Information-theoretic security with asymmetries

Tim Beyne and Yu Long Chen

imec-COSIC, KU Leuven, Belgium
name.lastname@esat.kuleuven.be

Abstract. In this paper, we study the problem of lower bounding any given cost function depending on the false positive and false negative probabilities of adversaries against indistinguishability security notions in symmetric-key cryptography. We take the cost model as an input, so that this becomes a purely information-theoretical question.
We propose power bounds as an easy-to-use alternative for advantage bounds in the context of indistinguishability with asymmetric cost functions. We show that standard proof techniques such as hybrid arguments and the H-coefficient method can be generalized to the power model, and apply these techniques to the PRP-PRF switching lemma, the Even-Mansour (EM) construction, and the sum-of-permutations (SoP) construction.
As the final and perhaps most useful contribution, we provide two methods to convert single-user power bounds into multi-user power bounds, and investigate their relation to the point-wise proximity method of Hoang and Tessaro (Crypto 2016). These method are applied to obtain tight multi-user power bounds for EM and SoP.

Keywords: Information theoretic security · Asymmetrical statistical costs · Power bounds · Multi-user security

1 Introduction

In the formal security analysis of cryptographic systems, the power of an adversary is always measured by the *advantage* that it achieves over a naive approach to attacking the system. For example, for indistinguishability games, the advantage of an adversary is the probability that it correctly identifies the real oracle, minus the probability that it mistakes the ideal oracle for the real one.

Traditionally, starting with the work of Goldwasser and Micali [14], a cryptosystem is considered to be secure if the advantage of all adversaries is a negligible function of the security parameter. However, in order to translate formal analysis into practical guidance, concrete estimates are often necessary. For example, at Crypto 1994, Bellare, Kilian and Rogaway [3] studied the security of the cipher block chaining mode, quantifying precisely how the advantage depends on the number of queries made by the adversary. A complete concrete treatment of symmetric-key cryptography was given by Bellare et al. [2], and concrete security has since become the norm in symmetric-key cryptography [4, 5, 22]. In the

same direction of bringing formal analysis closer to reality, Bellare, Boldyreva and Micali [1] have proposed that advantages should be bounded in a multi-user security model. Similarly, but in the context of symmetric-key provable security, a multi-key security model was proposed by Mouha and Luykx [19] at Crypto 2015. Many other refinements of the security model have been proposed, for example taking into account computational aspects such as preprocessing [13] and memory limitations [12].

There is one aspect of the security model that has not been subject to further refinement: the notion of advantage itself. In this paper, we argue that it can be useful to measure the power of adversaries in a different way because security games often exhibit asymmetric statistical costs. For example, in a real attack, the value of the target must always be weighted against processing costs. If the expected gains of the attack (for example measured in dollars) are negative, then the adversary will not proceed. Conversely, if the value of the target is immense, then making mistakes matters little to the adversary. In other words, we argue that – given a realistic cost model – security bounds should allow users (and adversaries) to predict the expected gain or loss of attacking the system, i.e. provide a *concrete* answer to the question 'is the attack worthwhile?'

In this paper, we study this question in the indistinguishability setting for symmetric-key constructions and take the cost model as an input, so that it becomes a purely information-theoretical problem: what is the optimal trade-off between false positives and false negatives, for a given cost function? This issue is intrinsically related to multi-user security, because the presence of multiple users implicitly tilts the cost function towards more tolerance for statistical errors. Indeed, standard definitions such as [19] consider it sufficient to compromise the security of a single user.

Related work. For information-theoretic indistinguishability games, the maximum possible advantage is equal to the statistical distance between the transcript distributions produced over the course of the game. Several alternative distances or contrast functions have been considered in previous work, but this was always for purely technical reasons. For example, Dai, Hoang and Tessaro [9] show that using the χ^2-divergence can be a powerful tool to obtain tighter bounds on the statistical distance.

