a classical problem in two-party communication complexity.

In the set disjointness problem, we have two players, Alice and Bob. The players receive inputs $X, Y \subseteq [n]$, respectively, of size $|X| = |Y| = \ell$, and the players must determine whether $X \cap Y = \emptyset$. To do this, the players communicate with each other, and the question is how many bits they must exchange. It is known [17] that for any sufficiently large $n \in \mathbb{N}$, if the size of the sets is $\ell = n/4$, then the players must exchange $\Omega(n)$ bits to solve set disjointness, and this holds even

[3] A lower bound on key-agreement protocols implies a lower bound for the Set-Intersection problem. This fact suggests that the proof cannot be simple.

for randomized protocols where the players have access to shared randomness and only need to succeed with probability $2/3$. Here, we require high success probability *on any input*, not over some specific input distribution. We note that in the 2-party communication complexity model there is no random oracle.

The connection between set disjointness and key agreement comes from the fact that the only *correlation* between the parties' views in a key agreement protocol comes from the *intersection queries*, the queries that both players ask and Eve does not know. Indeed, if Alice asks $A \subseteq [n]$ and Bob asks $B \subseteq [n]$, and the random oracle is $F : [n] \to [n]$, then $F(A \setminus (A \cap B))$ and $F(B \setminus (A \cap B))$ are *independent of each other*. In particular, if $A \cap B = \emptyset$, then $F(A)$ and $F(B)$ are independent, and intuitively, in this case the players cannot securely agree on a secret key, because they have no advantage over the eavesdropper. On the other hand, if $A \cap B \neq \emptyset$, then the players can exploit the correlation induced by $F(A \cap B)$ to securely agree on a secret key. Thus, any secure key agreement protocol "behaves differently" depending on whether $A \cap B = \emptyset$ or not, and we can use this to solve the set disjointness problem.

Suppose that we are given a secure key-agreement protocol Π, where the players make ℓ uniformly-random queries to an oracle $\mathrm{F} : [n] \to [n]$. For simplicity we assume that the protocol has *perfect agreement*, that is, the players always output the same key, and that the security parameter is $3/4$, that is, an eavesdropper has probability at most $3/4$ of outputting the same key as the players. Our full proof does not make these assumptions.

Now, we want to construct from the key-agreement protocol Π, which uses a random oracle, a protocol Π' for set disjointness, *without* a random oracle (as usual in communication complexity). To this end, we consider two possible ways of simulating Π without an oracle:

- Λ_{Com}: the players use their shared randomness to simulate the oracle. They interpret the shared randomness as a random function $\mathrm{F} : [n] \to [n]$, and whenever Π wants to query some element $q \in [n]$, the players use $\mathrm{F}(q)$ as the oracle's answer.

- Λ_{Dist}: the players use their *private* randomness to simulate the oracle. Alice and Bob interpret their private randomness as random functions $\mathrm{F}_A, \mathrm{F}_B : [n] \to [n]$, respectively. Whenever Π indicates that Alice should query an element $q \in [n]$, she uses $\mathrm{F}_A(q)$ as the answer, while Bob uses $\mathrm{F}_B(q)$.

The first simulation, Λ_{Com}, is "perfect": it produces exactly the correct distribution of transcripts and outputs under our key-agreement protocol Π. In particular, the keys produced by the players in Λ_{Com} always agree, and an eavesdropper that sees the transcript of Λ_{Com} (but not the shared randomness) can find the key with probability at most $3/4$.

On the other hand, the second simulation Λ_{Dist} is "wrong", because the players do not use the same random function to simulate the random oracle. In fact, it is known that without shared randomness, secure key agreement is *impossible*, as an eavesdropper that sees the transcript can find the key with the same probability that the players have of agreeing with each other. Therefore there are two possible cases:

Agreement gap: The probability that the players agree on the key in Λ_{Dist} is at most $7/8$ (compared to one in Λ_{Com}), or

Secrecy gap: There is an eavesdropper E that guesses Alice's key in Λ_{Dist} with probability at least $7/8$ (compared to $3/4$ in Λ_{Com}).

(Instead of $7/8$ we could have used here any constant probability in $(3/4, 1)$, but in the full proof this choice depends on the agreement and security parameters of Π.)

We divide into cases, depending on which of the two gaps we have.

Agreement gap. Assume that the players agree with probability at most $7/8$ in Λ_{Dist}. For simplicity, let us make the stronger assumption that for any intersection size $c > 0$, the probability of agreement between the players is at most $7/8$, even *conditioned* on the event that $|A \cap B| = c$. A general key-agreement protocol might not satisfy this assumption, which complicates the full proof significantly; see Section Section 5 for the details.

So, we assumed that whenever the intersection is non-empty, the players agree with probability at most $7/8$. Observe, however, that when the intersection *is* empty ($A \cap B = \emptyset$), the distribution of transcript and outputs in Λ_{Dist} is the same as in Π: although each player uses a different random function, they never ask the same query, so there is no inconsistency. Therefore, conditioned on $A \cap B = \emptyset$, in Λ_{Dist} the players have perfect agreement (as in Π). In other words, Λ_{Dist} behaves very differently when $A \cap B = \emptyset$, in which case the players always agree on the key, compared to the general case, where the players agree with probability at most $7/8$. We use this fact to *check* whether $A \cap B = \emptyset$. Thus, by checking whether or not they got the same key in Λ_{Dist}, the players get an indication for whether or not $A \cap B = \emptyset$.

Our set disjointness protocol Π' is defined as follows. Given inputs $X, Y \subseteq [n]$, respectively, the players simulate Λ_{Dist} several times. In each simulation, the players agree on a random permutation $\sigma : [n] \to [n]$ using their shared randomness, and then the players simulate Λ_{Dist} using their permuted inputs as the query set; that is, Alice feeds $A = \sigma(X)$ to Λ_{Dist} as her query set, and Bob feeds $B = \sigma(Y)$ to Λ_{Dist} as his query set. Note that A, B are uniformly random, *subject to* having an intersection of size $|X \cap Y|$.

After each simulation of Λ_{Dist}, the players send each other the keys output under Λ_{Dist}, and check if they got the same key. Finally, they output "$X \cap Y = \emptyset$" iff they got the same key in all the simulations of Λ_{Dist}.

Since Λ_{Dist} has perfect agreement when there is no intersection, the players always succeed when $X \cap Y = \emptyset$. However, by assumption, whenever $X \cap Y \neq \emptyset$, the probability of agreement in Λ_{Dist} is at most $7/8$, so if we repeat Λ_{Dist} sufficiently many times, the probability that all instances output the same key will be at most $1/3$.

Secrecy gap. In this case we convert Λ_{Com} and Λ_{Dist} into a pair of protocols with an agreement gap, and then proceed as above.

Consider the protocol Λ'_{Dist} where the parties acts as in Λ_{Dist}, but at the end, Bob executes the eavesdropper E on the transcript, and outputs the key that E outputs. Define Λ'_{Com} analogously.

By assumption, E guesses Alice's output in Λ_{Dist} with probability at least $7/8$, but in Λ_{Com} it succeeds with probability at most $3/4$. Thus, in Λ'_{Dist} the players agree with probability at least $7/8$, but in Λ'_{Com} they agree with probability at most $3/4$; there is a gap of at least $1/8$ between the probability of agreement in the two protocols (although they have switched roles and now Λ'_{Dist} has the higher agreement probability). Note also that Λ'_{Dist} does not have agreement probability 1, as we assumed for simplicity above, but our full proof can handle this case.

What about general protocols? It was important for our reduction to assume that the key-agreement protocol makes uniformly-random queries. Indeed, this reduction fails in the general case: consider the protocol where Alice and Bob always query 1, and output $\mathrm{F}(1)$ as their secret key. This protocol is completely insecure, since the eavesdropper can also query 1 and output $\mathrm{F}(1)$. But our reduction would not work for it, because the input distribution where both players get the set $\{1\}$ *is not hard for set disjointness* (indeed it is trivial). We see that the "hardness" of secure key-agreement is not necessarily that it is hard for the players to find their intersection queries, but that the eavesdropper should not be able to *predict* the intersection queries that the players use. Our second lower bound makes this intuition explicit and uses it to get a lower bound on two-round protocols with arbitrary (but non-adaptive) query distributions.

3 Two-Message Non-Adaptive Protocols: Proof Outline

In this section we describe a lower bound on the communication cost of any key-agreement protocol that makes non-adaptive queries and uses two rounds of communication: we show that any such protocol that makes ℓ queries and is secure against an adversary that makes q queries must send a total of $\Omega(q/\ell)$ bits. In particular, taking $q = \Theta(\ell^2)$, this shows that Merkle's puzzles is optimal in its communication cost.

In this proof, we once again relate the parties' advantage over the eavesdropper to the information they gained about the intersection of their query sets. We show that to produce a shared key, the parties need to learn a lot of information about this intersection. Moreover, the query sets and their intersection need to be "unpredictable" (have high min-entropy) given the transcript, otherwise an eavesdropper could make the same queries and output the same key.

Preliminaries. In the proof we often need to measure differences between various distributions. For this purpose we use *f-divergences*: given a convex function $f : \mathbb{R} \to \mathbb{R}$ with $f(1) = 0$, and distributions P, Q, the *f-divergence of P from Q* is defined as

$$\mathsf{D}_f(\mathrm{P} \parallel \mathrm{Q}) = \sum_{q \in \mathrm{Q}} \Pr\left[\mathrm{Q} = q\right] f\left(\frac{\Pr\left[\mathrm{P} = q\right]}{\Pr\left[\mathrm{Q} = q\right]}\right).$$

Specifically, the two f-divergences we use in this paper are the *statistical distance*, obtained by taking $f(x) = |x-1|/2$, and the *KL divergence*, obtained by taking $f(x) = x \log x$. Each has its own nice properties and disadvantages: statistical distance is bounded in $[0,1]$ but it is not additive, while KL divergence is additive but unbounded (more on this below).

We frequently need to measure the "amount of dependence" between two random variables. Let $(\mathrm{X}, \mathrm{Y}) \sim P_{\mathrm{X},\mathrm{Y}}$ be random variables jointly distributed according to $P_{\mathrm{X},\mathrm{Y}}$, and let $P_{\mathrm{X}}, P_{\mathrm{Y}}$ be the marginal distribution of X and Y, respectively. Also, let $P_{\mathrm{X}} \times P_{\mathrm{Y}}$ be the product distribution where X and Y are sampled independently of each other, each from its marginal distribution $P_{\mathrm{X}}, P_{\mathrm{Y}}$ (respectively). To quantify the dependence between X and Y, we measure the difference between their joint distribution and the product of the marginals: formally, we define

$$\mathsf{I}_f(\mathrm{X}; \mathrm{Y}) = \mathsf{D}_f(P_{\mathrm{X},\mathrm{Y}} \parallel P_{\mathrm{X}} \times P_{\mathrm{Y}}).$$

This generalizes the usual notion of mutual information, which is the special case of I_f where we use KL divergence (i.e., when $f = x \log x$). For clarity, when we use KL divergence we omit the subscript f, and when using statistical distance, we use the notation I_{SD} (instead of $\mathrm{I}_{f(x)=|x-1|/2}$).

Finally, we also need the notion of *conditional mutual information*, which is simply the average mutual information between two variables X, Y, where the average is taken over a third random variable Z. Formally, let $(\mathrm{X}, \mathrm{Y}, \mathrm{Z}) \sim P_{\mathrm{X,Y,Z}}$. For any value z, let $P_{\mathrm{X,Y|Z}=z}, P_{\mathrm{X|Z}=z}, P_{\mathrm{Y|Z}=z}$ be the joint distribution of X, Y and the marginals of X and Y, respectively, all conditioned on the event $\mathrm{Z} = z$. Then we define $\mathrm{I}_f(\mathrm{X}; \mathrm{Y}|\mathrm{Z}) = \mathrm{E}_{z \sim P_\mathrm{Z}}\left[\mathsf{D}_f(P_{\mathrm{X,Y|Z}=z} \parallel P_{\mathrm{X|Z}=z} \times P_{\mathrm{Y|Z}=z})\right]$.

Some examples. Let us illustrate the ideas behind the lower bound by way of some examples.

Example 1: We already discussed the na ive example where both players query 1 and output $\mathrm{F}(1)$, and said that it is insecure because the eavesdropper can *predict* the intersection query. Here is another instantiation of this idea: Alice and Bob view the domain ℓ^2 as an $\ell \times \ell$ matrix, so that the oracle queries are represented by pairs $(i, j) \in [\ell]^2$. Alice chooses a row $a \in [\ell]$, and queries all the elements of the row (that is, all pairs (a, j) where $j \in [\ell]$); Bob chooses a column $b \in [\ell]$ and queries all the elements of the column (all pairs (i, b) where $i \in [\ell]$). Then, Alice sends a to Bob, who responds with $\mathrm{F}(a, b)$. From $\mathrm{F}(a, b)$, Alice can compute b, by finding the (w.h.p. unique) index j such that $\mathrm{F}(a, b) = \mathrm{F}(a, j)$. Both players output the first bit of b as the key.

This protocol is slightly less naïve than the previous one: now there are no queries that have high prior probability of being asked, and the index b of the query that determines the key is uniformly random a-priori. However, once Alice sends a to Bob, the game is up: Eve can also query row a and find b the same way Alice does.

We see that in addition to queries that have a high prior probability of being asked, Eve also needs to ask queries that have a high *posterior* probability of being asked, after she sees M_1. It turns out that this is enough: if we were to continue for more than 2 rounds, then Eve would also need to ask queries that become likely after seeing M_2, and so on, but to prove a 2-round lower bound, Eve does not need to ask these queries. Intuitively, if a query only becomes likely after M_2 is sent, then this is "too late" for it to be useful to the players, and Eve can ignore it.

Example 2: First, both players query 1. Then they carry out the protocol from Example 1, but all messages are "encrypted" by XOR-ing them with $\mathrm{F}(1)$.

From this example we see that Eve needs to be somewhat adaptive: when she decides what queries to ask after seeing M_1, she must incorporate the queries she asked before the first round (in this case, she would query 1). Essentially, when Eve tries to understand what the players have done in round i, she should take into account all the queries she made up to round i.

Should Eve be adaptive *inside* each round? In other words, after seeing M_1, should she ask all queries $\mathcal{E}_1$ that became likely, then compute which new queries are now likely given $\mathrm{M}_1, \mathcal{E}_1$, and so on, until she reaches a fixpoint?

It turns out that for our purposes here, because we consider non-adaptive protocols, Eve does not need to do this.

Heavy queries. Our attacker Eve tries to break the security of the protocol by asking all queries that are "somewhat likely" to be asked by the players; these queries are called *heavy queries*. Informally, a query $q \in \{0, 1\}^n$ is *heavy* after round i if given the transcript up to round i (inclusive), and given Eve's queries up to round i, the probability that q is asked by one (or both) of the players exceeds some threshold δ which is fixed in advance.

More formally, the set $\mathcal{E}_i$ of heavy queries after round i is defined by induction on rounds, as follows: the a-priory heavy queries, $\mathcal{E}_0$, are given by

$$\mathcal{E}_0 = \{q \in \{0,1\}^n : \Pr[q \in X \cup Y] \geq \delta\}.$$

These are queries that are "somewhat likely" to be asked before the protocol begins. For $i > 0$, we define

$$\mathcal{E}_i = \mathcal{E}_{i-1} \cup \{q \in \{0,1\}^n : \Pr[q \in X \cup Y \mid M_{\leq i}, \mathrm{F}(\mathcal{E}_{i-1})] \geq \delta\}.$$

In other words, after round i, Eve asks all queries $q \in \{0,1\}^n$ that have probability at least δ of being queried by the players, given the messages $M_{\leq i}$ that Eve observed up to round i and the heavy queries she asked before, $\mathcal{E}_{i-1}$.

A simplified normal form for protocols. To simplify the proof of the lower bound, we first apply an easy transformation to the protocol: given a key agreement protocol Π, we construct a protocol Π', which has nearly the same communication and query complexity as Π, the same number of rounds, and the same agreement and nearly the same security parameters. But Π' also has the following properties: first, Π' has no a-priori heavy queries, that is, $\mathcal{E}_0 = \emptyset$; and second, the secret key output by Bob in Π' is the first bit of Bob's last query. This easy transformation is omitted in this overview.

3.1 Measuring the Players' Advantage Over Eve

As we saw in the examples above, the players' ability to produce a shared secret key is closely tied to how much information *the players* have that *Eve does not have* about the intersection of the query sets, $X \cap Y$.

To quantify this advantage, define the following random variables:

- $\mathrm{S}_i = (\mathrm{X} \cap \mathrm{Y}) \setminus \mathcal{E}_{i-1}$, the intersection queries that have not been asked by Eve.
- $\mathrm{F}(\mathrm{S}_i)$: the answers to the queries in S_i.
- $V_E^i = (\mathcal{E}_i', \mathrm{F}(\mathcal{E}_i'))$: a subset of the heavy queries for the previous round, and the answers to them. Here, $\mathcal{E}_i' \subseteq \mathcal{E}_i$ is a subset that will be defined later (and depends on the round number i). For technical reasons, it is convenient to use only some of the heavy queries in some contexts; (essentially, in some places in our proof, Eve uses only some of her power. This helps us avoid some unnecessary dependencies.

We measure the advantage gained by the players in round i by an expression of the form:

$$\mathrm{I}_f(\mathrm{S}_i, \mathrm{F}(\mathrm{S}_i); \mathrm{M}_i | \mathrm{Z}, \mathrm{V}_E^{i-1}, \mathrm{M}_{<i}), \tag{1}$$

where I_f is the information with respect to the f-divergence, $\mathrm{M}_i, \mathrm{M}_{<i}$ are the i-th message and the messages of rounds $1, \ldots, i-1$, Z is the query set of the player that sent M_i (either X or Y, depending on the round number i). Note that Eve uses her heavy queries from the previous rounds, V_E^{i-1}, to "try to understand" what is going on in the current round.

Intuitively, this expression measures how much information the i-th message conveys about the intersection queries and their answers, which Eve *cannot guess*. For this reason, the random variable S_i excludes intersection queries that were asked by Eve. Notice that on the right-hand side we condition on Eve's view (or on things Eve can sample): Eve has already seen the messages $\mathrm{M}_{<i}$ and asked the heavy queries $\mathrm{V}_E^{i-1} = (\mathcal{E}'_{i-1}, \mathrm{F}(\mathcal{E}'_{i-1}))$, and she can sample the queries Z, either X or Y, from the correct distribution given the transcript and her queries. Crucially, this does not require her to make any oracle queries: we do not require her to sample the answers F(Z), only the queries Z. In other words, Eve can *pretend* to be whichever player the query set Z belongs to, and by conditioning on her view, we essentially neutralize all the information that Eve can extract about the intersection. Thus, the expression in (1) measures the information the players gain about the intersection but that is hidden from Eve.[4]

Our proof consists of showing:

Step I: After the first message M_1 is sent, the advantage gained is small, only $O(\delta|\mathrm{M}_1|)$. For this part of the proof we use KL-divergence to measure the advantage.

Step II: After the second message M_2 is sent, the advantage is still small, only $O(\sqrt{\delta(|\mathrm{M}_1| + |\mathrm{M}_2|)})$. Here we use statistical distance to measure the advantage, for reasons we will explain below.

Step III: When the expression in (1) is small (i.e., the players only have a small "advantage"), then indeed, Eve can break the security of the protocol, by pretending to be one of the players and sampling the secret key that this player would output.

Next we explain in more detail how each step is carried out.

3.2 Outline of the Proof

Step III: How Eve breaks security. Let us start from the end: suppose that after the second round, the "advantage" is small:

$$\mathrm{I}_f(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathrm{V}_E^1) \leq \beta,$$

where $\beta = O(\sqrt{\delta(|\mathrm{M}_1| + |\mathrm{M}_2|)}) \ll 1$. Here, the advantage is measured in statistical distance (that is, we take $f(t) = |t-1|/2$). We want to show that Eve can break the security of the protocol, by guessing the secret key.

As we said, Eve's strategy is to "pretend" that she is Bob, and sample Bob's output, $\mathrm{out}^{\mathsf{B}}$.

In a general protocol, to do this, Eve needs to sample Bob's queries Y and the answers F(Y), and then she can compute $\mathrm{out}^{\mathsf{B}} = \mathrm{out}^{\mathsf{B}}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}), \mathrm{M}_1, \mathrm{M}_2)$. However, recall that we transformed the protocol so that $\mathrm{out}^{\mathsf{B}}$ is a fixed function of Y; therefore, Eve in fact needs to do nothing clever, only sample Y given her view $\mathrm{M}_1, \mathrm{M}_2, \mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)$ and compute $\mathrm{out}^{\mathsf{B}}$ from Y.

We need to show that Eve's key is close to the correct distribution, the one used by the players. In general, if too much communication is allowed, this is not true, as shown by the following example.

[4] As we said above, we use only some of Eve's heavy queries, $\mathcal{E}'_{i-1} \subseteq \mathcal{E}_i$, so this intuition is not completely accurate; specifically, when (1) is *large*, it does not mean that Eve cannot guess a lot about the intersection, because she could use the full set $\mathcal{E}_i$. However, when (1) is *small*, then indeed Eve knows almost as much about the intersection as the players do, because her view *includes* V_E^{i-1} (and possibly more).

Example: In Merkle's puzzles, Alice's message is $\mathrm{F(X)}$, and Bob responds with $\mathrm{F(s)}$, where $\mathrm{s} \in \mathrm{X} \cap \mathrm{Y}$ is some intersection query. The original secret key (before our transformation) is the first bit s^1. After our transformation, the secret key is $\mathrm{Y}^1_{\ell+1}$, and as part of M_2, Bob sends Alice the bit $\mathrm{b} = \mathrm{s}^1 \oplus \mathrm{Y}^1_{\ell+1}$ so that she can extract $\mathrm{Y}^1_{\ell+1}$.
From Alice's perspective, given $\mathrm{X}, \mathrm{F(X)}$, Bob's message $\mathrm{M}_2 = \mathrm{F(s)}, \mathrm{b}$ fixes $\mathrm{Y}^1_{\ell+1}$ to the value $\mathrm{b} \oplus \mathrm{s}^1$. (We ignore here the tiny probability that s cannot be uniquely computed from $\mathrm{X}, \mathrm{F(X)}$ and $\mathrm{F(s)}$, i.e., the probability of a collision in F.) However, from Eve's perspective, because she does not know $\mathrm{X}, \mathrm{F(X)}$ and she asks no queries (there are no heavy queries in Merkle's puzzles), the intersection element s remains uniformly random. When Eve samples $\mathrm{Y}^1_{\ell+1}$ given $\mathrm{M}_1, \mathrm{M}_2$ and her non-existent heavy queries, the result is random, and completely independent from the true secret key.

We need to show that when the players' advantage is small, then the example above cannot happen, and Eve's key agrees with the players' w.h.p. To this end, we are interested in the difference between Eve's "pretend distribution", and the true distribution that the players use to produce the key: if the two distributions are close, then Eve's chances of guessing the right secret key are roughly the same as Bob's. The *only difference* between these two distributions is that given $\mathrm{M}_1, \mathrm{M}_2$ and $\mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)$ (which the players do not use),

- The players' keys are produced according to the *joint distribution* $(\mathrm{out}^{\mathsf{A}}, \mathrm{out}^{\mathsf{B}})$, and in particular, both players have the same answers $\mathrm{F(S_2)}$ to the non-heavy intersection queries $\mathrm{S}_2 = (\mathrm{X} \cap \mathrm{Y}) \setminus \mathcal{E}_1$.
- Eve's pretense that she is Bob is carried out *independently* from Alice's view: Eve cannot use the true intersection queries (which she does not know), only what she has learned about them from $\mathrm{M}_1, \mathrm{M}_2, \mathrm{V}^i_E$. The joint distribution of Alice and Eve's keys is therefore given by the product distribution $\mathrm{out}^{\mathsf{A}} \times \mathrm{out}^{\mathsf{B}}$.

So, we would like to bound the difference between the joint distribution and the product distribution, i.e.,

$$\mathrm{I}_f(\mathrm{out}^{\mathsf{A}}; \mathrm{out}^{\mathsf{B}} \,|\mathrm{M}_1, \mathrm{M}_2, \mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)).$$

Given the conditioning, Alice's output $\mathrm{out}^{\mathsf{A}}$ is a function of her view, $\mathrm{X}, \mathrm{F(X)}$. Also, we assumed that Bob's output is a function of his queries Y. Therefore,

$$\mathrm{I}_f(\mathrm{out}^{\mathsf{A}}; \mathrm{out}^{\mathsf{B}} \,|\mathrm{M}_1, \mathrm{M}_2, \mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)) \leq \mathrm{I}_f(\mathrm{X}, \mathrm{F(X)}; \mathrm{Y}|\mathrm{M}_1, \mathrm{M}_2, \mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)). \tag{2}$$

Now we need to show that given Eve's view, the dependence between $\mathrm{X}, \mathrm{F(X)}$ and Y is bounded in terms of the advantage:

$$\mathrm{I}_f(\mathrm{X}, \mathrm{F(X)}; \mathrm{Y}|\mathrm{M}_1, \mathrm{M}_2, \mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)) \leq \mathrm{I}_f(\mathrm{S}_2, \mathrm{F(S_2)}; \mathrm{M}_2|\mathrm{M}_1, \mathrm{Y}, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})). \tag{3}$$

This proof is somewhat tedious; it relies on the fact that M_2 is a function of M_1, Y and $\mathrm{F(Y)}$, and on the fact that $\mathrm{X}, \mathrm{F(X)}$ are independent of $\mathrm{Y}, \mathrm{F(Y)}$ given the intersection queries and answers, $\mathrm{S}_2, \mathrm{F(S_2)}$ and $\mathcal{E}_1, \mathrm{F}(\mathcal{E}_1)$. Intuitively, all the dependence between $\mathrm{X}, \mathrm{F(X)}$ and Y "flows through" what the players learn about the intersection, and the proof of (3) formalizes this intuition.

Step I: Bounding the advantage after the first round. For the first round, we analyze the players' advantage in terms of KL-divergence, and bound

$$\mathrm{I}(\mathrm{S}_1, \mathrm{F(S_1)}; \mathrm{M}_1|\mathrm{X}).$$

Notice that we do not use Eve at this point, because we eliminated any a-priory heavy queries, so there is nothing Eve needs to query in order to "understand" M_1. For the same reason, $\mathrm{S}_1 = \mathrm{X} \cap \mathrm{Y}$ (there are no heavy queries to remove from the intersection).

We claim that

$$\mathrm{I}(\mathrm{S}_1, \mathrm{F}(\mathrm{S}_1); \mathrm{M}_1 | \mathrm{X}) \leq \delta |\mathrm{M}_1|. \tag{4}$$

This is not hard to see: suppose $\mathrm{X} = x$. Because we got rid of the a-priori heavy queries, every individual query $q \in x$ has probability at most δ of being asked by Bob (otherwise, q would be heavy). Therefore, for every $q \in x$, we have $\Pr[q \in \mathrm{S}_1 | \mathrm{X} = x] \leq \delta$. Because M_1 is generated by Alice without knowing S_1, and every query is in S_1 only w.p. at most δ, intuitively, the information in M_1 "only applies" to the queries in S_1 with probability δ. Therefore the information that M_1 gives about $\mathrm{S}, \mathrm{F}(\mathrm{S}_1)$ is at most $\delta|\mathrm{M}_1|$.

The actual proof involves a Shearer-like argument for mutual information, similar to the ones used in [6, 16].

Step II: Bounding the advantage after the second round. Now we must bound the advantage the players gain after the second round, and show that

$$\mathrm{I}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) = O(\sqrt{|\mathrm{M}_1| + |\mathrm{M}_2|}). \tag{5}$$

As we said, we switch here to using statistical distance, and we will see why below.

Following the first round, we know that not much is known about the intersection, because Alice's message M_1 did not convey a lot of information about it. So, our proof here proceeds in two steps: first, we "pretend" that *nothing* is known about the intersection, and consider the distribution μ' where given M_1 the distribution of $\mathrm{Y}, \mathrm{F}(\mathrm{Y})$ is completely independent from X. We show that under μ', Bob's message M_2 would only convey $\delta|\mathrm{M}_2|$ bits of information about the intersection. This is very similar to the analysis of the first round, and it is also carried out using KL-divergence. Formally, we show that for the distribution μ' where $\mathrm{Y}, \mathrm{F}(\mathrm{Y})$ are drawn independently of of X, we have

$$\mathrm{I}^{\mu'}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) \leq \delta |\mathrm{M}_2|. \tag{6}$$

The proof relies on the fact that we excluded heavy queries from S_2 (recall that $\mathrm{S}_2 = (\mathrm{X} \cap \mathrm{Y}) \setminus \mathcal{E}_1$), so given the conditioning, any query in Y can only belong to S_2 with probability at most δ.

However, μ' is not the real distribution: given M_1, we do know a little about the intersection, so $\mathrm{Y}, \mathrm{F}(\mathrm{Y})$ are not completely independent from X. Our next step is to switch to statistical distance, and show that the real distribution μ (where X, Y are not independent) and μ' (where they are) are close to each other. Therefore, what we showed for μ' is also true for μ, with the addition of a small penalty corresponding to the distance between μ and μ'.

Formally, we prove that

$$\begin{aligned} &\mathrm{I}^{\mu}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) \\ &\leq O\left(\mathrm{I}^{\mu'}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) + \mathsf{D}_{SD}(\mu' \parallel \mu)\right) \end{aligned} \tag{7}$$

Under μ', by (6) and Pinsker's inequality, we have:

$$\mathrm{I}^{\mu'}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2 | \mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) \leq \sqrt{\delta |\mathrm{M}_2|}. \tag{8}$$

So, under μ' the expected amount of information revealed is small.

Next, we bound the difference between μ and μ'. We show that:

$$\mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1) \leq \mathrm{I}(\mathrm{S}_1, \mathrm{F}(\mathrm{S}_1); \mathrm{M}_1|\mathrm{X}).$$

This is quite similar to the proof of Step III above — here we do use standard mutual information, so the proof uses the chain rule, just as we did above. Since we have shown in Step I that $\mathrm{I}(\mathrm{S}_1, \mathrm{F}(\mathrm{S}_1); \mathrm{M}_1|\mathrm{X}) \leq \delta|\mathrm{M}_1|$, we conclude using Pinsker's inequality that

$$\mathsf{D}_{SD}(\mu' \parallel \mu) \leq \sqrt{\mathsf{D}_{KL}(\mu' \parallel \mu)} \leq \sqrt{\delta|\mathrm{M}_1|}. \tag{9}$$

Together, (8) and (9) are the ingredients we need to apply (7), and obtain:

$$\begin{aligned}
&\mathrm{I}^{\mu}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2|\mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y}))\\
&\leq O\left(\mathrm{I}^{\mu'}_{SD}(\mathrm{S}_2, \mathrm{F}(\mathrm{S}_2); \mathrm{M}_2|\mathrm{Y}, \mathrm{M}_1, \mathcal{E}_1 \cap \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y})) + \mathsf{D}_{SD}(\mu' \parallel \mu)\right)\\
&\leq O(\sqrt{\delta(|\mathrm{M}_1| + |\mathrm{M}_2|)}).
\end{aligned}$$

4 Preliminaries

4.1 Notations

We use calligraphic letters to denote sets, uppercase for random variables and lowercase for values. For $m \in \mathbb{N}$, let $[m] = \{1, \ldots, m\}$. For a random variable X, let $x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}$ to denote that x is chosen according to X. Similarly, for a set S let $s \overset{\mathrm{R}}{\leftarrow} \mathsf{S}$ to denote that s is chosen according to the uniform distribution over S. The support of the distribution D, denoted $\mathrm{Supp}(D)$, is defined as $\{u \in \mathcal{U} : \Pr_D[u] > 0\}$. The statistical distance between two distributions P and Q over a finite set $\mathcal{U}$, denoted $\mathrm{SD}(P, Q)$, is defined as $\frac{1}{2}\sum_{u \in \mathcal{U}} |\Pr_P[u] - \Pr_Q[u]|$, which is equal to $\max_{\mathcal{S} \subset \mathcal{U}}(\Pr_P[\mathcal{S}] - \Pr_Q[\mathcal{S}])$.

For a vector $\mathbf{X} = X_1, ..., X_n$ and an index $i \in [n]$, let $X_{<i}$ denote the vector $X_1, ..., X_{i-1}$ and $X_{\leq i}$ denote the vector $X_1, ..., X_i$. For a set of indexes $T = \{i_1, \ldots, i_k\} \subseteq [n]$ such that $i_1 < i_2 < \cdots < i_k$, let X_T denote the vector $X_{i_1}, \ldots, X_{i_k}$. Similarly, $X_{T,<i}$ denotes the vector $X_{T \cap \{1,\ldots,i-1\}}$. For a function f, let $f(\mathbf{X}) = (f(X_1), ..., f(X_n))$.

For random variables A and B we use $\mathrm{A}|_{\mathrm{B}=b}$ to denote the distribution of A condition on the event $\mathrm{B} = b$, and $\mathrm{A} \times \mathrm{B}$ to denote the product between the marginal distributions of A and B. When A is independent from B we write $\mathrm{A} \perp \mathrm{B}$ to emphasize that this is the case.

4.2 Interactive Protocols

A two-party protocol $\Pi = (\mathsf{A}, \mathsf{B})$ is a pair of probabilistic interactive Turing machines. The communication between the Turing machines A and B is carried out in rounds, where in each round one of the parties is active and the other party is idle. In the j-th round of the protocol, the currently active party P acts according to its partial view, writing some value on its output tape, and then sending a message to the other party (i.e., writing the message on the common tape). The communication transcript (henceforth, the transcript) of a given execution of the protocol $\Pi = (\mathsf{A}, \mathsf{B})$, is the list of messages m exchanged between the parties in an execution of the protocol, where $m_{1,\ldots,j}$ denotes the first j messages in m. A view of a party contains its input, its random tape and the messages exchanged by the parties during the execution. Specifically, A's view is a tuple $v_{\mathsf{A}} = (i_{\mathsf{A}}, r_{\mathsf{A}}, m)$, where i_{A} is A's input, r_{A} are A's random coins, and m is the transcript of the execution. Let out^{A} denote the output of A in the end of the protocol, and out^{B} B's output. Notice that given a protocol, the transcript and the outputs are deterministic function of the joint view $(i_{\mathsf{A}}, r_{\mathsf{A}}, i_{\mathsf{B}}, r_{\mathsf{B}})$. For a joint view v, let $\mathsf{trans}(v)$, $\text{out}^{\mathsf{A}}(v)$ and $\text{out}^{\mathsf{B}}(v)$ be the transcript of the protocol and the parties' outputs determined by v. For a distribution D we denote the distribution over the parties' joint view in a random execution of Π, with inputs drawn from D by $\Pi(D)$.

A protocol Π has r rounds, if for every possible random tapes for the parties, the number of rounds is exactly r. The Communication Complexity of a protocol Π, denoted as $\text{CC}(\Pi)$ is the length of the transcript of the protocol in the worst case.

4.3 Oracle-Aided Protocols

An oracle-aided two-party protocol $\Pi = (\mathsf{A}, \mathsf{B})$ is a pair of interactive Turing machines, where each party has an additional tape called the oracle tape; the Turing machine can make a query to the oracle by writing a string q on its tape. It then receives a string ans (denoting the answer for this query) on the oracle tape. An oracle-aided protocol is ℓ-queries protocol if each party makes at most ℓ queries during each run of the protocol. In a *non-adaptive* oracle-aided protocol, the parties choose their queries before the protocol starts and before querying the oracle. A *uniform query* oracle-aided protocol, is a non-adaptive protocol in which the parties queries are chosen uniformly form a predetermined set.

4.4 Key-Agreement Protocols

Since we are giving lower bounds, we focus on single bit protocols.

Definition 4.1 (key-agreement protocol)**.** *Let $0 \leq \gamma$, $\alpha \leq 1$ and $q \in N$. A two-party boolean output protocol $\Pi = (\mathsf{A}, \mathsf{B})$ is a (q, α, γ)-key-agreement relative to a function family $\mathcal{F}$, if the following hold:*

Accuracy: Π has $(1-\alpha)$-accuracy. *For every $f \in \mathcal{F}$:*

$$\Pr_{v \overset{R}{\leftarrow} \Pi^f}\left[\text{out}^{\mathsf{A}}(v) = \text{out}^{\mathsf{B}}(v)\right] \geq 1 - \alpha.$$

Secrecy: Π has (q, γ)-secrecy. *For every q-query oracle-aided algorithm E:*

$$\Pr_{f \overset{R}{\leftarrow} \mathcal{F}, v \overset{R}{\leftarrow} \Pi^f}\left[\mathsf{E}^f(\mathsf{trans}(v)) = \text{out}^{\mathsf{A}}(v)\right] \leq \gamma.$$

If $\mathcal{F}$ is a trivial function family (e.g., $\mathcal{F}$ contains only the identity function), then all correlation between the parties' view is implied by the transcript. Hence, an adversary that on a given transcript τ samples a random view for A that is consistent with τ, and outputs whatever A would upon this view, agrees with B with the same probability as does A. This simple argument yields the following fact.