Perhaps not surprisingly, alternatives to the statistical distance have proven to be useful to obtain multi-user security bounds from single-user bounds. If users are independent, then a folklore result says that the multi-user advantage is at most u times larger than the single-user advantage. There are currently only a handful of known ways to avoid this factor without resorting to a dedicated multi-user security analysis. At Crypto 2016, Hoang and Tessaro [16] have shown that *point-wise proximity* can sometimes be lifted from the single-user to the multi-user setting. This technique was refined to *approximate point-wise proximity* by the same authors at Eurocrypt 2017 [17]. The only other general method is the squared-ratio method from Crypto 2023 by Chen, Choi and Lee [7].

However, all of these methods still have significant limitations. The squared-ratio method is limited to independent users, and it only reduces the multi-user

security loss to a factor $\sqrt{u}$ and cannot avoid a factor of u loss for bad events. Point-wise proximity can be used to show that there is no multi-user security degradation and is applicable even when there are dependencies between users such as in the ideal permutation model, but it comes with technical limitations. For example, the bound must be a super-linear function of the number of queries. Some of the technical limitations can be overcome using approximate point-wise proximity, at the cost of making proofs more difficult.

Perhaps the most important limitation of all of these techniques is that they require a detailed analysis of transcript probabilities, which is typically much more difficult. Single-user hybrid proofs, on the contrary, are easy to understand because the advantages for intermediate steps add up and are easy to bound by analyzing the probabilities of bad events. This is an important reason why the statistical distance remains popular, despite the fact that it does not lift well to multi-user bounds.

Contribution. We propose *power bounds* as an easy-to-use alternative for advantage bounds in the context of indistinguishability. These bounds limit the statistical power (one minus the false negative rate) of any statistical test in terms of its false positive rate. A power bound implicitly lower bounds the cost of all distinguishing attacks *in any cost model*. It is shown that standard proof techniques (such as hybrid arguments, bounding bad events and the H-coefficient method) have natural counterparts for power bounds. To demonstrate that these bounds are no more difficult to obtain than advantage bounds for many constructions, we derive power bounds for the Even-Mansour and sum-of-permutations constructions. It is also shown that 'good' (linear) power bounds automatically lift to the multi-user setting. This allows us to deduce the first tight multi-user advantage bound for the sum-of-permutation construction.

From a probability theorist's point of view, we essentially apply the Neyman-Pearson theory of hypothesis testing to information-theoretic security and explore the consequences. As we explain in Section 3, this classical theory describes hypothesis tests as a decision-theoretic problem based on two types of errors. The first error probability is the false positive rate α, i.e. the probability that the null hypothesis is falsely rejected. The second error probability is the false negative rate β, or the probability that the alternative hypothesis is incorrectly rejected. A power bound is simply any upper bound of the form $1 - \beta \leq f(\alpha)$ (in statistics, $1-\beta$ is called the power of a test). For any statistical cost function $\mathscr{C}$ of the error probabilities α and β, the power bound implies the lower bound $\mathscr{C}(\alpha, \beta) \geq \mathscr{C}(\alpha, 1 - f(\alpha))$. This bound is tight if f is tight. Advantages only give a tight bound if $\mathscr{C}$ is a symmetric function. More precisely, the maximum advantage attack minimizes the cost function $\mathscr{C}(\alpha, \beta) = \alpha + \beta$.

It may seem nontrivial to obtain such bounds for realistic constructions, but in Section 4 we show that standard proof techniques can be generalized. For example, Theorem 4 shows that the power bound of a hybrid system with non-decreasing power bounds f and g is given by $g \circ f$. Theorem 5 shows how bad events can be ruled out by bounding their probabilities in the usual way. The widely used H-coefficient technique of Patarin [21] and its generalization by Chen

and Steinberger [6] are extended to power bounds by Theorem 7. An interesting aspect of our extension is that it suggests that the H-coefficient technique should not be over-used, in the sense that one should avoid excluding bad events to point that the probability ratio for good transcripts becomes greater than one. As our applications show, this is often feasible and results in better power bounds – which in turn is important to obtain a good multi-user bound.