Fact 4.2. *For every $0 \leq \alpha \leq 1$ and $0 \leq \gamma < 1 - \alpha$, there exists no (q, α, γ)-key-agreement protocol relative to the trivial family.*

4.5 Entropy and Information

The Shannon Entropy of a random variable A is defined as $\mathrm{H}(\mathrm{A}) = \sum_{a\in\mathsf{Supp}(\mathrm{A})} \Pr_{\mathrm{A}}[a] \log \frac{1}{\Pr_{\mathrm{A}}[a]}$. The conditional entropy of a random variable A given B is defined as $\mathrm{H}(\mathrm{A}|\mathrm{B}) = E_{b \xleftarrow{\mathrm{R}} \mathrm{B}}[\mathrm{H}(\mathrm{A}|_{\mathrm{B}=b})]$. The following fact is called the chain rule of Shannon Entropy:

Fact 4.3 (Chain rule for entropy)**.** *For a random variable* $\boldsymbol{A} = \mathrm{A}_1, ..., \mathrm{A}_n$ *the following holds:*

$$\mathrm{H}(\mathrm{A}_1, ..., \mathrm{A}_n) = \sum_{i=1}^{n} \mathrm{H}(\mathrm{A}_i|\mathrm{A}_1, ...\mathrm{A}_{i-1}).$$

For a function f, the f-divergence between random variables A and B, denoted as $D_f(\mathrm{A}, \mathrm{B})$, is defined as $D_f(\mathrm{A}, \mathrm{B}) = \sum_{b\in\mathrm{B}} \Pr[\mathrm{B} = b]\, f(\frac{\Pr[\mathrm{A}=b]}{\Pr[\mathrm{B}=b]})$. We use I_f as "mutual information with respect to the f-divergence", $\mathrm{I}_f(\mathrm{A};\mathrm{B}) = D_f((A, B), (A \times B))$. The conditional mutual information, $\mathrm{I}_f(\mathrm{A};\mathrm{B}|\mathrm{C})$ is defined as $\mathrm{E}_{c \xleftarrow{\mathrm{R}} \mathrm{C}}[\mathrm{I}_f(\mathrm{A}|_{\mathrm{C}=c}; \mathrm{B}|_{\mathrm{C}=c})]$

For $f(t) = 1/2|t - 1|$, $\mathrm{I}_f = \mathrm{I}_{SD}$ is the statistical distance between the joint distribution to the product, that is, $\mathrm{I}_f(\mathrm{A};\mathrm{B}) = \mathrm{SD}((\mathrm{AB}), (\mathrm{A} \times \mathrm{B}))$.

For $f(t) = t \log t$, the f-divergence is called the KL-divergence, and I_f (from here denoted as I_{KL} or simply I) is the mutual information $\mathrm{I}(\mathrm{A};\mathrm{B}) = \mathrm{H}(\mathrm{A}) - \mathrm{H}(\mathrm{A}|\mathrm{B})$. The mutual information is known to be symmetric, and the following facts are known:

Fact 4.4 (Chain rule for information)**.** *For random variables* $\boldsymbol{A} = \mathrm{A}_1, ..., \mathrm{A}_n$ *and* B*,*

$$\mathrm{I}(\mathrm{A};\mathrm{B}) = \sum_{i=1}^{n} \mathrm{I}(\mathrm{A}_i;\mathrm{B}|\mathrm{A}_1, ..., \mathrm{A}_{i-1}).$$

Fact 4.5. *For every random variables* A *and* B*,* $0 \leq \mathrm{I}(\mathrm{A};\mathrm{B}) \leq \mathrm{H}(\mathrm{A}) \leq |\mathrm{A}|$.

Fact 4.6 (Data processing inequality)**.** *Let* A, B *be random variables, and* f *a function. Then:* $\mathrm{I}(f(\mathrm{A});\mathrm{B}) \leq \mathrm{I}(\mathrm{A};\mathrm{B})$ *and* $\mathrm{H}(f(\mathrm{A})) \leq \mathrm{H}(\mathrm{A})$.

Lastly, a connection between mutual information and statistical distance is known:

Fact 4.7 (Pinsker's inequality)**.**

$$I_{SD}(\mathrm{A};\mathrm{B}) \leq 2\sqrt{\mathrm{I}(\mathrm{A};\mathrm{B})}.$$

We will also use the next general lemmas in our proof. The proofs are in Appendix B.

Lemma 4.8. *For every random variables* $\mathrm{A}, \mathrm{B}, \mathrm{C}$ *and* D *it holds that*

$$-\mathrm{I}(\mathrm{A};\mathrm{D}|\mathrm{C}) \leq \mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{C},\mathrm{D}) - \mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{C}) \leq \mathrm{I}(\mathrm{A};\mathrm{D}|\mathrm{C},\mathrm{B})$$

.

The next two lemmas are useful in bounding information by using Bernoulli random variables:

Lemma 4.9. *Let* J *be a Bernoulli random variable, s.t.* $\Pr[\mathrm{J} = 1] \leq 1/2$*. Then*

$$\mathrm{H}(\mathrm{J}) \leq \Pr[\mathrm{J} = 1] (\log \frac{1}{\Pr[\mathrm{J} = 1]} + 4).$$

Lemma 4.10. *Let* $\mathrm{A}, \mathrm{B}, \mathrm{M}$ *and for each* $m \in M$ E_m *be random variables. Let* J_m *be the indicator for the event* $\mathrm{M} = m$*, then*

$$\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{M},\mathrm{E}_M) \leq \sum_{m\in M} \left[I(\mathrm{A};\mathrm{B}|\mathrm{E}_m) + I(\mathrm{J}_m;\mathrm{B}|\mathrm{E}_m,\mathrm{A})\right].$$

4.5.1 Some Useful Facts

Fact 4.11 (Data processing inequality for statistical distance)**.** *Let* A, B *be random variables, and* f *a function. Then:* $\mathrm{SD}(f(\mathrm{A}), f(\mathrm{B})) \leq \mathrm{SD}(\mathrm{A}, \mathrm{B})$.

Fact 4.12. *Let* $\mathrm{A}, \mathrm{B}, \mathrm{C}$ *be random variables. Then:*

$$\mathrm{SD}((\mathrm{A}, \mathrm{B}), (\mathrm{A}, \mathrm{C})) = \mathop{\mathrm{E}}_{a \overset{R}{\leftarrow} \mathrm{A}} \left[\mathrm{SD}(\mathrm{B}|_{\mathrm{A}=a}, \mathrm{C}|_{\mathrm{A}=a})\right].$$

Fact 4.13. *Let* $\mathrm{A}, \mathrm{B}, \mathrm{C}$ *be random variables. Then:* $\mathrm{SD}((\mathrm{A} \times \mathrm{B}), (\mathrm{A} \times \mathrm{C})) = \mathrm{SD}(\mathrm{B}, \mathrm{C})$.

Fact 4.14 (Hoeffding's inequality[9])**.** *Let* $\mathrm{A}_1, ..., \mathrm{A}_n$ *be independent random variables s.t.* $\mathrm{A}_i \in [0, 1]$ *and let* $\widehat{\mathrm{A}} = \frac{1}{n}\Sigma_{i=1}^{n}\mathrm{A}_i$*. It holds that:*

$$\Pr\left[\widehat{\mathrm{A}} - \mathrm{E}\left[\widehat{\mathrm{A}}\right] \geq t\right] \leq e^{-2nt^2}.$$

Fact 4.15 (Jensen's inequality)**.** *Let* f *be some convex function, and* $x_1, ..., x_n$ *some numbers in* f*'s domain. And let* $w_1, ..., w_n$ *be positive weights such that* $\Sigma w_i = 1$*. Then:*

$$f(\Sigma w_i x_i) \geq \Sigma w_i f(x_i).$$

The proofs for the next three lemmas are appear in Appendix B:

Lemma 4.16. *Let* A, B *and* C *be random variables. Then*

$$\mathop{\mathrm{E}}_{c \overset{R}{\leftarrow} \mathrm{C}} \left[I_{SD}\left(\mathrm{A}; \mathrm{B}|_{\mathrm{C}=c}\right)\right] \leq 2 I_{SD}\left(\mathrm{A}, \mathrm{C}; \mathrm{B}\right).$$

Lemma 4.17. *Let* A, B *and* M *be random variables. Then*

$$I_{SD}\left(\mathrm{M}; \mathrm{A}\right) \leq \mathop{\mathrm{E}}_{b \overset{R}{\leftarrow} \mathrm{B}} \left[I_{SD}\left(\mathrm{M}; \mathrm{A}|_{\mathrm{B}=b}\right)\right] + I_{SD}\left(\mathrm{A}; \mathrm{B}\right).$$

Lemma 4.18. *Let* A, B *and* M *be random variables. Then*

$$\mathop{\mathrm{E}}_{m \overset{R}{\leftarrow} \mathrm{M}} \left[I_{SD}\left(\mathrm{A}; \mathrm{B}|_{\mathrm{M}=m}\right)\right] \leq 2 \mathop{\mathrm{E}}_{b \overset{R}{\leftarrow} \mathrm{B}} \left[I_{SD}\left(\mathrm{A}; \mathrm{M}|_{\mathrm{B}=b}\right)\right] + 2 I_{SD}\left(\mathrm{A}; \mathrm{B}\right).$$

For our proof we need only the following specific case of Lemma 4.18:

Corollary 4.19. *Let* A, B *and* M *be random variables, such that* $\mathrm{A} \perp \mathrm{B}$*. Then*

$$\mathop{\mathrm{E}}_{m \overset{R}{\leftarrow} \mathrm{M}} \left[I_{SD}\left(\mathrm{A}; \mathrm{B}|_{\mathrm{M}=m}\right)\right] \leq 2 \mathop{\mathrm{E}}_{b \overset{R}{\leftarrow} \mathrm{B}} \left[I_{SD}\left(\mathrm{A}; \mathrm{M}|_{\mathrm{B}=b}\right)\right].$$

5 Uniform-Query Protocols

In this section, we prove a lower bound on the communication complexity of *uniform-query* key-agreement protocols. Recall that an oracle-aided protocol has *uniform-queries*, if the queries made by the parties are uniformly chosen independently from an (a-priori fixed) domain. Our bound is that an ℓ-uniform-query protocol secure against ℓ^2-query eavesdropper, must have communication complexity $\Omega(\ell)$. It follows that the uniform-query protocol of Merkle [13] (i.e., Merkle puzzle) has optimal communication complexity (up to a log factor) for such protocols. We prove the bound by exhibiting a reduction from uniform-query key-agreement protocol to (no oracle) protocol for solving the *set-disjointness problem.*

Definition 5.1 (Set-disjointness)**.** *Protocol* $\Pi = (\mathsf{A},\mathsf{B})$ solves set-disjointness with error ϵ over distribution D (with support $(\{0,1\}^*)^* \times (\{0,1\}^*)^*$)*, if*

$$\Pr_{\substack{(\mathcal{X},\mathcal{Y})\overset{R}{\leftarrow}D \\ r_{\mathsf{A}}\overset{R}{\leftarrow}\{0,1\}^*, r_{\mathsf{B}}\overset{R}{\leftarrow}\{0,1\}^* \\ r_p\overset{R}{\leftarrow}\{0,1\}^*}} \left[(\mathsf{A}(\mathcal{X};r_{\mathsf{A}}),\mathsf{B}(\mathcal{Y};r_{\mathsf{B}}))(r_{\mathsf{P}}) = (\mathcal{X}\cap\mathcal{Y}=\emptyset \wedge \mathcal{X}\cap\mathcal{Y}=\emptyset)\right] \geq 1-\varepsilon.$$

Namely, with save but probability ε over the instance in hand and their private and public randomness, the parties find outs whether their two input sets intersect. Our reduction is to solving set-disjointness over the distribution below, known to be hard for low complexity protocols.

Definition 5.2 (hard distribution for set-disjointness)**.** *For* $\ell \in \mathbb{N}$*, let*
$\mathcal{Q}_\ell^0 = \{\mathcal{X},\mathcal{Y} \subset [\ell]\colon\ |\mathcal{X}| = |\mathcal{Y}| = \lfloor \ell/4 \rfloor\ ,\ \mathcal{X}\cap\mathcal{Y}=\emptyset\}$ *and let*
$\mathcal{Q}_\ell^1 = \{\mathcal{X},\mathcal{Y} \subset [\ell]\colon\ |\mathcal{X}| = |\mathcal{Y}| = \lfloor \ell/4 \rfloor, |\mathcal{X}\cap\mathcal{Y}| = 1\}$*. Let* D_ℓ^0 *and* D_ℓ^1 *be the uniform distribution over* $\mathcal{Q}_\ell^0$ *and* $\mathcal{Q}_\ell^1$ *respectively, and let* $D_\ell = \frac{3}{4}\cdot D_\ell^0 + \frac{1}{4}\cdot D_\ell^1$*.*

Razborov [17] has shown that solving set-disjointness D_ℓ with small error require high communication complexity.

Theorem 5.3 (hardness of D_ℓ, [17])**.** *Exists* $\epsilon > 0$ *such that for every* $\ell \in \mathbb{N}$ *and a protocol* Π *that solves set-disjointness over* D_ℓ *with error* ϵ*, it holds that* $\mathrm{CC}(\Pi) \geq \Omega(\ell)$*.*

For a finite set $\mathcal{S}$, let $\mathcal{F}_{\mathcal{S}} = \{f : \mathcal{S} \mapsto \{0,1\}^*\}$ be the family of all functions from $\mathcal{S}$ to binary strings. Our reduction is stated in the following theorem.

Theorem 5.4 (from uniform-query key-agreement protocols to set-disjointness)**.** *Assume exists an* ℓ*-uniform-query* $(0,\alpha,\gamma)$*-key agreement protocol relative to* $\mathcal{F}_{\mathcal{S}}$*, for some set* $\mathcal{S}$*, of communication complexity* c*. Then there exists a protocol for solving set-disjointness over* D_ℓ *with* ϵ *error and communication complexity* $\frac{2^{15}\cdot\ell^4\cdot\log 1/\epsilon}{|\mathcal{S}|^2(1-\alpha-\gamma)^4}\cdot c$*.*

Note that the above theorem holds also for protocols that are only secure against eavesdropper without access to the oracle. Combining Theorems 5.3 and 5.4 yields the following bound on the communication complexity of uniform-query key-agreement protocols.

Theorem 5.5 (Main result for uniform-inputs protocols)**.** *For any* ℓ*-uniform-query* (q,α,γ)*-key agreement protocol* Π *relative to* $\mathcal{F}_{\mathcal{S}}$*, it holds that* $\mathrm{CC}(\Pi) \in \Omega((1-\alpha-\gamma)^4 q^2/\ell^3)$*.*

Proof. By Theorems 5.3 and 5.4, protocol Π has communication complexity $\Omega((1-\alpha-\gamma)^4\left|\mathcal{S}\right|^2/\ell^3)$. By Fact 4.2, an eavesdropper that queries all the elements in $\mathcal{S}$ can guess the key with probability $1-\alpha$. Since without loss of generality $1-\alpha > \gamma$, it must hold that $q < |\mathcal{S}|$. Hence, $\mathrm{CC}(\Pi) \in \Omega((1-\alpha-\gamma)^4 q^2/\ell^3)$. □

The rest of this section is devoted for proving Theorem 5.4. Assume there exists an ℓ-uniform-query $(0,\alpha,\gamma)$-key-agreement protocol $\Pi = (\mathsf{A},\mathsf{B})$ relative to the function family $\mathcal{F}_{\mathcal{S}}$. We use Π to create a (no-oracle) protocol of about the same communication complexity that finds out the intersection size of parties inputs. We complete the proof showing that the latter protocol can be used to solve set-disjointness over the hard distribution D_ℓ.

Protocol Λ_{Com} below emulates protocol Π relative to the family $\mathcal{F}_{\mathcal{S}}$, in the communication complexity model (where no oracle is given). The parties of Λ_{Com} emulate of the random oracle using their shared public randomness interpreted as (description of a) function from the function family.

Protocol 5.6 ($\Lambda_{\mathsf{Com}} = (\mathsf{A}_{\mathsf{Com}}, \mathsf{B}_{\mathsf{Com}})$)**.**

$\mathsf{A}_{\mathsf{Com}}$*'s input: an ℓ-element set $\mathcal{X} \subseteq \mathcal{S}$.*

$\mathsf{B}_{\mathsf{Com}}$*'s input: an ℓ-element set $\mathcal{Y} \subseteq \mathcal{S}$.*

Public randomness: (description of a) function $f \in \mathcal{F}_{\mathcal{S}}$.

Operation:

$\mathsf{A}_{\mathsf{Com}}$ *and* $\mathsf{B}_{\mathsf{Com}}$ *interact in an execution* $(\mathsf{A}(\mathcal{X}, f(\mathcal{X})), \mathsf{B}(\mathcal{Y}, f(\mathcal{Y})))$ *of* Π*, taking the roles of* A *and* B *respectively:* $\mathsf{A}_{\mathsf{Com}}$ *acts as* A *with queries $\mathcal{X}$ and answers $f(\mathcal{X})$, and* $\mathsf{B}_{\mathsf{Com}}$ *as* B *with queries $\mathcal{Y}$ and answers $f(\mathcal{Y})$. At the end of the interaction,* $\mathsf{A}_{\mathsf{Com}}$ *and* $\mathsf{B}_{\mathsf{Com}}$ *output the outputs of* A *and* B *respectively.*

We compare the above protocol to a protocol that emulates a run of Π *without* using the shared oracle; each party sets the answers of the oracle using its *private* randomness, and acts accordingly.

The private-oracle emulation. In this protocol, each party sample a random function using private randomness. The parties then interact according to Λ_{Com}, while treating the private function as the shared oracle.

Protocol 5.7 ($\Lambda_{\mathsf{Dist}} = (\mathsf{A}_{\mathsf{Dist}}, \mathsf{B}_{\mathsf{Dist}})$)**.**

$\mathsf{A}_{\mathsf{Com}}$*'s input: an ℓ-element set $\mathcal{X} \subseteq \mathcal{S}$.*

$\mathsf{B}_{\mathsf{Com}}$*'s input: an ℓ-element set $\mathcal{Y} \subseteq \mathcal{S}$.*

Public randomness: none.

Operation:

1. $\mathsf{A}_{\mathsf{Dist}}$ *samples* $g \overset{R}{\leftarrow} \mathcal{F}_{\mathcal{S}}$.
2. $\mathsf{B}_{\mathsf{Dist}}$ *samples* $f \overset{R}{\leftarrow} \mathcal{F}_{\mathcal{S}}$.
3. $\mathsf{A}_{\mathsf{Dist}}$ *and* $\mathsf{B}_{\mathsf{Dist}}$ *interact in protocol* $(\mathsf{A}(\mathcal{X}, g(\mathcal{X})), \mathsf{B}(\mathcal{Y}, f(\mathcal{Y})))$ *taking the roles of* A *and* B *respectively:* $\mathsf{A}_{\mathsf{Dist}}$ *acts as* A *with queries $\mathcal{X}$ and answers $g(\mathcal{X})$, and* $\mathsf{B}_{\mathsf{Dist}}$ *as* B *with queries $\mathcal{Y}$ and answers $f(\mathcal{Y})$. At the end of the interaction,* $\mathsf{A}_{\mathsf{Dist}}$ *and* $\mathsf{B}_{\mathsf{Dist}}$ *output the outputs of* A *and* B *respectively.*

Let (X, Y) be distributed as the queries of parties A and B respectively in Π (that is, uniform sets in $\mathcal{S}$ of size ℓ), and recall that $\Lambda_{\mathsf{Dist}}(\mathrm{X}, \mathrm{Y})$ and $\Lambda_{\mathsf{Com}}(\mathrm{X}, \mathrm{Y})$ denote the parties' joint view in a random execution of Λ_{Dist} and Λ_{Com} respectively, with inputs drawn from (X, Y). We first show that $\Lambda_{\mathsf{Dist}}(\mathrm{X}, \mathrm{Y})$ is far from $\Lambda_{\mathsf{Com}}(\mathrm{X}, \mathrm{Y})$. Indeed, since Λ_{Dist} is a no-oracle protocol (and has no common randomness), Fact 4.2 yields that there is an algorithm E such that

$$\Pr_{v \overset{R}{\leftarrow} \Lambda_{\mathsf{Dist}}(\mathrm{X},\mathrm{Y})}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right] = \Pr_{v \overset{R}{\leftarrow} \Lambda_{\mathsf{Dist}}(\mathrm{X},\mathrm{Y})}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Dist}}}(v) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right] \tag{10}$$

In contrast, since Λ_{Com} is an emulation of the protocol Π with a random oracle, the secrecy of Π and the fact that E sees not the common randomness, yields that

$$\Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right] = \Pr_{f \xleftarrow{R} F, v \xleftarrow{R} \Pi^f}\left[\mathsf{E}^f(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}}(v)\right] \leq \gamma \tag{11}$$

Finally, since the joint distribution of the outputs of the parties in Λ_{Com} is exactly as in Π, it holds that

$$\Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}(v) = \mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right] = \Pr_{f \xleftarrow{R} F, v \xleftarrow{R} \Pi^f}\left[\mathrm{out}^{\mathsf{A}}(v) = \mathrm{out}^{\mathsf{B}}(v)\right] \geq 1 - \alpha \tag{12}$$

It follows that at least one of the two equations below holds:

$$\begin{aligned}&\text{Agreement gap:}\\&\Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}(v) = \mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right] - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Dist}}(X,Y)}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Dist}}}(v) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right]\\&\geq (1-\alpha-\gamma)/2\end{aligned} \tag{13}$$

$$\begin{aligned}&\text{Secrecy gap:}\\&\Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Dist}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right] - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}}(v)\right]\\&\geq (1-\alpha-\gamma)/2\end{aligned} \tag{14}$$

Namely, wither Λ_{Com} is significantly more accurate than Λ_{Dist}, or Λ_{Com} is significantly more secure than protocol Λ_{Dist} (or both). We claim that without loss of generality one can assume that Equation (13) holds (i.e., there is agreement gap). Assuming otherwise (i.e., Equation (14) holds), we build a new protocol with inaccurate no-oracle emulation, and then continue the proof assuming Equation (13) holds.

Consider protocols $\Lambda'_{\mathsf{Com}} = (\mathsf{A}'_{\mathsf{Com}}, \mathsf{B}'_{\mathsf{Com}})$ and $\Lambda'_{\mathsf{Dist}} = (\mathsf{A}'_{\mathsf{Dist}}, \mathsf{B}'_{\mathsf{Dist}})$, in which the parties interact according to Λ_{Com} and Λ_{Dist} respectively, but parties $\mathsf{B}'_{\mathsf{Com}}$ and $\mathsf{B}'_{\mathsf{Dist}}$ output $\neg\mathsf{E}(\mathsf{trans})$. By the secrecy gap assumption,

$$\begin{aligned}&\Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Dist}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right] - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right]\\&\geq (1-\alpha-\gamma)/2\end{aligned} \tag{15}$$

Hence,

$$\begin{aligned}&\Pr_{v \xleftarrow{R} \Lambda'_{\mathsf{Com}}(X,Y)}\left[\mathrm{out}^{\mathsf{B}'_{\mathsf{Com}}}(v) = \mathrm{out}^{\mathsf{A}'_{\mathsf{Com}}}(v)\right] - \Pr_{v \xleftarrow{R} \Lambda'_{\mathsf{Dist}}(X,Y)}\left[\mathrm{out}^{\mathsf{B}'_{\mathsf{Dist}}}(v) = \mathrm{out}^{\mathsf{A}'_{\mathsf{Dist}}}(v)\right]\\
&=(1 - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}(v)\right])\\
&\quad - (1 - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Dist}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right])\\
&= \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Dist}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right] - \Pr_{v \xleftarrow{R} \Lambda_{\mathsf{Com}}(X,Y)}\left[\mathsf{E}(\mathsf{trans}(v)) = \mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right]\\
&\geq (1-\alpha-\gamma)/2.\end{aligned}$$

That is, protocol Λ'_{Dist} is less accurate than Λ'_{Com} by $(1-\alpha-\gamma)/2$. Namely, we are exactly in the same situation as if Equation (13) holds, but with respect to protocols Λ'_{Dist} and Λ'_{Com}. From hereafter, we assume for concreteness that Equation (13) holds with respect to the original protocols Λ_{Com} and Λ_{Dist}.

5.1 From Agreement Gap to Set Disjointness

Since, by assumption, Λ_{Dist} is less accurate than Λ_{Com} in (i.e., Equation (13) holds), it is less accurate for some specific intersection size; when the parties have *no* common query, Λ_{Dist} behaves just like Λ_{Com}, and thus Λ_{Dist} is (perfectly) accurate in this case. We exploit this observation to show that the accuracy difference between the protocols enables us to distinguish between disjoint inputs and intersecting inputs, yielding a protocol that solves set intersection over certain distributions.

For $z \in \{\mathsf{Com}, \mathsf{Dist}\}$ and a joint view $v = (\mathcal{X}, r_{\mathsf{A}}, \mathcal{Y}, r_{\mathsf{B}}, r_{\mathsf{P}}) \in \mathrm{Supp}(\Lambda_z)$, let $x(v) = \mathcal{X}$ and $y(v) = \mathcal{Y}$. For $i \in [\ell]$, let $\mathrm{Acc}_z(i)$ be the accuracy of Λ_z on inputs with intersection size i. Namely,

$$\mathrm{Acc}_z(i) := \Pr_{v \xleftarrow{\mathrm{R}} \Lambda_z(\mathrm{X},\mathrm{Y})} \left[\mathrm{out}^{\mathsf{B}_z}(v) = \mathrm{out}^{\mathsf{A}_z}(v) \mid |x(v) \cap y(v)| = i \right].$$

Let $\mathrm{AccGap}(i)$ be the accuracy advantage of Λ_{Com} over Λ_{Dist} on inputs with intersection size i. That is,

$$\mathrm{AccGap}(i) := \mathrm{Acc}_{\mathsf{Com}}(i) - \mathrm{Acc}_{\mathsf{Dist}}(i)$$

A key observation is that for some intersection size, protocol Λ_{Com} is more accurate than Λ_{Dist}.

Claim 5.8. $\exists d < \frac{4\ell^2}{|\mathcal{S}|(1-\alpha-\gamma)}$ *such that* $\mathrm{AccGap}(d) \geq (1-\alpha-\gamma)/4$.

The proof for this claim appears in Appendix B.

In contrast to the above claim, if the inputs are *disjoint* then there is no agreement gap. That is, we have the following fact.

Claim 5.9. $\mathrm{AccGap}(0) = 0$.

Proof. It is clear that for $(\mathrm{F}, \mathrm{G}) \xleftarrow{\mathrm{R}} \mathcal{F}_{\mathcal{S}}^2$ and pair of sets $\mathcal{X} \subseteq \mathcal{S}, \mathcal{Y} \subseteq \mathcal{S}$ with $\mathcal{X} \cap \mathcal{Y} = \emptyset$, the distributions of $(\mathcal{X}, \mathcal{Y}, \mathrm{F}(\mathcal{X}), \mathrm{F}(\mathcal{Y}))$ and of $(\mathcal{X}, \mathcal{Y}, \mathrm{F}(\mathcal{X}), \mathrm{G}(\mathcal{Y}))$ are the same. It follows that the distribution $\Lambda_{\mathsf{Dist}}|_{x \cap y = \emptyset}$ is identical to that of $\Lambda_{\mathsf{Com}}|_{x \cap y = \emptyset}$, meaning that the protocols act the same. □

Combining Claim 5.8 and Claim 5.9 yields there exists some constant $0 < c \leq d$ such that

$$\mathrm{AccGap}(c) - \mathrm{AccGap}(c-1) \geq \mathrm{AccGap}(d)/d \geq \frac{(1-\alpha-\gamma)^2 \cdot |\mathcal{S}|}{16\ell^2} \tag{16}$$

Hence,

$$\begin{aligned} \frac{(1-\alpha-\gamma)^2 |\mathcal{S}|}{16\ell^2} &\leq \mathrm{AccGap}(c) - \mathrm{AccGap}(c-1) \qquad (17) \\ &= (\mathrm{Acc}_{\mathsf{Com}}(c) - \mathrm{Acc}_{\mathsf{Dist}}(c)) - (\mathrm{Acc}_{\mathsf{Com}}(c-1) - \mathrm{Acc}_{\mathsf{Dist}}(c-1)) \\ &= \mathrm{Acc}_{\mathsf{Com}}(c) - \mathrm{Acc}_{\mathsf{Com}}(c-1) + \mathrm{Acc}_{\mathsf{Dist}}(c-1) - \mathrm{Acc}_{\mathsf{Dist}}(c). \end{aligned}$$

Therefore, either

$$\mathrm{Acc}_{\mathsf{Com}}(c) - \mathrm{Acc}_{\mathsf{Com}}(c-1) \geq \frac{(1-\alpha-\gamma)^2 |\mathcal{S}|}{32\ell^2}, \tag{18}$$

or

$$\mathrm{Acc}_{\mathsf{Dist}}(c-1) - \mathrm{Acc}_{\mathsf{Dist}}(c) \geq \frac{(1-\alpha-\gamma)^2 |\mathcal{S}|}{32\ell^2}. \tag{19}$$

Namely, at least, one of protocols Λ_{Com} and Λ_{Dist} can be used to distinguish between input of intersection of size c and input of $c-1$ with good probability. We conclude the proof showing how to use this ability to solve set-disjointness on the hard distribution D_ℓ.

The set intersection protocol. In the following we assume for concreteness that Equation (18) holds, where the proof assuming Equation (19) holds follows analogously by replacing Λ_{Com} with Λ_{Dist}. Consider the following protocol for solving set intersection (in the standard communication complexity model). For simplicity, we assume that ℓ is a multiple of 4, and that $\mathcal{S} = \{1, \ldots, |\mathcal{S}|\}$.

Protocol 5.10 ($\Lambda_{\mathsf{Set}} = (\mathsf{A}_{\mathsf{Set}}, \mathsf{B}_{\mathsf{Set}})$)**.**

Parameter: $k \in N$.

$\mathsf{A}_{\mathsf{Set}}$*'s input: an* $\ell/4$*-element set* $\mathcal{X} \subseteq [\ell]$.

$\mathsf{B}_{\mathsf{Set}}$*'s input: an* $\ell/4$*-element set* $\mathcal{Y} \subseteq [\ell]$.

Public randomness: (description of) k *permutations* $\sigma_1, ..., \sigma_n$ *over* $\mathcal{S}$.

Operation:

1. $\mathsf{A}_{\mathsf{Set}}$ *sets* $\mathcal{X}' = \mathcal{X} \cup \{\ell+1, \ell+2, ..., \ell+c-1\} \cup \{2\ell, 2\ell+1, ..., 3\ell - \ell/4 - c + 1\}$ *and* $\mathsf{B}_{\mathsf{Set}}$ *sets* $\mathcal{Y}' = \mathcal{Y} \cup \{\ell+1, \ell+2, ..., \ell+c-1\} \cup \{3\ell, 3\ell+1, ..., 4\ell - \ell/4 - c + 1\}$.
2. $\mathsf{A}_{\mathsf{Set}}$ *sets* $\mathsf{counter} = 0$.
3. *For* $j = 1$ *to* k*:*
   (a) $\mathsf{A}_{\mathsf{Set}}$ *and* $\mathsf{B}_{\mathsf{Set}}$ *interact in random execution of* $(\mathsf{A}_{\mathsf{Com}}(\sigma_j(\mathcal{X}')), \mathsf{B}_{\mathsf{Com}}(\sigma_j(\mathcal{Y}')))$*, with fresh randomness, taking the roles of* $\mathsf{A}_{\mathsf{Com}}$ *and* $\mathsf{B}_{\mathsf{Com}}$ *respectively. Let* $\mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}$ *and* $\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}$ *be the parties outputs in the execution.*
   (b) $\mathsf{B}_{\mathsf{Set}}$ *sends* $\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}$ *to* $\mathsf{A}_{\mathsf{Set}}$.
   (c) *If* $\mathrm{out}^{\mathsf{A}_{\mathsf{Com}}} = \mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}$, $\mathsf{A}_{\mathsf{Set}}$ *increases* $\mathsf{counter}$ *by one.*
4. $\mathsf{A}_{\mathsf{Set}}$ *informs* $\mathsf{B}_{\mathsf{Set}}$ *whether* $\mathsf{counter}/k > (\mathrm{Acc}_{\mathsf{Com}}(c) + \mathrm{Acc}_{\mathsf{Com}}(c-1))/2$. *If positive, both parties output zero; otherwise, they output one.*

In the following we analyze the success probability and communication complexity of protocol Λ_{Set} for $k = k^* := \frac{2^{13}\ell^4 \log 1/\epsilon}{|\mathcal{S}|^2(1-\alpha-\gamma)^4}$.

Success probability of Λ_{Set}. We show that for $k = k^*$ it holds that

$$\Pr_{\substack{(\mathcal{X},\mathcal{Y}) \overset{\mathrm{R}}{\leftarrow} D_\ell \\ r_{\mathsf{A}} \overset{\mathrm{R}}{\leftarrow} \{0,1\}^* \\ r_{\mathsf{B}} \overset{\mathrm{R}}{\leftarrow} \{0,1\}^* \\ r_p \overset{\mathrm{R}}{\leftarrow} \{0,1\}^*}} [(\mathsf{A}_{\mathsf{Set}}(\mathcal{X}; r_{\mathsf{A}}), \mathsf{B}_{\mathsf{Set}}(\mathcal{Y}; r_{\mathsf{B}}))(r_{\mathsf{P}}) = (\mathcal{X} \cap \mathcal{Y} = \emptyset, \mathcal{X} \cap \mathcal{Y} = \emptyset)] \geq 1 - \varepsilon \tag{20}$$

We prove that Equation (20) holds for any fixed $(\mathcal{X}, \mathcal{Y}) \in \mathrm{Supp}(D_\ell)$. Fix such a pair $(\mathcal{X}, \mathcal{Y})$, and assume without loss of generality that $|\mathcal{S}| > 3\ell/(1 - \alpha - \gamma)$ (as otherwise the proof of Theorem 5.4 is immediate). By this assumption, it holds that $c \leq d \leq 3/4\ell$. By construction, the sets $\mathcal{X}'$ and $\mathcal{Y}'$ set by the parties in Step 1 of the protocol, are both of size ℓ. Since, by definition, $(\mathcal{X}, \mathcal{Y})$ have at most one shared element, it holds that

$$|\mathcal{X}' \cap \mathcal{Y}'| = \begin{cases} c & \mathcal{X} \cap \mathcal{Y} \neq \emptyset \\ c - 1, & \text{otherwise.} \end{cases} \tag{21}$$

It follows that if $|\mathcal{X}' \cap \mathcal{Y}'| = c$ and $\mathsf{counter}/k > (\mathrm{Acc}_{\mathsf{Com}}\,(c) + \mathrm{Acc}_{\mathsf{Com}}\,(c-1))/2$, then the protocol outputs the right answer. Similarly, this is the case if $|\mathcal{X}' \cap \mathcal{Y}'| = c - 1$ and $\mathsf{counter}/k < (\mathrm{Acc}_{\mathsf{Com}}\,(c) + \mathrm{Acc}_{\mathsf{Com}}\,(c-1))/2$. Given these observations concerning the protocol correctness, we conclude the proof by bounding the probability that $\mathsf{counter}/k$ is far from $\mathrm{Acc}_{\mathsf{Com}}\,(|\mathcal{X} \cap \mathcal{Y}|)$.

Claim 5.11. *Let* $\mathsf{Counter}$ *be the value of* $\mathsf{counter}$ *in a random execution of* Λ_{Set} *on inputs* $(\mathcal{X}, \mathcal{Y})$*. Then for every* $\epsilon > 0$, $\delta > 0$ *and* $k = \lceil \log(1/\epsilon)/2\delta^2 \rceil$*, it holds that*
$\Pr[\mathsf{Counter}/k - \mathrm{Acc}_{\mathsf{Com}}\,(|\mathcal{X}' \cap \mathcal{Y}'|) > \delta] < \epsilon$ *and*
$\Pr[\mathrm{Acc}_{\mathsf{Com}}\,(|\mathcal{X}' \cap \mathcal{Y}'| - \mathsf{Counter}/k) > \delta] < \epsilon$.

Proof. Since the parties randomly permute their inputs, for every $j \in [k]$ it holds that $\sigma_j(\mathcal{X}')$ and $\sigma_j(\mathcal{Y}')$ are random sets drawn (independently of other iteration) from the distribution $(\mathrm{X}, \mathrm{Y})|_{|\mathrm{X} \cap \mathrm{Y}| = |\mathcal{X}' \cap \mathcal{Y}'|}$. Therefore, the probability of the parties to have the same output in each run of Π is exactly $\mathrm{Acc}_{\mathsf{Com}}\,(|\mathcal{X} \cap \mathcal{Y}|)$. The stated bound thus follows by by Hoffeding inequality (Fact 4.14). □

Let $\delta = \frac{(1-\alpha-\gamma)^2|\mathcal{S}|}{2^7\ell^2}$. By Equation (18), it holds that
$\delta < (\mathrm{Acc}_{\mathsf{Com}}\,(c) - \mathrm{Acc}_{\mathsf{Com}}\,(c-1))/2$. Hence, Claim 5.11 yields that protocol Λ_{Set} error probability on the input pair $(\mathcal{X}, \mathcal{Y})$ for parameter $k = k^*$ is less than ϵ, and Equation (20) follows.

Communication complexity. In each iteration of protocol Λ_{Set}, the parties run protocol Λ_{Com} and send one additional bit. Since $\mathrm{CC}(\Lambda_{\mathsf{Com}}) = \mathrm{CC}(\Pi)$, for $k = k^*$ we get that

$$\mathrm{CC}(\Lambda_{\mathsf{Set}}) \leq k(\mathrm{CC}(\Lambda_{\mathsf{Com}}) + 1) + 1 \leq 4k \cdot \mathrm{CC}(\Pi) = \frac{2^{15}\ell^4 \log 1/\epsilon}{|\mathcal{S}|^2 (1 - \alpha - \gamma)^4} \cdot \mathrm{CC}(\Pi) \tag{22}$$

Proving Theorem 5.4. The proof of Theorem 5.4 immediately follows that above observations.