Our first applications beyond the PRP-PRF switching lemma, which we use as a running example, are presented in Section 5. We first consider the classical Even-Mansour construction [11], arguably the simplest construction to build a pseudorandom permutation from a public permutation. Given an n-bit permutation π, as well as an n-bit key K, the Even-Mansour construction $\mathsf{EM}_K[\pi]$ is defined by

$$\mathsf{EM}_K[\pi](M) = \pi(M \oplus K) \oplus K\,.$$

A tight power bound for Even-Mansour can be obtained by a simple hybrid argument. If the ideal world is taken as the null hypothesis, then the power bound we obtain is given by (with $N = 2^n$)

$$1 - \beta \leq \frac{\alpha}{1 - \frac{2qp}{N}}\,,$$

where q is the number of construction queries and p is the number of queries made to π. This implies the standard advantage bound, but provides more detailed information in the face of asymmetric statistical costs. An important difference in our security proof is the way the bad events are defined. They are defined in a way such that they never appear in the real world, unlike in traditional proofs for advantage bounds of the Even-Mansour construction such as [19].

The second application we consider is the sum-of-permutations construction proposed by Bellare et al. [4] and Hall et al. [15]. The SoP construction turns block ciphers into pseudorandom functions with security beyond the birthday-bound barrier, i.e., above $2^{n/2}$. Given an n-bit (keyed) pseudorandom permutation π, the sum-of-permutations is defined by

$$\mathsf{SoP}[\pi](M) = \pi(0\|M) + \pi(1\|M)\,.$$

For this construction, a simple hybrid proof does not suffice. Instead, we rely on a combination of our hybrid and H-coefficient theorems. Again, interestingly from a technical point of view, out approach requires that the bad events only happen in one of the two worlds – but this is almost always possible by redefining the events.

Our final and perhaps most useful contribution from a proof-technical point of view are two methods to convert single-user power bounds into multi-user power bounds. Both are presented in Section 6. The first method is based on the observation that linear power bounds imply point-wise proximity as defined by Hoang and Tessaro [16] *and*, conversely, point-wise proximity implies a linear power bound. This equivalence allows us to lift linear single-user power bounds

to multi-user power bounds, though the technical conditions from [16] related to the dependence of the bound on the number of queries still apply. We apply this result to obtain the multi-user power bound of the EM construction. Another consequence of this result is that our proof techniques for power bounds can be used to obtain point-wise proximity estimates in an easier way.

The second and more interesting method is proposed in Section 6.2. It still assumes that the single-user power bound is linear, and additionally assumes that users are independent. The latter condition is satisfied for standard model proofs. The benefit of this method is that it completely removes the technical conditions related to the dependence of the bound on the number of queries that are required by the proof of Hoang and Tessaro [16] – without making the analysis more complicated, as done in [17]. Since it applies to linear power bounds, this result can also be understood as a way to obtain tighter point-wise proximity estimates with fewer assumptions. We apply the second method to the sum-of-permutations construction to obtain multi-user power bounds. Our bound implies the first tight advantage bound for this construction that does not contain a term of the form uq/N. This should be compared to previous partial results in this direction by Choi et al. [8], who obtain a bound that still contains terms of the form $\sqrt{u}q^2/N^2$ and relies on an irregular security model where the ideal world oracle is forbidden to output zero (and with a difficult proof).

2 Notation

Throughout this paper, we work with probabilities on a finite sample space $\mathcal{T}$ with its power set $2^{\mathcal{T}}$ as the space of events. In this context, a probability distribution is a function $P : 2^{\mathcal{T}} \to [0, 1]$ such that $P(\mathcal{T}) = 1$. In information-theoretical security proofs, $\mathcal{T}$ is typically the set of all possible transcripts that can be produced during the course of some security game.

In the following, we will consider indistinguishability security games. In terms of probability theory, a deterministic information-theoretical distinguisher is the same as a hypothesis test between two transcript distributions P and Q. Traditionally, indistinguishability insecurity is measured by the statistical distance

$$\Delta\big(P\,;\,Q\big) = \max_{\mathcal{R}\subseteq\mathcal{T}} \big|P(\mathcal{R}) - Q(\mathcal{R})\big|\,.$$

This is the same as the maximum *advantage* over all possible distinguishers, i.e. hypothesis tests with critical region $\mathcal{R}$. It is not difficult to see that the maximum is achieved for some $\mathcal{R}$ with $P(\mathcal{R}) \geq Q(\mathcal{R})$.