Proof of Theorem 5.4. Fix $k = k^* = \frac{2^{13}\ell^4 \log 1/\epsilon}{|\mathcal{S}|^2(1-\alpha-\gamma)^4}$. Equation (20) yields that protocol Λ_{Set} solves set-disjointness over D_ℓ with error ϵ, and Equation (22) yields that $\mathrm{CC}(\Lambda_{\mathsf{Set}}) \leq \frac{2^{15}\ell^4 \log 1/\epsilon}{|\mathcal{S}|^2(1-\alpha-\gamma)^4} \cdot \mathrm{CC}(\Pi)$. □

6 Two-Messages Non-Adaptive Protocols

In this section we prove a lower bound on the communication complexity of any non-adaptive key agreement protocol that uses only two messages. We consider protocols with respect to the family $\mathcal{F}_n$ of all functions from $\{0,1\}^n$ to $\{0,1\}^n$.

Theorem 6.1 (Main theorem for two-message, non-adaptive protocols)**.** *For any $n \in \mathbb{N}$, the communication complexity of a two-message, non-adaptive, ℓ-query (q, α, γ)-key-agreement protocol relative to $\mathcal{F}_n$ is at least*

$$\frac{(1-\alpha-\gamma)^2 q}{50^2 \ell} - 6.$$

Fix a two-message, non-adaptive, ℓ-query protocol $\Pi = (\mathsf{A}, \mathsf{B})$. Each execution of the protocol specifies the following:

- X and Y, the queries made by A and B, respectively;
- $\mathrm{M}_1, \mathrm{M}_2$, the messages sent in the two rounds;
- $\mathrm{out}^{\mathsf{A}}$ and $\mathrm{out}^{\mathsf{B}}$, the outputs of the parties.

Where M_1 is a function (not necessarily deterministic) of X and $\mathrm{F}(\mathrm{X})$, and M_2 is a function of $\mathrm{Y}, \mathrm{F}(\mathrm{Y})$ and M_1. We define an eavesdropper $\mathrm{Eve} = \mathrm{Eve}_\delta$, where δ is a parameter we will specify later, and show that Eve violates the secrecy of Π if $\mathrm{CC}(\Pi)$ is too small. Loosely speaking, the eavesdropper, which is described below, queries all "heavy" queries and outputs what B would output given these queries.

Algorithm 6.2 (The eavesdropper Eve)**.**

Oracle: $f \in \mathcal{F}_n$.

Parameter: $\delta > 0$.

Operation: *Let $\boldsymbol{m} = m_1, m_2$ be the messages exchanged in the protocol.*

1. *Query f on all elements in $\mathcal{E}_0 \cup \mathcal{E}_1$, defined as*

$$\mathcal{E}_0 = \{q \in \{0,1\}^n : \Pr[q \in X \cup Y] \geq \delta\}.$$

and

$$\mathcal{E}_1 = \left\{q \in \{0,1\}^n : \Pr\left[q \in X \cup Y \;\middle|\; M_1 = m_1, \mathrm{F}\big|_{\mathcal{E}_0} = f\big|_{\mathcal{E}_0}\right] \geq \delta\right\}.$$

2. *Sample and output*

$$k \overset{R}{\leftarrow} \mathrm{out}^{\mathsf{B}}\big|_{\mathrm{M}_{\leq 2} = m_{\leq 2}, \mathrm{F}|_{\mathcal{E}_0 \cup \mathcal{E}_1} = f|_{\mathcal{E}_0 \cup \mathcal{E}_1}}.$$

It does not matter if Eve asks her queries *during* the protocol's run or *afterwards*. It is convenient to assume that Eve asks the queries $\mathcal{E}_{i-1}$ after observing $\mathrm{M}_{\leq i-1}$ and before the next message is sent. In particular, $\mathcal{E}_0$ denotes the queries that are heavy *before* the messages are sent. These queries are a function of Π itself.

6.1 Simplifying the Structure of the Protocol

For our lower bound it is convenient to assume that the protocol has two structural properties:

(1) There are no queries that are a-priori heavy, that is, $\mathcal{E}_0 = \emptyset$.

(2) The secret key chosen by the players is the first bit in B's last query; that is, if B's queries are $\mathrm{Y}_1, \ldots, \mathrm{Y}_s$, then the secret key is the first bit of Y_s.

We show that any key agreement protocol can be transformed into one that has these properties, with minor loss in the parameters. The proofs for the next two lemmas appears in Appendix B.

Eliminating the a priori heavy queries. First we show that if $\mathcal{E}_0 \neq \emptyset$, we can *fix* the answers to $\mathcal{E}_0$ in advance, eliminating the need for the players and for Eve to ask these queries.

Lemma 6.3. *Let* Π *be any* ℓ*-query* (q, α, γ)*-key-agreement protocol. Then there is an* ℓ*-query* $(q - |\mathcal{E}_0|, \alpha, \gamma)$*-protocol* Θ *with the same communication complexity as* Π*, such that* Θ *has no queries that are heavy a priori, that is, for each* $q \in \{0,1\}^n$*, and for any oracle* $f \in \mathcal{F}_n$*,*

$$\Pr_{\theta^f}\left[q \in \mathrm{X} \cup \mathrm{Y}\right] \leq \delta.$$

The key can be B's last query. Next we show that we can transform any protocol into one where the secret key is the first bit of B's last query.

Lemma 6.4. *Let* Π *be an* ℓ*-query* (q, α, γ)*-key-agreement protocol with two messages and communication complexity* C*. Then there is an* $(\ell+1)$*-query* (q, α, γ)*-protocol* Θ *with two messages and communication complexity* $C + 1$*, in which the secret key is the first bit of* $\mathrm{Y}_{\ell+1}$*.*

6.2 Proof of the Main Theorem

We are now ready to prove Theorem 6.1. Given a (q, α, γ)-protocol, we showed in the previous section that we can construct a $(q - |\mathcal{E}_0|, \alpha, \gamma)$-protocol with one extra query and one extra bit of communication, which has the two properties we need. Henceforth, we assume that the two structural properties hold.

The heart of the lower bound is the following lemma, which asserts that the eavesdropper Eve defined above is able to ask enough queries so that B has very little advantage over Eve when it comes to outputting a secret key shared with A.

Let Π^F_{Eve} denote the distribution of Eve's view under Π^F. Namely, it is the joint distribution of $(\mathrm{M}_1, \mathrm{M}_2, \mathrm{F}(\mathcal{E}_1))$. We use v_E to denote a view of Eve drawn from this distribution.

Lemma 6.5.

$$\mathop{\mathrm{E}}_{v_E \xleftarrow{R} \Pi^{\mathrm{F}}_{\mathrm{Eve}}}\left[I_{SD}\left(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|_{v_E}\right)\right] \leq 25\sqrt{\delta(\mathrm{CC}(\Pi) + 5)}. \tag{23}$$

simplicity of notation, here and below we use $\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{Y}|_{v_E}$ to denote $\big(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{Y}\big)|_{v_E}$ (we condition all the three random variables not just Y), and similarly in other cases. We prove Lemma 6.5 below, but let us first use it to prove Theorem 6.1.

Proof of Theorem 6.1. First, let us fix δ such that Eve does not ask more than q queries. Let[5] $\delta = 4\ell/q$. Since both A and B ask together at most 2ℓ queries,

$$2\ell \geq \mathop{\mathrm{E}}_{\Pi^{\mathrm{F}}}[|\mathrm{X} \cup \mathrm{Y}|] = \sum_{q \in \{0,1\}^n} \Pr_{\Pi^{\mathrm{F}}}[q \in \mathrm{X} \cup \mathrm{Y}].$$

Since every heavy-query contributes to the sum at least δ, the size[6] of $\mathcal{E}_0$ is at most $2\ell/\delta = q/2$. Similarly, for every m_1 and $f\big|_{\mathcal{E}_0}$,

$$2\ell \geq \sum_{q \in \{0,1\}^n} \Pr_{\Pi^{\mathrm{F}}}\left[q \in \mathrm{X} \cup \mathrm{Y} \mid \mathrm{M}_1 = m_1, \mathrm{F}\big|_{\mathcal{E}_0} = f\big|_{\mathcal{E}_0}\right].$$

So, the size of $\mathcal{E}_1$ is also at most $q/2$. Overall, Eve asks no more than q queries.

Now, recall that $\mathrm{out}^{\mathsf{B}}$ is assumed to be the first bit of B's last query. In particular, $\mathrm{out}^{\mathsf{B}}$ is a deterministic function of Y. From Equation (23) and Fact 4.11,

$$\begin{aligned}&\mathop{\mathrm{E}}_{v_E \xleftarrow{\mathrm{R}} \Pi^{\mathrm{F}}_{\mathrm{Eve}}}\left[I_{SD}\left(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{out}^{\mathsf{B}}|_{v_E}\right)\right] \\ &\qquad \leq 25\sqrt{\delta(\mathrm{CC}(\Pi)+5)}.\end{aligned}$$

A's output is a function of her view $(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_1, \mathrm{M}_2)$, so conditioned on $v_E = (\mathrm{M}_1, \mathrm{M}_2, \mathrm{F}(\mathcal{E}_1))$, it is a function of $(\mathrm{X}, \mathrm{F}(\mathrm{X}))$. Using the data processing inequality again, we obtain

$$\mathop{\mathrm{E}}_{v_E \xleftarrow{\mathrm{R}} \Pi^{\mathrm{F}}_{\mathrm{Eve}}}\left[I_{SD}\left(\mathrm{out}^{\mathsf{A}}; \mathrm{out}^{\mathsf{B}}|_{v_E}\right)\right] \leq 25\sqrt{\delta(\mathrm{CC}(\Pi)+5)}.$$

Eve samples her output $\mathrm{out}^{\mathrm{Eve}}$ from $\mathrm{out}^{\mathsf{B}}|v_E$. Therefore,

$$\begin{aligned}&\Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathsf{B}}\right] - \Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathrm{Eve}}\right] \\ &= \mathop{\mathrm{E}}_{v_E \xleftarrow{\mathrm{R}} \Pi^{\mathrm{F}}_{\mathrm{Eve}}}\left[\Pr_{\Pi^{\mathrm{F}}|v_E}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathsf{B}}\right] - \Pr_{\Pi^{\mathrm{F}}|v_E}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathrm{Eve}}\right]\right] \\ &\leq 25\sqrt{\delta(\mathrm{CC}(\Pi)+5)}.\end{aligned} \tag{24}$$

In words, Eve's probability of guessing A's output is close to B's when $\mathrm{CC}(\Pi)$ is small.

On the other hand, we know that Π is α-consistent and γ-secure, so Eve *cannot* have a success probability too close to B's: By the α-consistency of Π, we have $\Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathsf{B}}\right] \geq 1 - \alpha$. By the γ-secrecy, we have
$\Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathrm{Eve}}\right] \leq \gamma$. Together,

$$\Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathsf{B}}\right] - \Pr_{\Pi^{\mathrm{F}}}\left[\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathrm{Eve}}\right] \geq 1 - \alpha - \gamma. \tag{25}$$

Combining (24) and (25) we see that we must have

$$\mathrm{CC}(\Pi) \geq \frac{(1-\alpha-\gamma)^2}{25^2\delta} - 5.$$

□

[5] In general, for an r-message protocol, we would set $\delta = 2r\ell/q$.

[6] Recall that we assumed that $\mathcal{E}_0 = \emptyset$. This assumption caused a loss in parameters, so here we need to bound the size of $\mathcal{E}_0$.

6.3 Proving Lemma 6.5

We prove Lemma 6.5 by considering each message separately. We start with an informal exposition of the proof. The advantage the players obtain over Eve is encapsulated by the difference between

- what A and B learn about the intersection $\mathrm{X} \cap \mathrm{Y}$ of their query sets given the transcript *and their queries* X or Y; and
- what Eve knows about the intersection $\mathrm{X} \cap \mathrm{Y}$ given the transcript and *her* queries $\mathrm{F}(\mathcal{E}_1)$.

To bound this advantage, we argue that

I. After the first message (A's message), all the knowledge that B has about A's queries X comes from her first message M_1. Any advantage he has over Eve comes from what he has learned about the intersection $\mathrm{X} \cap \mathrm{Y}$ of their query sets. Because M_1 is short, B cannot learn too much about this intersection. From his point of view, the posterior distribution of the intersection given M_1 remains close to the prior (which is known to Eve).

   To establish this part of the argument we use the language of mutual information.

II. Similarly, after the second message (B's message), all the knowledge that A has gained about B's queries Y comes from M_2 *and* what B already learned about the intersection $\mathrm{X} \cap \mathrm{Y}$ from M_1. In particular, there is a small probability that after seeing M_1, B has learned too much about the intersection, and can use this knowledge to communicate with A securely (as Eve does not know the intersection).

   To deal with this low-probability bad event, we need to switch to the language of statistical distance, and use Lemma 6.6 below.

The following technical lemma is useful in the analysis of the second message, as it allows to ignore the knowledge B gained about the intersection in the first message. This lemma can be useful in other contexts as well. Its proof appears in Appendix B.

Lemma 6.6. *Let $\mathrm{A} = \mathrm{A}_1, \ldots, \mathrm{A}_n$, let $\mathrm{T} \subseteq [n]$ and let B be random variables. Let Z be a random variable taking values in the set $\mathcal{Z}$, and let $g : \mathcal{Z} \to \mathcal{P}([n])$ be a function mapping the domain of Z to subsets of $[n]$. Let*

$$\epsilon = \mathop{\mathrm{E}}_{z \overset{R}{\leftarrow} \mathrm{Z}} \left[\mathop{\mathrm{E}}_{t \overset{R}{\leftarrow} \mathrm{T}|_z} \left[\mathrm{I}(\mathrm{A}_t ; \mathrm{B} | \mathrm{A}_{g(z)}, z) \right] \right] \qquad \text{and}$$

$$\delta = \mathop{\mathrm{E}}_{z \overset{R}{\leftarrow} \mathrm{Z}} \left[I_{SD} \left(\mathrm{A}, \mathrm{B}; \mathrm{T}|_z \right) \right].$$

Then

$$\mathop{\mathrm{E}}_{z, a_{g(z)} \overset{R}{\leftarrow} \mathrm{Z}, \mathrm{A}_{g(z)}} \left[\mathrm{SD} \left(\left(\mathrm{A}_\mathrm{T}, \mathrm{T}, \mathrm{B}|_{z, a_{g(z)}} \right), \left((\mathrm{A}_\mathrm{T}, \mathrm{T}|_{z, a_{g(z)}}) \times \mathrm{B}|_{z, a_{g(z)}} \right) \right) \right] \leq 2\sqrt{\epsilon} + 2\delta.$$

Analyzing the first message.

We start by proving that in expectation, the first message does not create too much dependence between the players' views:

Claim 6.7. *The following statements hold after seeing* A*'s message:*

1. A*'s view remain independent of* B*'s queries:* $\mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|\mathrm{M}_1) = 0$.
2. *The same holds conditioned on Eve's queries:* $\mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|\mathrm{M}_1, \mathrm{F}(\mathcal{E}_1)) = 0$.
3. *Not much dependence is created between* B*'s view and* A*'s queries:* $\mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1) \le \delta|\mathrm{M}_1|$.

Proof for Claim 6.7. The proof of the first item:

$$\begin{aligned} 0 \le \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|\mathrm{M}_1) &\le \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_1; \mathrm{Y}) && \text{(Chain rule)} \\ &= \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}) && \text{(Since } \mathrm{M}_1 \text{ is a function of } \mathrm{X}, \mathrm{F}(\mathrm{X})\text{)} \\ &= 0. && \text{(Because } \mathrm{Y} \perp (\mathrm{X}, \mathrm{F}(\mathrm{X}))\text{)} \end{aligned}$$

The proof of the second item:

$$\begin{aligned} 0 \le \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|\mathrm{M}_1, \mathrm{F}(\mathcal{E}_1)) &\le \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_1, \mathrm{F}(\mathcal{E}_1); \mathrm{Y}) && \text{(Chain rule)} \\ &= \mathrm{I}(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{F}(\mathcal{E}_1); \mathrm{Y}) && \text{(Since } \mathrm{M}_1 \text{ is a function of } \mathrm{X}, \mathrm{F}(\mathrm{X})\text{)} \\ &\le \mathrm{I}(\mathrm{X}, \mathrm{F}; \mathrm{Y}) && \text{(Data processing)} \\ &= 0. && \text{(Because } \mathrm{Y} \perp (\mathrm{X}, \mathrm{F})\text{)} \end{aligned}$$

To prove the third item, we first show that all the "secret information" B has about X after seeing M_1 — that is, the dependence between his view and X given M_1 — comes from the intersection between A and B's sets.

Let $\mathrm{T} := \{i \,:\, \mathrm{X}_i \in \mathrm{Y}\}$ be the indexes of the intersection queries.

Claim 6.8. $\mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1) \le \mathrm{I}(\mathrm{M}_1; \mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T}, \mathrm{X})$.

Proof.

$$\begin{aligned} \mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1) &= \mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1) - \mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}) && \text{(Because } \mathrm{X} \perp (\mathrm{Y}, \mathrm{F}(\mathrm{Y}))\text{)} \\ &\le \mathrm{I}(\mathrm{M}_1; \mathrm{Y}, \mathrm{F}(\mathrm{Y})|\mathrm{X}) && \text{(Lemma 4.8)} \\ &\le \mathrm{I}(\mathrm{M}_1; \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T}), \mathrm{Y}, \mathrm{F}(\mathrm{Y})|\mathrm{X}) \\ &= \mathrm{I}(\mathrm{M}_1; \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{X}) + \mathrm{I}(\mathrm{M}_1; \mathrm{Y}, \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T})). && \text{(Chain rule)} \end{aligned}$$

The second term is 0: because M_1 is a function of $\mathrm{X}, \mathrm{F}(\mathrm{X})$, we have

$$\begin{aligned} &\mathrm{I}(\mathrm{M}_1; \mathrm{Y}, \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T})) \\ &\le \mathrm{I}(\mathrm{F}(\mathrm{X}); \mathrm{Y}, \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T})) && \text{(Data processing)} \\ &= \mathrm{I}(\mathrm{F}(\mathrm{X}); \mathrm{Y}|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T})) + \mathrm{I}(\mathrm{F}(\mathrm{X}); \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T}), \mathrm{Y}) && \text{(Chain rule)} \\ &\le \mathrm{I}(\mathrm{F}(\mathrm{X}), \mathrm{F}(\mathrm{X}_\mathrm{T}); \mathrm{Y}|\mathrm{X}, \mathrm{T}) + \mathrm{I}(\mathrm{F}(\mathrm{X}); \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T}), \mathrm{Y}) && \text{(Chain rule)} \\ &= 0 + \mathrm{I}(\mathrm{F}(\mathrm{X}); \mathrm{F}(\mathrm{Y})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T}), \mathrm{Y}) && \text{((}\mathrm{X}, \mathrm{Y}, \mathrm{T}) \perp \mathrm{F}\text{)} \\ &= \mathrm{I}(\mathrm{F}(\mathrm{X} \setminus \mathrm{X}_\mathrm{T}); \mathrm{F}(\mathrm{Y} \setminus \mathrm{X}_\mathrm{T})|\mathrm{X}, \mathrm{T}, \mathrm{F}(\mathrm{X}_\mathrm{T}), \mathrm{Y}) \\ &= 0. && \text{(Since } F \text{ is a random function and } (\mathrm{X} \setminus \mathrm{X}_\mathrm{T}) \cap (\mathrm{Y} \setminus \mathrm{X}_\mathrm{T}) = \emptyset\text{)} \end{aligned}$$

Bound the first term:

$$\begin{aligned}
&\mathrm{I}(\mathrm{M}_1;\mathrm{T},\mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{X}) \\
&= \mathrm{I}(\mathrm{M}_1;\mathrm{T}|\mathrm{X}) + \mathrm{I}(\mathrm{M}_1;\mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T},\mathrm{X}) && \text{(Chain rule)} \\
&\leq \mathrm{I}(\mathrm{M}_1;\mathrm{Y}|\mathrm{X}) + \mathrm{I}(\mathrm{M}_1;\mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T},\mathrm{X}) && \text{(Data processing: T is a function of Y given X)} \\
&= \mathrm{I}(\mathrm{M}_1;\mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T},\mathrm{X}). && (\mathrm{M}_1 \perp \mathrm{Y}|\mathrm{X})
\end{aligned}$$

□

Next, we bound the information M_1 conveys about $\mathrm{F}(\mathrm{X}_\mathrm{T})$, using the fact that every element in X is in the intersection only with small probability (less than δ). The proof of the claim is similar to the proof of Shearer's inequality and appears in Appendix B.