Although advantages are symmetric, in our setting, the choice of either P or Q as the null hypothesis is actually important. In Section 3 we will introduce other measures of insecurity that are not symmetrical in P and Q. For this reason, we will use the phrase 'distinguisher from P to Q' to emphasize that P is taken to be the null hypothesis.

3 Asymmetrical statistical costs

In the Neyman-Pearson theory of hypothesis testing [20], distinguishing between two distributions P and Q is interpreted as a statistical decision problem. The goal is to minimize two types of statistical errors. Assume that P is the null hypothesis, and Q the alternative – i.e. a distinguisher from P to Q. For every hypothesis test, there exists a critical region $\mathcal{R}$ such that the null hypothesis is rejected if the sample is in $\mathcal{R}$.

The first type of errors are *false positives*: the null hypothesis P is rejected, even though it is true. Note that, as is common practice, we consider rejecting the null hypothesis a positive outcome. The probability of these errors is equal to $\alpha = P(\mathcal{R})$. The error probability α is also called the significance level of the test. The second type of errors are *false negatives*, meaning that the null hypothesis is accepted in spite of the alternative being true. The false negative probability is $\beta = 1 - Q(\mathcal{R})$. The probability $1 - \beta = Q(\mathcal{R})$ is called the *power* of the test.

The advantage of a distinguisher with critical region $\mathcal{R}$ is equal to $Q(\mathcal{R}) - P(\mathcal{R}) = (1-\beta)-\alpha$. In Section 3.1, we argue why is interesting to upper bound the *power* of distinguishers rather than just their advantage. A few basic properties of power bounds are discussed in Section 3.2. A first application to the PRP-PRF switching lemma is given in Section 3.3.

3.1 Bounding the power of a hypothesis test

Since choosing a smaller critical region $\mathcal{R}$ decreases α but increases β, there exists a trade-off between the two types of errors. From a decision-theoretic point of view, one should choose $\mathcal{R}$ to minimize some application-dependent cost function $\mathscr{C}(\alpha, \beta)$. It is reasonable to assume that $\mathscr{C}$ is non-decreasing in both arguments. An example is the function

$$\mathscr{C}(\alpha, \beta) = \alpha + \beta\,.$$

Minimizing this expression with respect to $\mathcal{R}$ is equivalent to maximizing the advantage, since $\mathscr{C}(\alpha, \beta) = 1 - (Q(\mathcal{R}) - P(\mathcal{R}))$. However, in many applications, the cost function $\mathscr{C}$ is not symmetric in α and β.

Example 1 (Asymmetrical costs). Suppose a malicious actor wants to deploy a critical vulnerability at scale, but does not know which targets are vulnerable[1]. For simplicity, assume that half of all users are vulnerable. To test for vulnerability, the actor develops a distinguishing attack. In practice, exploiting the vulnerability comes at some cost (for example computational) that can often be expressed in monetary terms. Assume that the marginal cost is \$1. If the value of an average target is \$10, then the cost function is

$$\mathscr{C}(\alpha, \beta) = \$9\,\alpha + \$1\,\beta\,.$$

[1] Realistic examples are provided by weak-key attacks on block ciphers.

The cost of a false negative is the \$1 marginal cost of exploiting the vulnerability. Every false positive comes at a \$9 cost, i.e. a \$10 opportunity cost (missed target) minus a \$1 saving because the vulnerability does not have to be used. This example ignores the costs of the distinguisher itself. In the information-theoretical setting, one can conservatively assume that these costs do not depend on α and β. In the computational setting, this may not be the case. ▷

Let f be an increasing function so that $1 - \beta \leq f(\alpha)$ for all possible distinguishers from P to Q with false positive rate α and false negative rate β. For any non-decreasing cost function $\mathscr{C}$, one has a corresponding lower bound

$$\mathscr{C}(\alpha, \beta) \geq \mathscr{C}(\alpha, 1 - f(\alpha)) \,.$$

Hence, our upper bound $1 - \beta \leq f(\alpha)$ gives a lower bound on the cost of any distinguisher – no matter how costs are defined. This is one reason why we argue that power bounds are the right way to deal with asymmetrical statistical costs.

There is also a more concrete motivation for investigating power bounds. In Section 6 of this paper, we show that upper bounds of the form $1 - \beta \leq f(\alpha)$ automatically imply tight multi-user security power bounds in the standard model. As shown below in Section 3.2, this in turns implies tight multi-user advantage bounds. For example, our Theorem 13 implies the following informal statement. From a probability theory point of view, the number of queries q mentioned in the claim below is simply a restriction on the probability space.

Claim (Informal)**.** *Given a single-user power bound $1-\beta \leq c(q)\,\alpha$ for q queries, the multi-user power bound for u independent and identical users making a total of q queries is given by*

$$1 - \beta' \leq \alpha' \max_{q_1 + \cdots + q_u \leq q} \prod_{i=1}^{q} c(q_i) \,,$$

where α' and β' are the multi-user false positive and negative rates, respectively.

Variants of this result with some dependency between the users, such as a shared public cryptographic primitive, will also be obtained. Given the strength of such results, one might expect power bounds $1 - \beta \leq f(\alpha)$ to be much more difficult to prove. However, this is not the case: it will be shown in Section 4 that common proof techniques for advantages such as hybrid arguments and the H-coefficient method have analogues for power bounds.

3.2 Some properties of power bounds

Before we begin analyzing specific problems, a few properties of power bounds are worth pointing out. First, given a bound of the form $1 - \beta \leq f(\alpha)$ for distinguishers from P to Q, one can always deduce a bound on the statistical distance:

$$\Delta\big(P\,;\,Q\big) = \max_{\mathcal{R}} \big|Q(\mathcal{R}) - P(\mathcal{R})\big| \leq \sup_{\alpha} f(\alpha) - \alpha \,.$$

Indeed, the maximum is achieved for some $\mathcal{R}$ such that $Q(\mathcal{R}) \geq P(\mathcal{R})$ and if $P(\mathcal{R}) = \alpha$, then $Q(\mathcal{R}) \leq f(\alpha)$. The advantage is symmetric in P and Q, so one might conclude that – for the purpose of deriving advantage bounds – the choice of the null hypothesis matters little. However, this is not the case: a good choice of the null hypothesis can make multi-user security proofs much simpler.

Since it so important throughout this paper, it is worth discussing the asymmetry between the null hypothesis and the alternative hypothesis in more detail. If the problem was symmetric, then one would expect the best distinguisher from Q to P to be obtained by replacing the critical region of the best distinguisher from P to Q by its complement. The best possible critical region is characterized by Theorem 1, i.e. the Neyman-Pearson lemma. In this result, $\mathcal{R}_t$ is the optimal critical region among all tests from P to Q with false positive rate $P(\mathcal{R}_t)$. The false positive rate can be decreased by reducing the threshold t.

Theorem 1 (Neyman-Pearson lemma). *Let P and Q be distributions on a finite set $\mathcal{T}$ with probability mass functions p and q. For all $t > 0$, let*

$$\mathcal{R}_t = \left\{ x \in \mathcal{T} \;\middle|\; p(x) \leq t\, q(x) \right\}.$$

If $\mathcal{S}$ is a subset of $\mathcal{T}$ such that $P(\mathcal{S}) \leq P(\mathcal{R}_t)$, then $Q(\mathcal{S}) \leq Q(\mathcal{R}_t)$.

Proof. The result is well known, but we provide the proof for completeness to help the reader familiarize with our notation. The definition of $\mathcal{R}_t$ implies the following inequalities:

$$Q(\mathcal{S} \setminus \mathcal{R}_t) \leq \frac{1}{t}\, P(\mathcal{S} \setminus \mathcal{R}_t)\,,$$
$$P(\mathcal{R}_t \setminus S) \leq t\, Q(\mathcal{R}_t \setminus \mathcal{S})\,.$$