Claim 6.9. $\mathrm{I}(\mathrm{M}_1;\mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T},\mathrm{X}) \leq \delta|\mathrm{M}_1|$.

The proof of the third item is complete. □

Analyzing the second message.

We now want to show that the second message also does not create much dependence between A and B's views. As with Claim 6.7 for the first message, we first want to show that all the dependence between A's view and B's queries comes from B's message, and that this dependence goes through the intersection between A and B's queries and what the players learn about the intersection from the transcript. This is done by the next claim. Let

$$\mathrm{T}_1 := \{i \ : \ \mathrm{Y}_i \in \mathrm{X} \setminus \mathcal{E}_1\}\,.$$

In words, it is the set of the indices of B's queries in the intersection that were not queried by Eve. Recall that $\Pi^\mathrm{F}_\mathrm{Eve}$ is the distribution of Eve's view, which includes $\mathrm{M}_1, \mathrm{M}_2$ and $\mathrm{F}(\mathcal{E}_1)$. Let $\mathrm{B}_E = (\mathrm{M}_1, \mathrm{Y}, \mathrm{F}(\mathrm{Y} \cap \mathcal{E}_1))$.

Claim 6.10.

$$\mathop{\mathrm{E}}_{v_E \xleftarrow{R} \Pi^\mathrm{F}_\mathrm{Eve}} \left[I_{SD}\left(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|_{v_E}\right)\right] \leq 4 \mathop{\mathrm{E}}_{b_E \xleftarrow{R} \mathrm{B}_E} \left[I_{SD}\left(\mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1}); \mathrm{M}_2|_{b_E}\right)\right].$$

The proof for the claim appears in Appendix B.

Now we left to show that on average, B's message cannot convey too much information about the intersection queries and their answers, as we did in Claim 6.9 for the first message. Specifically, we want to bound

$$\mathop{\mathrm{E}}_{b_E \xleftarrow{\mathrm{R}} \mathrm{B}_E} \left[I_{SD}\left(\mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1}); \mathrm{M}_2|_{b_E}\right)\right].$$

It would be easier if B knew *nothing* about the intersection (i.e. M_2 was independent of T_1 given M_1). But this is not the case, as B can learn some info from A's message. However, from Claim 6.7, we know that he does not learn a lot, and his message does not strongly depend on the intersection. Formally,

Claim 6.11.

$$\mathop{\mathrm{E}}_{b_E \overset{R}{\leftarrow} \mathrm{B}_E} \left[I_{SD}\left(\mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1}); \mathrm{M}_2|_{b_E}\right)\right] \leq 6\sqrt{\delta(|\mathrm{M}_1| + |\mathrm{M}_2| + 5)}.$$

The two claims above complete the proof of Lemma 6.5.

Proof. By definition of B_E,

$$\mathop{\mathrm{E}}_{b_E \overset{\mathrm{R}}{\leftarrow} \mathrm{B}_E} I_{SD}\left(\mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1}); \mathrm{M}_2|_{b_E}\right) = \mathop{\mathrm{E}}_{b_E \overset{\mathrm{R}}{\leftarrow} \mathrm{B}_E} I_{SD}\left(\mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1}); \mathrm{M}_2|_{m_1, y, f(e_1 \cap y)}\right).$$

By Lemma 6.6, it is enough to show:

(1)

$$\mathop{\mathrm{E}}_{m_1, y \overset{\mathrm{R}}{\leftarrow} \mathrm{M}_1, \mathrm{Y}} \left[I_{SD}\left(\mathrm{F}(\mathrm{Y}), \mathrm{M}_2; \mathrm{T}_1|_{m_1, y}\right)\right] \leq 2\sqrt{\delta|\mathrm{M}_1|}.$$

(2) $\mathrm{E}_{m_1, y \overset{\mathrm{R}}{\leftarrow} \mathrm{M}_1, \mathrm{Y}} \left[\mathrm{E}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}_1|m_1, y} [\mathrm{I}(\mathrm{F}(y_t); \mathrm{M}_2|m_1, y, \mathrm{F}(e_1 \cap y))] \right] \leq \delta(|\mathrm{M}_1| + |\mathrm{M}_2| + 5).$

The proof of the first item is (which is similar to the analysis of the first message):

$$\begin{aligned}
&\mathop{\mathrm{E}}_{m_1, y \overset{\mathrm{R}}{\leftarrow} \mathrm{M}_1, \mathrm{Y}} \left[I_{SD}\left(\mathrm{F}(\mathrm{Y}), \mathrm{M}_2; \mathrm{T}_1|_{m_1, y}\right)\right] \\
&\leq 2\sqrt{\mathrm{I}(\mathrm{F}(\mathrm{Y}), \mathrm{M}_2; \mathrm{T}_1|\mathrm{M}_1, \mathrm{Y})} && \text{(Fact 4.7)} \\
&= 2\sqrt{\mathrm{I}(\mathrm{F}(\mathrm{Y}); \mathrm{T}_1|\mathrm{M}_1, \mathrm{Y})} && (\mathrm{M}_2 \text{ is a function of } \mathrm{Y}, \mathrm{F}(\mathrm{Y}), \mathrm{M}_1) \\
&\leq 2\sqrt{\mathrm{I}(\mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1, \mathrm{Y})} && (\mathrm{T}_1 \text{ is a function of } \mathrm{X}, \mathrm{Y} \text{ and } \mathrm{M}_1) \\
&\leq 2\sqrt{\mathrm{I}(\mathrm{Y}, \mathrm{F}(\mathrm{Y}); \mathrm{X}|\mathrm{M}_1)} && \text{(Chain rule)} \\
&\leq 2\sqrt{\delta|\mathrm{M}_1|}. && \text{(Claim 6.7)}
\end{aligned}$$

To bound the second item we use a similar argument to the proof of Claim 6.9. The proof is more complicated here, because when we condition on M_1 and on Eve's queries, the answers of the oracle F are no longer independent of each other (e.g., A could send the XOR of the answers to her queries). Nevertheless, because not much information was revealed about the oracle's answers, not much dependence is created between them. The proof consists of two steps. First, we show that this term is bounded by $\delta|\mathrm{M}_2|$, plus the dependency between the answers, created by the first message and Eve's queries (Claim 6.12). Next, we bound this dependency (Claim 6.13).

Claim 6.12.

$$\begin{aligned}
&\mathop{\mathrm{E}}_{m_1, y \overset{R}{\leftarrow} \mathrm{M}_1, \mathrm{Y}} \left[\mathop{\mathrm{E}}_{t \overset{R}{\leftarrow} \mathrm{T}_1|m_1, y} [\mathrm{I}(\mathrm{F}(y_t); \mathrm{M}_2|m_1, y, \mathrm{F}(e_1 \cap y))] \right] \\
&\leq \delta|\mathrm{M}_2| + \delta \mathop{\mathrm{E}}_{y \overset{R}{\leftarrow} \mathrm{Y}} \left[\sum_i \mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i})|\mathrm{M}_1, y, \mathrm{F}(\mathcal{E}_1 \cap y)) \right].
\end{aligned}$$

The proof for Claim 6.12 is similar to the proof of Claim 6.9 and appears in Appendix B.

Claim 6.13.

$$\mathop{\mathrm{E}}_{y \xleftarrow{R} \mathrm{Y}} \left[\sum_i \mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i}) | \mathrm{M}_1, y, \mathrm{F}(\mathcal{E}_1 \cap y)) \right] \leq |\mathrm{M}_1| + 5.$$

Proof. For every $m \in \mathrm{Supp}(\mathrm{M}_1)$, let $\mathcal{E}(m)$ be the set of queries Eve asks after seeing the message m. By Lemma 4.10 (recall that J_m is the indicator for the event $\mathrm{M} = m$),

$$\begin{aligned} & \mathop{\mathrm{E}}_{y \xleftarrow{\mathrm{R}} \mathrm{Y}} \left[\sum_i \mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i}) | \mathrm{M}_1, y, \mathrm{F}(\mathcal{E}_1 \cap y)) \right] \\ & \leq \mathop{\mathrm{E}}_{y \xleftarrow{\mathrm{R}} \mathrm{Y}} \sum_i \sum_{m \in \mathrm{M}_1} \big[\mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i}) | y, \mathrm{F}(\mathcal{E}(m) \cap y)) & \text{(Lemma 4.10)} \\ & \qquad + \mathrm{I}(\mathrm{F}(y_i); \mathrm{J}_m | y, \mathrm{F}(\mathcal{E}(m) \cap y), \mathrm{F}(y_{<i})) \big] \end{aligned}$$

For every m, y, i, by the structure of F, and since $\mathrm{F}(\mathcal{E}(m) \cap y)$ is a fixed set, we have $\mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i}) | y, \mathrm{F}(\mathcal{E}(m) \cap y)) = 0$. Thus,

$$\begin{aligned} & \leq \mathop{\mathrm{E}}_{y \xleftarrow{\mathrm{R}} \mathrm{Y}} \sum_i \sum_{m \in \mathrm{M}_1} \big[\mathrm{I}(\mathrm{F}(y_i); \mathrm{F}(y_{<i}) | y, \mathrm{F}(\mathcal{E}(m) \cap y)) \\ & \qquad + \mathrm{I}(\mathrm{F}(y_i); \mathrm{J}_m | y, \mathrm{F}(\mathcal{E}(m) \cap y), \mathrm{F}(y_{<i})) \big] \\ & = \mathop{\mathrm{E}}_{y \xleftarrow{\mathrm{R}} \mathrm{Y}} \left[\sum_i \sum_{m \in \mathrm{M}_1} \mathrm{I}(\mathrm{F}(y_i); \mathrm{J}_m | y, \mathrm{F}(\mathcal{E}(m) \cap y), \mathrm{F}(y_{<i})) \right] \\ & = \mathop{\mathrm{E}}_{y \xleftarrow{\mathrm{R}} \mathrm{Y}} \left[\sum_{m \in \mathrm{M}_1} \mathrm{I}(\mathrm{F}(y); \mathrm{J}_m | y, \mathrm{F}(\mathcal{E}(m) \cap y)) \right] & \text{(Chain rule)} \\ & \leq \sum_{m \in \mathrm{M}_1} \mathrm{H}(\mathrm{J}_m). & \text{(Fact 4.5)} \end{aligned}$$

There is at most one m' such that $\Pr[\mathrm{M}_1 = m'] \geq 1/2$, hence,

$$\begin{aligned} & \sum_{m \in \mathrm{M}_1} \mathrm{H}(\mathrm{J}_m) \\ & \leq 1 + \sum_{m \in \mathrm{M}_1} \Pr[\mathrm{M}_1 = m] \left(-\log\left(\Pr[\mathrm{M}_1 = m] \right) + 4 \right) & \text{(Lemma 4.9)} \\ & = \mathrm{H}(\mathrm{M}_1) + 5. \end{aligned}$$

□

The proof of Claim 6.11 is complete. □

6.4 Remarks

Adaptive Protocols. While we believe that the eavesdropper Eve we defined above should allow us to prove lower bounds for every non-adaptive protocol, Eve will not work for adaptive protocol, even if she can choose the sets adaptively as well. Protocol 6.14 is an example of a one-message protocol with only $O(\log(\ell))$ communication, but without any heavy query (for every $\delta > 1/\ell$). Specifically, Eve will not make any query, and can not, therefore, break the protocol. Notice, however, that every one-message protocol can be broken trivially by simulating B, so this protocol is not secure.

Protocol 6.14.

Parameters: n, $\ell = 2^{n/2}$

Common functions: $f, g : \{0,1\}^n \to \{0,1\}^n$

1. *A choses a random string $x \in \{0,1\}^n$ and queries $x, f(x), ..., f^{\ell-1}(x)$ and $g(f^{i-1}(x))$ for a random index $i \in [\ell]$.*
2. *B choses a random string $y \in \{0,1\}^n$ and queries $y, f(y), ..., f^{\ell-1}(y)$ and $g(y), ..., g(f^{\ell-1}(y))$.*
3. *A sends $M_1 = g(f^{i-1}(x))$ to B, and outputs $f^{i-1}(x)$.*
4. *If there is $j \in [\ell]$ so that $g(f^{j-1}(y)) = M_1$ then B outputs $f^{j-1}(y)$. Otherwise, B aborts.*

.

Constant Rounds Protocols We failed to continue the proof for multi-message protocol. The main reason is that we were not able to deal with the dependency caused by Eve's queries. In two-message protocol, Eve's only asks queries after the first message, which depends only on A's view. We show here that conditioning on Eve's view in this case, cannot add too much dependency between A and B. However, in protocols with more messages, the queries of Eve depend on the view of both sides, and conditioning on Eve's view can potentially make the dependency more significant.

Acknowledgement

We thank Yuval Ishai for challenging us with this intriguing question, and Omer Rotem for very useful discussions.

References

[1] M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 284–293. ACM, 1997. 1

[2] B. Barak and M. Mahmoody. Merkle puzzles are optimal - an $O(n^2)$-query attack on any key exchange from a random oracle. In *Advances in Cryptology – CRYPTO '09*, pages 374–390, 2009. 1, 2, 3

[3] D. J. Bernstein and T. Lange. ebacs: Ecrypt benchmarking of cryptographic systems. https://bench.cr.yp.to. accessed 15 May 2018. 1

[4] B. Chor and E. Kushilevitz. A zero-one law for boolean privacy. *SIAM Journal on Discrete Mathematics*, 4(1):36–47, 1991. 3

[5] W. Diffie and M. E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976. 1

[6] A. Ganor, G. Kol, and R. Raz. Exponential separation of information and communication for boolean functions. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 557–566. ACM, 2015. 11

[7] I. Haitner, J. J. Hoch, O. Reingold, and G. Segev. Finding collisions in interactive protocols - tight lower bounds on the round and communication complexities of statistically hiding commitments. *SIAM Journal on Computing*, 44(1):193–242, 2015. Preliminary version in *STOC'07*. 3

[8] I. Haitner, E. Omri, and H. Zarosim. Limits on the usefulness of random oracles. *Journal of Cryptology*, 29(2):283–335, 2016. 3

[9] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963. 15

[10] R. Impagliazzo and S. Rudich. Limits on the provable consequences of one-way permutations. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC)*, pages 44–61. ACM Press, 1989. 1, 2, 3

[11] M. Mahmoody, H. K. Maji, and M. Prabhakaran. Limits of random oracles in secure computation. *arXiv preprint arXiv:1205.3554*, 2012. 3

[12] R. J. McEliece. A public-key cryptosystem based on algebraic. *Coding Thv*, 4244:114–116, 1978. 1

[13] R. C. Merkle. Secure communications over insecure channels. In *SIMMONS: Secure Communications and Asymmetric Cryptosystems*, 1982. 1, 15

[14] R. C. Merkle. A digital signature based on a conventional encryption function. In *Advances in Cryptology – CRYPTO '87*, pages 369–378, 1987. 32

[15] M. O. Rabin. Digitalized signatures and public-key functions as intractable as factorization. Technical report, MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER SCIENCE, 1979. 1

[16] A. Rao and M. Sinha. Simplified separation of information and communication. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 22, pages 2–3, 2015. 11

[17] A. A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992. 3, 16

[18] R. L. Rivest, A. Shamir, and L. M. Adelman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978. 1

A Merkle's Puzzles

For completeness, we briefly describe here the Merkle Puzzles protocol [14]. Let $\mathcal{S}$ be a set of size ℓ^2, and $\mathcal{F}_{\mathcal{S}} = \left\{ f : \mathcal{S} \mapsto \{0,1\}^{2\log|\mathcal{S}|} \right\}$ be the family of all functions from $\mathcal{S}$ to binary strings of length $2\log|\mathcal{S}|$.

Protocol A.1 (Merkle's Puzzles protocol $\Pi = (\mathsf{A}, \mathsf{B})$)**.**

Oracle: $f \in \mathcal{F}_{\mathcal{S}}$.

Operation:

1. A *samples uniformly and independently* ℓ *elements* $x_1, ..., x_\ell \in \mathcal{S}$*, and sets* $a_1 = f(x_1), ..., a_\ell = f(x_\ell)$.

   B *samples uniformly and independently* ℓ *elements* $y_1, ..., y_\ell \in \mathcal{S}$*, and set* $b_1 = f(y_1), ..., b_\ell = f(y_\ell)$.

2. A *sends* $a_1, ..., a_\ell$ *to* B.

3. B *looks for indices* $i, j \in [\ell]$ *with* $a_i = b_i$. *If no such indices exists, it aborts.*

4. B *sends* i *to* A.

5. A *outputs* x_i *and* B *outputs* y_j.

Since each party samples $\ell = \sqrt{|\mathcal{S}|}$ uniform random element from $\mathcal{S}$, by the birthday paradox they have a common element (i.e., collision) with constant probability. By construction, the parties out the same collision, if such exists. On the other hand, from an attacker point of view the collision is a random element of $\mathcal{S}$, and therefore she cannot find it with good probability without querying a constant fraction of the element of $\mathcal{S}$, namely by making $\Theta(\ell^2)$ queries.

Note that Merkle Puzzles is non-adaptive, uniform-queries, two-message protocol with near linear communication, and therefore shows that our two lower bounds (Theorems 1.1 and 1.2) are tight.