It follows from $\mathcal{S} = (\mathcal{R}_t \cap \mathcal{S}) \cup (\mathcal{S} \setminus \mathcal{R}_t)$ that

$$Q(\mathcal{S}) = Q(\mathcal{R}_t \cap \mathcal{S}) + Q(\mathcal{S} \setminus \mathcal{R}_t) \leq Q(\mathcal{R}_t \cap \mathcal{S}) + \frac{1}{t} P(\mathcal{S} \setminus \mathcal{R}_t)$$

The condition $P(\mathcal{S}) \leq P(\mathcal{R}_t)$ implies that $P(\mathcal{S} \setminus \mathcal{R}_t) \leq P(\mathcal{R}_t \setminus \mathcal{S})$. Hence,

$$Q(\mathcal{S}) \leq Q(\mathcal{R}_t \cap \mathcal{S}) + \frac{1}{t} P(\mathcal{R}_t \setminus \mathcal{S}) \leq Q(\mathcal{R}_t \cap \mathcal{S}) + Q(\mathcal{R}_t \setminus \mathcal{S})\,.$$

It follows that $Q(\mathcal{S}) \leq Q(\mathcal{R}_t)$, as claimed. □

Let $\mathcal{R}_t^*$ denote the optimal critical region for tests from Q to P. Assume, without loss of generality, that the boundary $p(x) = tq(x)$ contains no points x. In this case, the complement of $\mathcal{R}_t^*$ is equal to $\mathcal{R}_{1/t}$. The asymmetry of the problem is then clear from the fact that, in most cases, $\mathcal{R}_{1/t} \neq \mathcal{R}_t$. Nevertheless, equality always holds in the singular case $t = 1$. This corresponds to $\alpha = \beta$, maximizing the advantage. There are of course well-known exceptions to this, such as when comparing two normal distributions with the same mean, where $\mathcal{R}_{1/t} = \mathcal{R}_t$ for all t.

Despite the asymmetry of hypothesis testing, a power bound for distinguishers from P to Q can always be converted to a power bound for distinguishers from Q to P. The idea is to take the complement of the critical region, as above. For all the power bounds that we prove in this paper, the following result provides an essentially tight bound when P and Q swapped. Note, however, that the two bounds will typically be completely different.

Theorem 2. *If f is an increasing function such that $1 - \beta \leq f(\alpha)$ for all distinguishers from P to Q with false positive and false negative rates α and β, then every distinguisher from Q to P with false positive and false negative rates α' and β' satisfies $1 - \beta' \leq 1 - f^{-1}(1 - \alpha')$.*

Proof. Let $\mathcal{T}$ denote the sample space of P and Q. Let us construct a test from Q to P by using $\mathcal{T} \setminus \mathcal{R}$ as the rejection region. Since $Q(\mathcal{R}) \leq f(P(\mathcal{R}))$ for any set $\mathcal{R}$, we have

$$\alpha' = Q(\mathcal{T} \setminus \mathcal{R}) = 1 - Q(\mathcal{R}) \geq 1 - f(P(\mathcal{R}))\,,$$

Since f is increasing, it follows that $P(\mathcal{R}) \geq f^{-1}(1 - \alpha')$. Hence,

$$1 - \beta' = P(\mathcal{T} \setminus \mathcal{R}) = 1 - P(\mathcal{R}) \leq 1 - f^{-1}(1 - \alpha')\,.$$

This is the desired result. □

3.3 Example: PRP-PRF switching lemma

The PRP-PRF switching lemma [3, 5, 18] is one of the most widely used results for proving the birthday-bound security of symmetric-key constructions. The lemma bounds the advantage of distinguishing a uniform random function ρ from a uniform random permutation π on a domain of size N and using q queries as $q(q-1)/N$. Theorem 3 gives an upper bound of the form $1 - \beta \leq f(\alpha)$ for this problem.