B Missing Proofs

Proof of Lemma 4.8.

$$
\begin{aligned}
&\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{C},\mathrm{D}) - I(\mathrm{A};\mathrm{B}|\mathrm{C}) \\
&\qquad = \mathrm{H}(\mathrm{A}|\mathrm{C},\mathrm{D}) - \mathrm{H}(\mathrm{A}|\mathrm{B},\mathrm{C},\mathrm{D}) - [\mathrm{H}(\mathrm{A}|\mathrm{C}) - \mathrm{H}(\mathrm{A}|\mathrm{B},\mathrm{C})] \\
&\qquad = \mathrm{H}(\mathrm{A}|\mathrm{C},\mathrm{D}) - \mathrm{H}(\mathrm{A}|\mathrm{C}) - [\mathrm{H}(\mathrm{A}|\mathrm{B},\mathrm{C},\mathrm{D}) - \mathrm{H}(\mathrm{A}|\mathrm{B},\mathrm{C})] \\
&\qquad = \mathrm{I}(\mathrm{A};\mathrm{D}|\mathrm{C},\mathrm{B}) - \mathrm{I}(\mathrm{A};\mathrm{D}|\mathrm{C})
\end{aligned}
$$

The inequalities hold by the fact that mutual information is always positive. □

Proof of Lemma 4.9.

$$\begin{aligned}\mathrm{H}(\mathrm{J}) =& \Pr[\mathrm{J}=1]\log\frac{1}{\Pr[\mathrm{J}=1]} + \Pr[\mathrm{J}=0]\log\frac{1}{\Pr[\mathrm{J}=0]}\\ \leq& \Pr[\mathrm{J}=1]\log\frac{1}{\Pr[\mathrm{J}=1]} + \log\frac{1}{1-\Pr[\mathrm{J}=1]}\end{aligned}$$

Let $f(x) = log\frac{1}{1-x} - 4x$. We need to show that $f(x) \leq 0$ for all $0 \leq x \leq 1/2$. $f(0) = 0$, therefore it is enough to show that $f'(x) \leq 0$.

$$\begin{aligned}f'(x) =& \frac{1}{\ln 2}\frac{1}{1-x} - 4\\ \leq& 2\frac{1}{1-x} - 4 \leq 4 - 4 = 0 && (0 \leq x \leq 1/2)\end{aligned}$$

□

Proof of Lemma 4.10.

$$\begin{aligned}\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{M},\mathrm{E}_M) =& \sum_{m\in\mathrm{M}} \Pr[\mathrm{M}=m]\,\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{M}=m,\mathrm{E}_m)\\ =& \sum_{m\in\mathrm{M}} \Pr[\mathrm{J}_m=1]\,\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{J}_m=1,\mathrm{E}_m)\\ \leq& \sum_{m\in\mathrm{M}} \big[\Pr[\mathrm{J}_m=1]\,\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{J}_m=1,\mathrm{E}_m)\\ & \qquad\qquad && \text{(Because I is non-negative)}\\ & \qquad + \Pr[\mathrm{J}_m=0]\,\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{J}_m=0,\mathrm{E}_m)\big]\\ =& \sum_{m\in\mathrm{M}} \mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{J}_m,\mathrm{E}_m)\\ \leq& \sum_{m\in\mathrm{M}} \mathrm{I}(\mathrm{A},\mathrm{J}_m;\mathrm{B}|\mathrm{E}_m) && \text{(Chain rule)}\\ =& \sum_{m\in\mathrm{M}} \big[\mathrm{I}(\mathrm{A};\mathrm{B}|\mathrm{E}_m) + \mathrm{I}(\mathrm{J}_m;\mathrm{B}|\mathrm{E}_m,\mathrm{A})\big] && \text{(Chain rule)}\end{aligned}$$

□

Proof of Lemma 4.16.

$$\begin{aligned}& \mathop{\mathrm{E}}_{c\overset{\mathrm{R}}{\leftarrow}\mathrm{C}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{C}=c}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B}|_{\mathrm{C}=c})\right)\right]\\ \leq& \mathop{\mathrm{E}}_{c\overset{\mathrm{R}}{\leftarrow}\mathrm{C}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{C}=c}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B})\right)+\mathrm{SD}\left((\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B}|_{\mathrm{C}=c})\right)\right]\\ & && \text{(Triangle inequality)}\\ =& \mathop{\mathrm{E}}_{c\overset{\mathrm{R}}{\leftarrow}\mathrm{C}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{C}=c}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B})\right)+\mathrm{SD}\left((\mathrm{B}),(\mathrm{B}|_{\mathrm{C}=c})\right)\right] && \text{(Fact 4.13)}\\ \leq& \mathop{\mathrm{E}}_{c\overset{\mathrm{R}}{\leftarrow}\mathrm{C}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{C}=c}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B})\right)+\mathrm{SD}\left((\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B}),(\mathrm{A},\mathrm{B}|_{\mathrm{C}=c})\right)\right] && \text{(Data procesing)}\\ =& 2\mathop{\mathrm{E}}_{c\overset{\mathrm{R}}{\leftarrow}\mathrm{C}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{C}=c}),(\mathrm{A}|_{\mathrm{C}=c}\times\mathrm{B})\right)\right]\\ =& 2I_{SD}\left(\mathrm{A},\mathrm{C};\mathrm{B}\right)\end{aligned}$$

□

Proof of Lemma 4.17.

$$\begin{aligned}
&\mathrm{SD}\left((\mathrm{M}\times\mathrm{A}),(\mathrm{M},\mathrm{A})\right)\\
&\le \mathrm{SD}\left((\mathrm{M},\mathrm{B})\times(\mathrm{A}),(\mathrm{M},\mathrm{B},\mathrm{A})\right) && \text{(Data processing)}\\
&= \underset{b\overset{\mathrm{R}}{\leftarrow}B}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}),(\mathrm{A},\mathrm{M}|_{\mathrm{B}=b})\right)\right] && \text{(Fact 4.12)}\\
&\le \underset{b\overset{\mathrm{R}}{\leftarrow}B}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}),(\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b})\right)+\mathrm{SD}\left((\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b}),(\mathrm{A},\mathrm{M}|_{\mathrm{B}=b})\right)\right]\\
& && \text{(Triangle inequality)}\\
&= \underset{b\overset{\mathrm{R}}{\leftarrow}B}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{A}),(\mathrm{A}|_{\mathrm{B}=b})\right)+\mathrm{SD}\left((\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b}),(\mathrm{A},\mathrm{M}|_{\mathrm{B}=b})\right)\right] && \text{(Fact 4.13)}\\
&= \mathrm{SD}\left((\mathrm{A}\times\mathrm{B}),(\mathrm{A},\mathrm{B})\right)+\underset{b\overset{\mathrm{R}}{\leftarrow}\mathrm{B}}{\mathrm{E}}\ \mathrm{SD}\left((\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b}),(\mathrm{A},\mathrm{M}|_{\mathrm{B}=b})\right) && \text{(Fact 4.12)}
\end{aligned}$$

□

Proof of Lemma 4.18.

$$\begin{aligned}
&\underset{m\overset{\mathrm{R}}{\leftarrow}\mathrm{M}}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{B}|_{\mathrm{M}=m}),(\mathrm{A}|_{\mathrm{M}=m}\times\mathrm{B}|_{\mathrm{M}=m})\right)\right]\\
&= \underset{m,b\overset{\mathrm{R}}{\leftarrow}\mathrm{M},\mathrm{B}}{\mathrm{E}}\left[\left[\mathrm{SD}\left((\mathrm{A}|_{\mathrm{M}=m,\mathrm{B}=b}),(\mathrm{A}|_{\mathrm{M}=m})\right)\right]\right] && \text{(Fact 4.12)}\\
&\le \underset{m,b\overset{\mathrm{R}}{\leftarrow}\mathrm{M},\mathrm{B}}{\mathrm{E}}\Big[\mathrm{SD}\left((\mathrm{A}|_{\mathrm{M}=m,\mathrm{B}=b}),(\mathrm{A}|_{\mathrm{B}=b})\right)\\
& && \text{(Triangle inequality)}\\
&\quad+\mathrm{SD}\left((\mathrm{A}|_{\mathrm{B}=b}),(\mathrm{A})\right)+\mathrm{SD}\left((\mathrm{A}),(\mathrm{A}|_{\mathrm{M}=m})\right)\Big]\\
&= \underset{b\overset{\mathrm{R}}{\leftarrow}\mathrm{B}}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{M}|_{\mathrm{B}=b}),(\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b})\right)\right]\\
& && \text{(Fact 4.12)}\\
&\quad+\mathrm{SD}\left((\mathrm{A},\mathrm{B}),(\mathrm{B}\times\mathrm{A})\right)+\mathrm{SD}\left((\mathrm{M}\times\mathrm{A}),(\mathrm{A},\mathrm{M})\right)\\
&\le 2\underset{b\overset{\mathrm{R}}{\leftarrow}\mathrm{B}}{\mathrm{E}}\left[\mathrm{SD}\left((\mathrm{A},\mathrm{M}|_{\mathrm{B}=b}),(\mathrm{M}|_{\mathrm{B}=b}\times\mathrm{A}|_{\mathrm{B}=b})\right)\right]+2\mathrm{SD}\left((\mathrm{A},\mathrm{B}),(\mathrm{A}\times\mathrm{B})\right) && \text{(Lemma 4.17)}
\end{aligned}$$

□

Proof of Claim 5.8. Let $t := \mathrm{E}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[|\mathcal{X}\cap\mathcal{Y}|\right]$ be the expected intersection size. We show below that

$$\sum_{i=0}^{\lfloor 4t/(1-\alpha-\gamma)\rfloor}\ \underset{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}{\Pr}\left[|\mathcal{X}\cap\mathcal{Y}|=i\right]\cdot\mathrm{AccGap}\left(i\right)\ge(1-\alpha-\gamma)/4 \tag{26}$$

It will then follows that $\exists d\le 4t/(1-\alpha-\gamma)$ such that $\mathrm{AccGap}\left(d\right)\ge(1-\alpha-\gamma)/4$. We conclude the proof by showing that $t=\ell^2/\left|\mathcal{S}\right|$, and therefore $d\le 4\ell^2/\left|\mathcal{S}\right|(1-\alpha-\gamma)$. By linearity of expectation,

$$t=\underset{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}{\mathrm{E}}\left[|\mathcal{X}\cap\mathcal{Y}|\right]=\sum_{i=1}^{\ell}\underset{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}{\mathrm{E}}\left[\mathcal{X}_i\in\mathcal{Y}\right]=\ell^2/\left|\mathcal{S}\right|.$$

So it is left to prove Equation (26). We first show that the expected value of $\mathrm{AccGap}\,(i)$ is at least $(1-\alpha-\gamma)/2$.

$$
\begin{aligned}
&\mathop{\mathrm{E}}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[\mathrm{AccGap}\,(|\mathcal{X}\cap\mathcal{Y}|)\right] \qquad (27)\\
&= \mathop{\mathrm{E}}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[\mathop{\Pr}_{v\overset{\mathrm{R}}{\leftarrow}\Lambda_{\mathsf{Com}}(\mathrm{X},\mathrm{Y})}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}(v)=\mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\mid |x(v)\cap y(v)|=|\mathcal{X}\cap\mathcal{Y}|\right]\right]\\
&\quad - \mathop{\mathrm{E}}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[\mathop{\Pr}_{v\overset{\mathrm{R}}{\leftarrow}\Lambda_{\mathsf{Dist}}(\mathrm{X},\mathrm{Y})}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Dist}}}(v)=\mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\mid |x(v)\cap y(v)|=|\mathcal{X}\cap\mathcal{Y}|\right]\right]\\
&= \mathop{\Pr}_{v\overset{\mathrm{R}}{\leftarrow}\Lambda_{\mathsf{Com}}(\mathrm{X},\mathrm{Y})}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Com}}}(v)=\mathrm{out}^{\mathsf{A}_{\mathsf{Com}}}(v)\right] - \mathop{\Pr}_{v\overset{\mathrm{R}}{\leftarrow}\Lambda_{\mathsf{Dist}}(\mathrm{X},\mathrm{Y})}\left[\mathrm{out}^{\mathsf{B}_{\mathsf{Dist}}}(v)=\mathrm{out}^{\mathsf{A}_{\mathsf{Dist}}}(v)\right]\\
&\geq (1-\alpha-\gamma)/2.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\sum_{i=0}^{\lfloor 4t/(1-\alpha-\gamma)\rfloor} \mathop{\Pr}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[|\mathcal{X}\cap\mathcal{Y}|=i\right]\cdot\mathrm{AccGap}\,(i)\\
&= \mathop{\mathrm{E}}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[\mathrm{AccGap}\,(|\mathcal{X}\cap\mathcal{Y}|)\right]\\
&\quad - \sum_{i=\lfloor 4t/(1-\alpha-\gamma)\rfloor+1}^{\ell} \mathop{\Pr}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[|\mathcal{X}\cap\mathcal{Y}|=i\right]\cdot\mathrm{AccGap}\,(i)\\
&\geq (1-\alpha-\gamma)/2 - \sum_{i=\lfloor 4t/(1-\alpha-\gamma)\rfloor+1}^{\ell} \mathop{\Pr}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[|\mathcal{X}\cap\mathcal{Y}|=i\right]\cdot\mathrm{AccGap}\,(i) \qquad \text{(Equation (27))}\\
&\geq (1-\alpha-\gamma)/2 - \sum_{i=\lfloor 4t/(1-\alpha-\gamma)\rfloor+1}^{\ell} \mathop{\Pr}_{(\mathcal{X},\mathcal{Y})\overset{\mathrm{R}}{\leftarrow}(\mathrm{X},\mathrm{Y})}\left[|\mathcal{X}\cap\mathcal{Y}|=i\right] \qquad (\mathrm{AccGap}\,(i)\leq 1)\\
&\geq (1-\alpha-\gamma)/2 - \mathop{\Pr}_{\mathcal{X}\overset{\mathrm{R}}{\leftarrow}\mathrm{X},\mathcal{Y}\overset{\mathrm{R}}{\leftarrow}\mathrm{Y}}\left[|\mathcal{X}\cap\mathcal{Y}|\geq 4t/(1-\alpha-\gamma)\right]\\
&\geq (1-\alpha-\gamma)/4., \qquad \text{(Markov inequality)}
\end{aligned}
$$

and the the proof of the claim follows. □

Proof of Lemma 6.3. For a mapping $R:\mathcal{E}_0\to\{0,1\}^n$ representing the answers to the queries in $\mathcal{E}_0$, let

$$
\mathcal{F}^R=\left\{f\in\mathcal{F}_n\;:\;f\big|_{\mathcal{E}_0}=R\right\}.
$$

In words, it is the set of oracles whose answers on $\mathcal{E}_0$ agree with R.

We show that there is R so that the protocol $\Pi^{\mathcal{F}^R}$, where the answers to $\mathcal{E}_0$ are fixed to agree with R, is a $(q-|\mathcal{E}_0|,\alpha,\gamma)$-key agreement protocol. We then define Θ to be the simulation of $\Pi^{\mathcal{F}^R}$ where for each query in $\mathcal{E}_0$, instead of querying the oracle the players use the answer from R.

In Θ, the queries in $\mathcal{E}_0$ are never asked, so they are no longer heavy. Moreover, no new heavy queries are created, because the protocol is non-adaptive; the queries X,Y asked by the players do not change when we fix the answers in $\mathcal{E}_0$.

Now let us choose R. First, observe that consistency is maintained for *any* setting of R: for each $f \in \mathcal{F}_n$,

$$\Pr_{v \xleftarrow{\mathrm{R}} \Pi^f} \left[\mathrm{out}^{\mathsf{A}}(v) = \mathrm{out}^{\mathsf{B}}(v) \right] \geq 1 - \alpha.$$

In particular this holds for $f \in \mathcal{F}^R$ for any R.

As for secrecy, assume for the sake of contradiction that there is no R under which Π is $(q - |\mathcal{E}_0|, \gamma)$-secure with respect to $\mathcal{F}^R$; that is, for each $R : \mathcal{E}_0 \to \{0,1\}^n$ there exists an attacker Eve_R that asks $q - |\mathcal{E}_0|$ queries such that

$$\Pr_{f \xleftarrow{\mathrm{R}} \mathcal{F}^R, v \xleftarrow{\mathrm{R}} \Pi^f} \left[\mathrm{Eve}_R^f(\mathit{trans}(v)) = \mathrm{out}^{\mathsf{A}}(v) \right] \geq \gamma.$$

Define an attacker Eve that breaks the original protocol Π as follows: First, Eve queries $\mathcal{E}_0$; let R be the answers she receives. Next, Eve simply runs Eve_R. We have:

$$\Pr_{\mathrm{f} \xleftarrow{\mathrm{R}} \mathcal{F}_n, v \xleftarrow{\mathrm{R}} \Pi^f} \left[\mathrm{Eve}^f(\mathit{trans}(v)) = \mathrm{out}^{\mathsf{A}}(v) \right] \geq \gamma.$$

This contradicts the secrecy of Π.

□

Proof of Lemma 6.4. In Θ, the players execute the original protocol Π, but with the following changes:

- In the beginning of the protocol, B asks one additional query $\mathrm{Y}_{\ell+1}$. This query is chosen uniformly at random and independently of his other queries (and is not used by Π).

- A then sends her message M_1 just as she would under Π, and B computes his message M_2 under Π, and the secret key $\mathrm{out}^{\mathsf{B}}$ that he would output in Π.

- B sends A the message M_2, b, where $\mathrm{b} = \mathrm{out}^{\mathsf{B}} \oplus (\mathrm{Y}_{\ell+1})_1$ is an additional bit B appends to the message.

- B outputs $(\mathrm{Y}_{\ell+1})_1$ as his secret key.

- A computes $\mathrm{out}^{\mathsf{A}}$ as in Π, and outputs $\mathrm{out}^{\mathsf{A}} \oplus \mathrm{b}$.

Whenever $\mathrm{out}^{\mathsf{A}} = \mathrm{out}^{\mathsf{B}}$, A's output agrees with B's. The consistency of the new protocol, therefore, is the same as Π's.

For secrecy, let F be the random oracle, and assume there is $\mathrm{Eve}^{\mathrm{F}}$ that breaks the secrecy of Θ. Namely, $\mathrm{Eve}^{\mathrm{F}}$ can guess the output of A with probability at least γ. Note that $(\mathrm{Y}_{\ell+1})_1$ is a uniform random bit independent of $\mathrm{M}_1, \mathrm{M}_2$ and F. Thus, we can think that in Θ, B chooses the value of $\mathrm{out}^{\mathsf{B}} \oplus (\mathrm{Y}_{\ell+1})_1$ *after* M_2 was sent.

- Given a transcript M_1 and M_2, the eavesdropper $\widehat{\mathrm{Eve}}^{\mathrm{F}}$ chooses a uniform random bit b.

- $\widehat{\mathrm{Eve}}^{\mathrm{F}}$ runs $\mathrm{Eve}^{\mathrm{F}}(\mathrm{M}_1, \mathrm{M}_2, \mathrm{b})$. Let $\mathrm{out}^{\mathrm{Eve}}$ be $\mathrm{Eve}^{\mathrm{F}}$'s output.

- $\widehat{\mathrm{Eve}}^{\mathrm{F}}$ outputs $\mathrm{b} \oplus \mathrm{out}^{\mathrm{Eve}}$.

Since $\mathrm{M}_1, \mathrm{M}_2, \mathrm{b}$ are distributed exactly as in Θ, we have that $\widehat{\mathrm{Eve}}^{\mathrm{F}}$ breaks Π with the same probability $\mathrm{Eve}^{\mathrm{F}}$ does, and with the same number of queries. □

Proof of Lemma 6.6. For $z \in \mathrm{Z}$, let $(\mathrm{T}'|_z)$ be distributed as the marginal distribution of $(\mathrm{T}|_{Z=z})$. From the triangle inequality for statistical distance, we get:

$$
\begin{aligned}
&\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left(\mathrm{A}_{\mathrm{T}}, \mathrm{T}, \mathrm{B}|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}}, \mathrm{T})|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right)\right)\right] \\
\leq &\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}}, \mathrm{T}, \mathrm{B})|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}', \mathrm{B})|_{z,a_{g(z)}}\right)\right)\right] \\
&+ \mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}', \mathrm{B})|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}')|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right)\right)\right] \\
&+ \mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}')|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}}, \mathrm{T})|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right)\right)\right].
\end{aligned}
$$

We bound each term above separately: the first term is bounded by δ, because by Fact 4.12 and the data processing inequality, we have

$$
\begin{aligned}
&\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}}, \mathrm{T}, \mathrm{B})|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}', \mathrm{B})|_{z,a_{g(z)}}\right)\right)\right] \\
= &\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[\mathrm{SD}\left((\mathrm{A}_{\mathrm{T}}, \mathrm{A}_{g(z)}, \mathrm{T}, \mathrm{B})|_z\right), (\mathrm{A}_{\mathrm{T}'}, \mathrm{A}_{g(z)}, \mathrm{T}', \mathrm{B})|_z)\right)\right] \\
\leq &\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[\mathrm{SD}\left(((\mathrm{A}, \mathrm{T}, \mathrm{B})|_z), ((\mathrm{A}, \mathrm{T}', \mathrm{B})|_z)\right)\right] \\
= &\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[I_{SD}\left(\mathrm{A}, \mathrm{B}; \mathrm{T}|_z\right)\right] = \delta.
\end{aligned}
$$

Similarly, the third term is also bounded by δ, as by data processing,

$$
\begin{aligned}
&\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}')|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}}, \mathrm{T})|_{z,a_{g(z)}} \times \mathrm{B}|_{z,a_{g(z)}}\right)\right)\right] \\
= &\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}'}, \mathrm{T}')|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}}, \mathrm{T})|_{z,a_{g(z)}}\right)\right)\right] \\
= &\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[\mathrm{SD}\left(((\mathrm{A}_{\mathrm{T}'}, \mathrm{A}_{g(z)}, \mathrm{T}')|_z), ((\mathrm{A}_{\mathrm{T}}, \mathrm{A}_{g(z)}, \mathrm{T})|_z)\right)\right] \\
\leq &\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[I_{SD}\left(\mathrm{A}, \mathrm{B}; \mathrm{T}|_z\right)\right] = \delta.
\end{aligned}
$$

Finally, for the second term, we can write

$$\begin{aligned}
&\mathop{\mathrm{E}}_{z,a_{g(z)} \overset{\mathrm{R}}{\leftarrow} \mathrm{Z},\mathrm{A}_{g(z)}} \left[\mathrm{SD}\left(\left((\mathrm{A}_{\mathrm{T}'},\mathrm{T}',\mathrm{B})|_{z,a_{g(z)}}\right), \left((\mathrm{A}_{\mathrm{T}'},\mathrm{T}')|_{z,a_{g(z)}} \times \mathrm{B}|_{a_{g(z)},z}\right)\right)\right] \\
&= \mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}, a_g(z) \overset{\mathrm{R}}{\leftarrow} \mathrm{A}_g(z)} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|_z} \left[I_{SD}\left(\mathrm{A}_t; \mathrm{B}|_{a_{g(z)},z}\right)\right]\right] && \text{(Fact 4.12)} \\
&\le \mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}, a_g(z) \overset{\mathrm{R}}{\leftarrow} \mathrm{A}_g(z)} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|_z} \left[2\sqrt{\mathrm{I}(\mathrm{A}_t; \mathrm{B}|z, a_g(z))}\right]\right] && \text{(Fact 4.7)} \\
&\le 2\sqrt{\mathop{\mathrm{E}}_{z \overset{\mathrm{R}}{\leftarrow} \mathrm{Z}} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} T|_z} \left[\mathrm{I}(\mathrm{A}_t; \mathrm{B}|z, \mathrm{A}_{g(z)})\right]\right]} && \text{(Fact 4.15)} \\
&= 2\sqrt{\epsilon}.
\end{aligned}$$

□

Proof of Claim 6.9. Recall that we denote by $X_{t,<i}$ the restriction of X to coordinates in t that are less than i. Write

$$\begin{aligned}
&\mathrm{I}(\mathrm{M}_1; \mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T}, \mathrm{X}) \\
&= \mathop{\mathrm{E}}_{x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|\mathrm{X}=x} \left[\mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_t)|\mathrm{T}=t, \mathrm{X}=x)\right]\right] \\
&= \mathop{\mathrm{E}}_{x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|\mathrm{X}=x} \left[\sum_{i \in t} \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{t,<i}))\right]\right]. && \text{(Chain rule)}
\end{aligned}$$

For fixed x, t, i, by the chain rule,

$$\begin{aligned}
&\mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{t,<i})) \\
&\le \mathrm{I}(\mathrm{M}_1, \mathrm{F}(X_{\{1,\dots,i-1\}\setminus t}); \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{t,<i})) \\
&= \mathrm{I}(\mathrm{F}(X_{\{1,\dots,i-1\}\setminus t}); \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{t,<i})) \\
&\quad + \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{<i})) \\
&= 0 + \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{<i})).
\end{aligned}$$

Conditioned on X, A's message M_1 and the oracle F are independent of B's queries Y and therefore also from the intersection T. Therefore,

$$\begin{aligned}
\mathrm{I}(\mathrm{M}_1; \mathrm{F}(\mathrm{X}_\mathrm{T})|\mathrm{T}, \mathrm{X}) &\le \mathop{\mathrm{E}}_{x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|\mathrm{X}=x} \left[\sum_{i \in t} \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{T}=t, \mathrm{X}=x, \mathrm{F}(X_{<i}))\right]\right] \\
&= \mathop{\mathrm{E}}_{x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}} \left[\mathop{\mathrm{E}}_{t \overset{\mathrm{R}}{\leftarrow} \mathrm{T}|\mathrm{X}=x} \left[\sum_{i \in t} \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{X}=x, \mathrm{F}(X_{<i}))\right]\right] \\
&= \mathop{\mathrm{E}}_{x \overset{\mathrm{R}}{\leftarrow} \mathrm{X}} \left[\sum_{i} \Pr\left[i \in T \mid \mathrm{X}=x\right] \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i)|\mathrm{X}=x, \mathrm{F}(X_{<i}))\right].
\end{aligned}$$

From the assumption that no queries are heavy a priori, $\Pr\left[i \in T \mid \mathrm{X} = x\right] \leq \delta$ for all i. Finally,

$$\begin{aligned}
\mathrm{I}(\mathrm{M}_1; \mathrm{F}(\mathrm{X}_\mathrm{T}) | \mathrm{T}, \mathrm{X}) &\leq \delta \sum_i \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X_i) | \mathrm{X}, \mathrm{F}(X_{<i})) \\
&= \delta\, \mathrm{I}(\mathrm{M}_1; \mathrm{F}(X) | \mathrm{X}) && \text{(Chain rule)} \\
&\leq \delta |M_1|. && \text{(Fact 4.5)}
\end{aligned}$$

□

Proof of Claim 6.10. From Claim 6.7 and Corollary 4.19 we get that:

$$\begin{aligned}
&\underset{v_E \xleftarrow{\mathrm{R}} \Pi^{\mathrm{F}}_{\mathrm{Eve}}}{\mathrm{E}} \left[I_{SD}\left(\mathrm{X}, \mathrm{F}(\mathrm{X}); \mathrm{Y}|_{v_E}\right)\right] \\
&\leq 2 \underset{m_1, f(e_1), y \xleftarrow{\mathrm{R}} \mathrm{M}_1, \mathrm{F}(\mathcal{E}_1), \mathrm{Y}}{\mathrm{E}} \mathrm{SD}\Big(\big(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{m_1, f(e_1), y}\big), \\
&\qquad\qquad \big((\mathrm{X}, \mathrm{F}(\mathrm{X})|_{m_1, f(e_1), y}) \times (\mathrm{M}_2|_{m_1, f(e_1), y})\big)\Big).
\end{aligned}$$

For every $b_E = (y, m_1, f(e_1 \cap y))$,

$$\begin{aligned}
&\underset{f(e_1) \xleftarrow{\mathrm{R}} \mathrm{F}(e_1)|_{\mathrm{B}_E = b_E}}{\mathrm{E}} \mathrm{SD}\left(\big(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E, f(e_1)}\big), \big((\mathrm{X}, \mathrm{F}(\mathrm{X}) \times \mathrm{M}_2)|_{b_E, f(e_1)}\big)\right) \\
&\leq 2\mathrm{SD}\left((\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{F}(e_1), \mathrm{M}_2|_{b_E}), (\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{F}(e_1)|_{b_E} \times \mathrm{M}_2|_{b_E})\right) && \text{(Lemma 4.16)} \\
&= 2 \underset{x, f(x) \xleftarrow{\mathrm{R}} \mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\big(\mathrm{F}(e_1), \mathrm{M}_2|_{b_E, x, f(x)}\big), \big(\mathrm{F}(e_1)|_{b_E, x, f(x)} \times \mathrm{M}_2|_{b_E}\big)\right)\right] && \text{(Fact 4.12)}
\end{aligned}$$

Alice's message M_1 is only a function of $\mathrm{X}, \mathrm{F}(\mathrm{X})$, and Eve's queries $\mathcal{E}_1$ are a function of M_1. Thus, since F is a random function, $\mathrm{F}(\mathcal{E}_1 \setminus \mathrm{Y})$ is independent from $\mathrm{F}(\mathrm{Y} \setminus \mathcal{E}_1)$ conditioned on $\mathrm{M}_1, \mathrm{Y}, \mathrm{F}(\mathcal{E}_1 \cap \mathrm{Y}), \mathrm{X}, \mathrm{F}(\mathrm{X})$. Next, because M_2 is a function of M_1,Y and F(Y), we have that M_2 is independent from $\mathrm{F}(\mathcal{E}_1 \setminus \mathrm{Y})$ under the same conditioning.

We get that the distribution $(\mathrm{F}(e_1), \mathrm{M}_2|_{b_E, x, f(x))}$ is equal to

$$\mathrm{F}(e_1)|_{b_E, x, f(x)} \times \mathrm{M}_2|_{b_E, x, f(x)},$$

and therefore,

$$\begin{aligned}
&\underset{x, f(x) \xleftarrow{\mathrm{R}} \mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\big(\mathrm{F}(e_1), \mathrm{M}_2|_{b_E, x, f(x)}\big), \big(\mathrm{F}(e_1)|_{b_E, x, f(x)} \times \mathrm{M}_2|_{b_E}\big)\right)\right] \\
&= \underset{x, f(x) \xleftarrow{\mathrm{R}} \mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\big(\mathrm{F}(e_1)|_{b_E, x, f(x)} \times \mathrm{M}_2|_{b_E, x, f(x)}\big), \big(\mathrm{F}(e_1)|_{b_E, x, f(x)} \times \mathrm{M}_2|_{b_E}\big)\right)\right] \\
&= \underset{x, f(x) \xleftarrow{\mathrm{R}} \mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\big(\mathrm{M}_2|_{b_E, x, f(x)}\big), \big(\mathrm{M}_2|_{b_E}\big)\right)\right] && \text{(Fact 4.13)} \\
&= \mathrm{SD}\left((\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E}), ((\mathrm{X}, \mathrm{F}(\mathrm{X}) \times \mathrm{M}_2)|_{b_E})\right). && \text{(Fact 4.12)}
\end{aligned}$$

Now we can show all the dependence comes from the intersection. Since T_1 is a function of Y, X and $\mathcal{E}_1$, and $\mathcal{E}_1$ is a function of M_1, we get that

$$\begin{aligned}
&\mathrm{SD}\left((\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E}\right), ((\mathrm{X}, \mathrm{F}(\mathrm{X}) \times \mathrm{M}_2)|_{b_E})) \\
&= \mathrm{SD}\left((\mathrm{X}, \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1}), \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E}\right), (\mathrm{X}, \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1}), \mathrm{F}(\mathrm{X})|_{b_E} \times \mathrm{M}_2|_{b_E})) \\
&= \underset{t,f(y_t) \overset{\mathrm{R}}{\leftarrow} \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\left(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E,t,f(y_t)}\right), \left(\mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E,t,f(y_t)} \times \mathrm{M}_2|_{b_E}\right)\right)\right] \qquad \text{(Fact 4.12)}
\end{aligned}$$

Again, M_2 is a function of $\mathrm{Y}, \mathrm{F}(\mathrm{Y})$ and M_1, and $\mathrm{X}, \mathrm{F}(\mathrm{X})$ are independent from $\mathrm{F}(\mathrm{Y})$ conditioned on $\mathrm{M}_1, \mathrm{Y}, \mathrm{F}(\mathcal{E} \cap \mathrm{Y}), \mathrm{T}_1, \mathrm{F}(\mathrm{Y}_{\mathrm{T}_1})$. Thus, the distribution $(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E,t,f(y_t))}$ is equal to

$$\mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E,t,f(y_t)} \times \mathrm{M}_2|_{b_E,t,f(y_t)},$$

and we get:

$$\begin{aligned}
&\underset{t,f(y_t) \overset{\mathrm{R}}{\leftarrow} \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\left(\mathrm{X}, \mathrm{F}(\mathrm{X}), \mathrm{M}_2|_{b_E,t,f(y_t)}\right), \left(\mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E,t,f(y_t)} \times \mathrm{M}_2|_{b_E}\right)\right)\right] \\
&= \underset{t,f(y_t) \overset{\mathrm{R}}{\leftarrow} \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1})|_{b_E}}{\mathrm{E}} \Big[\mathrm{SD}\big(\left(\mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E,t,f(y_t)} \times \mathrm{M}_2|_{b_E,t,f(y_t)}\right), \\
&\qquad\qquad \left(\mathrm{X}, \mathrm{F}(\mathrm{X})|_{b_E,t,f(y_t)} \times \mathrm{M}_2|_{b_E}\right)\big)\Big] \\
&= \underset{t,f(y_t) \overset{\mathrm{R}}{\leftarrow} \mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1})|_{b_E}}{\mathrm{E}} \left[\mathrm{SD}\left(\left(\mathrm{M}_2|_{b_E,t,f(y_t)}\right), (\mathrm{M}_2|_{b_E})\right)\right] \qquad \text{(Fact 4.13)} \\
&= \mathrm{SD}\left((\mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1}), \mathrm{M}_2|_{b_E}\right), ((\mathrm{T}_1, \mathrm{F}(y_{\mathrm{T}_1}) \times \mathrm{M}_2)|_{b_E})). \qquad \text{(Fact 4.12)}
\end{aligned}$$

To conclude the proof, we take the expectation over B_E, and the claim follows by the monotonicity of expectation. □

Proof of Claim 6.12.

$$
\begin{aligned}
&\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\left[\mathrm{I}(\mathrm{F}(y_t);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y))\right]\\
&=\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\left[\sum_{i\in t}\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{t,<i}))\right] && \text{(Chain rule)}\\
&=\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\left[\sum_{i\in t}\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i}))\right]\\
&\qquad\qquad+\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\Bigg[\sum_{i\in t}\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{t,<i}))\\
&\qquad\qquad-\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i}))\Bigg]\\
&\le\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\left[\sum_{i\in t}\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i}))\right]\\
&\qquad\qquad+\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\left[\sum_{i\in t}\mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|m_1,y,\mathrm{F}(e_1\cap y))\right] && \text{(Lemma 4.8)}\\
&=\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\ \underset{t\xleftarrow{\mathrm{R}}\mathrm{T}_1|_{m_1,y}}{\mathrm{E}}\sum_{i\in t}\Big[\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i}))\\
&\qquad\qquad+\mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|m_1,y,\mathrm{F}(e_1\cap y))\Big]\\
&=\underset{m_1,y\xleftarrow{\mathrm{R}}\mathrm{M}_1,\mathrm{Y}}{\mathrm{E}}\sum_{i\in[\ell]}\Pr\left[i\in\mathrm{T}_1|m_1,y\right]\Big[\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i}))\\
&\qquad\qquad+\mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|m_1,y,\mathrm{F}(e_1\cap y))\Big]
\end{aligned}
$$

Since we excluded the heavy queries $\mathcal{E}_1$ from T_1, and y_i is some fixed query, and since X is independent from Y conditioned on M_1 we have

$$
\begin{aligned}
\Pr\left[i\in\mathrm{T}_1|m_1,y\right]=\Pr\left[y_i\in(\mathrm{X}\setminus\mathcal{E}_1)|m_1,y\right]&\le\Pr\left[y_i\in(\mathrm{X}\setminus\mathcal{E}_1)|m_1\right]\\
&\le\Pr\left[y_i\in((\mathrm{X}\cup\mathrm{Y})\setminus\mathcal{E}_1)|m_1\right]\le\delta.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\mathop{\mathrm{E}}_{m_1,y\overset{\mathrm{R}}{\leftarrow}\mathrm{M}_1,\mathrm{Y}} \sum_{i\in[\ell]} \Pr\left[i\in \mathrm{T}_1|_{m_1,y}\right] \Big[\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i})) \\
&\qquad\qquad + \mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|m_1,y,\mathrm{F}(e_1\cap y))\Big] \\
&\le \mathop{\mathrm{E}}_{m_1,y\overset{\mathrm{R}}{\leftarrow}\mathrm{M}_1,\mathrm{Y}} \sum_{i} \delta\Big[\mathrm{I}(\mathrm{F}(y_i);\mathrm{M}_2|m_1,y,\mathrm{F}(e_1\cap y),\mathrm{F}(y_{<i})) \\
&\qquad\qquad + \mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|m_1,y,\mathrm{F}(e_1\cap y))\Big] \\
&\le \delta \mathop{\mathrm{E}}_{y\overset{\mathrm{R}}{\leftarrow}\mathrm{Y}} \Big[\mathrm{I}(\mathrm{F}(y);\mathrm{M}_2|\mathrm{M}_1,y,\mathrm{F}(\mathcal{E}_1\cap y)) \\
&\qquad\qquad + \sum_i \mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|\mathrm{M}_1,y,\mathrm{F}(\mathcal{E}_1\cap y))\Big] && \text{(Chain rule)} \\
&\le \delta \mathop{\mathrm{E}}_{y\overset{\mathrm{R}}{\leftarrow}\mathrm{Y}} \Big[|\mathrm{M}_2| + \sum_i \mathrm{I}(\mathrm{F}(y_i);\mathrm{F}(y_{<i})|\mathrm{M}_1,y,\mathrm{F}(\mathcal{E}_1\cap y))\Big] && \text{(Fact 4.5)}
\end{aligned}
$$

□