Theorem 3. *Let ρ be a uniform random function and π a uniform random permutation, both on a domain of size N. For every distinguisher from ρ to π with error probability α and power $1-\beta$ that makes at most $q \leq \sqrt{2N}+1$ queries,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{q(q-1)}{2N}}\,.$$

Proof. Since the resulting bound is non-decreasing in q, we can assume that the distinguisher never queries duplicate inputs. Hence, under the null hypothesis, the transcript consists of q outputs of ρ. In the ideal world, it consists of q outputs of π. Let P denote the distribution of transcripts for ρ. The probability that a transcript is in the rejection region $\mathcal{R}$ is equal to

$$\alpha = P(\mathcal{R}) \geq \frac{|\mathcal{R}|}{N^q}\,,$$

since every input-output pair occurs with probability $1/N$, independently of other pairs. Here, we use the fact that the rejection region does not contain transcripts with duplicate inputs. Let Q denote the distribution of transcripts for π. The probability that the transcript is in the rejection region $\mathcal{R}$ is

$$1 - \beta = Q(\mathcal{R}) \leq \frac{(N-q)!}{N!} \left|\mathcal{R}\right| .$$

Indeed, there are $(N-q)!$ permutations with q specified outputs, assuming the outputs are distinct. Hence, every transcript contains q distinct outputs has probability $(N-q)!/N!$ under Q. If two of the outputs in a transcript are equal, then its probability is always zero. Hence, we obtain an upper bound.

The lower bound on $1-\beta$ and the upper bound on α can now be combined to obtain the result. Substituting the upper bound $|\mathcal{R}| \leq N^q \, \alpha$ into the upper bound on $1-\beta$ yields

$$1 - \beta \leq N^q \, \frac{(N-q)!}{N!} \, \alpha .$$

The simplified bound follows from the inequality

$$N^q \, \frac{(N-q)!}{N!} \leq \prod_{i=1}^{q-1} \frac{1}{1-\frac{i}{N}} \leq \frac{1}{1-\sum_{i=1}^{q-1} \frac{i}{N}} = \frac{1}{1-\frac{q(q-1)}{2N}} ,$$

where the second inequality only works if $q(q-1) < 2N$. □

Despite the fact that the PRP-PRF switching lemma is quite simple, it already leads to several interesting conclusions. Figure 1 compares Theorem 3 to the bound obtained from the advantage, which corresponds to

$$1 - \beta \leq \alpha + \frac{q(q-1)}{N} .$$

As can be seen from Figure 1, the gap between both bounds is proportional to $1-\alpha$. This means that it is considerably more difficult to achieve a low error probability α than one might expect from the advantage bound. Taking a closer look at the problem reveals why this is the case. The only way to reject the null hypothesis that the primitive is a uniform random function, is to note the *absence of collisions*. However, there is always a chance that the absence of collisions is just bad luck. Hence, if one wants to be highly confident (low α) in the decision, the power cannot be too high.

If the null and alternative hypothesis are swapped, i.e. the null hypothesis is that the primitive is a uniform random permutation, then the results are completely different. By Theorem 2, for this case we have

$$1 - \beta \leq 1 - \left(1 - \frac{q(q-1)}{2N}\right) (1-\alpha) \leq \frac{q(q-1)}{2N} + \left(1 - \frac{q(q-1)}{2N}\right) \alpha .$$

It turns out that this is essentially tight, and for a good reason: if even a single collision is observed, then the null hypothesis can be rejected with certainty

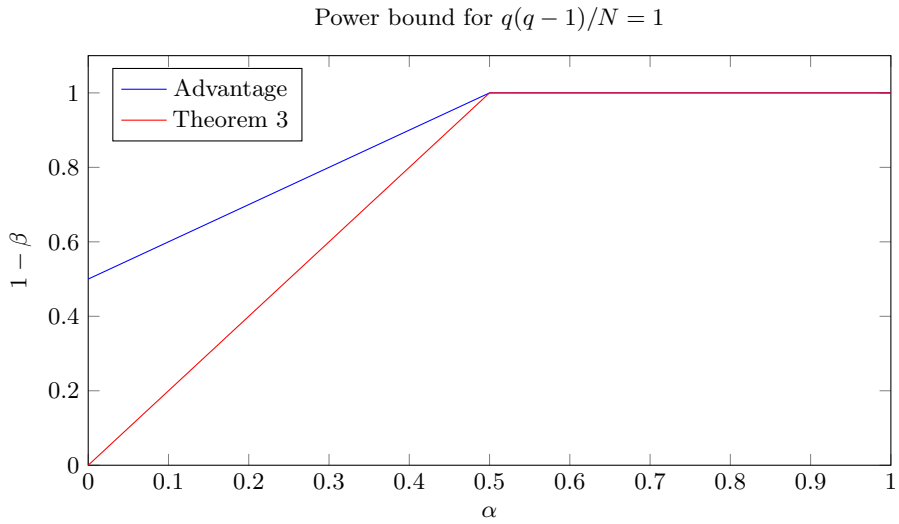

Fig. 1. Comparison of the advantage-based bound for the PRP-PRF switching lemma and Theorem 3. The relative tightness loss increases as q decreases.

(power equal to one). This case corresponds to the first term in the bound. The overall power of the test still depends on α, but only in a trivial way: if no collision occurs, then the decision is down to random guessing. This corresponds to the second term in the bound.

In Section 6, we will see that Theorem 3 implies a multi-user security bound that does not incur a security loss. However, if the advantage bound is used or if one chooses the uniform random permutation as the null hypothesis, then the resulting multi-user advantage bound would be worse by a factor equal to the number of users. Before turning to multi-user security, we first introduce a number of proof techniques that will help to simplify proofs – including the above proof of Theorem 3.

4 Proof techniques

In this section, we extend the most commonly used proof techniques to the asymmetric setting. More precisely, in Section 4.1 we show how to use hybrid arguments with power bounds. Hybrid arguments are often used to exclude 'bad events', so we discuss this specific case in Section 4.2 and illustrate it by giving an alternative and shorter proof of Theorem 3. Finally, Section 4.3 extends the H-coefficient method of Patarin [21].

4.1 Hybrid arguments

Probably the most widely used technique in provable security is the 'hybrid argument', or equivalently the triangle inequality for the total variation distance.

That is, for distributions P, Q and X on the same domain,

$$\Delta\big(P\,;\,Q\big) \le \Delta\big(P\,;\,X\big) + \Delta\big(X\,;\,Q\big)\,.$$

Power bounds are not additive. However, as shown by Theorem 4, the hybrid argument has a natural generalization: the power bound for distinguishers from P to Q is the functional composition of the bounds for distinguishers from P to X and from X to Q.

Theorem 4. *Let P, Q and X be distributions on a common domain. If every distinguisher from P to X with false positive and negative rates α_1 and β_1 satisfies $1-\beta_1 \le f(\alpha_1)$, and if every distinguisher from X to Q with false positive and negative rates α_2 and β_2 satisfies $1-\beta_2 \le g(\alpha_2)$ with f and g non-decreasing, then for all distinguishers between P and Q*

$$1-\beta \le g(f(\alpha))\,,$$

with α the false positive rate and β the false negative rate.

Proof. The proof is straightforward. Since f and g are non-decreasing functions,

$$1-\beta = Q(\mathcal{R}) \le g(X(\mathcal{R})) \le g(f(P(\mathcal{R}))) \le g(f(\alpha))\,,$$

where we choose $\alpha_1 = P(\mathcal{R})$, $\beta_1 = 1-X(\mathcal{R})$, $\alpha_2 = X(\mathcal{R})$ and $\beta_2 = 1-Q(\mathcal{R})$. □

An interesting consequence of Theorem 4 is that when multiple hybrid arguments are applied sequentially, in general the order matters. This is again related to the asymmetry of hypothesis testing.

Example 2 (Counter mode). To illustrate Theorem 4, we prove the security of the nonce-based counter mode based on a uniform random permutation π (see Figure 2). Let us take an ideal encryption scheme as the null hypothesis, with transcript distribution P. The transcript distribution of counter mode will be denoted by Q. To prove a power bound, define an intermediate transcript distribution X equal to the distribution of transcripts produced by counter mode based on a uniform random function ρ.

Under the assumption that nonces are unique, the input to the random function ρ is always fresh. Hence, it is clear that P and X are actually indistinguishable. That is, $1-\beta_1 \le f(\alpha_1)$ with $f(\alpha_1) = \alpha_1$ for all distinguishers from P to X with false positive rate α_1 and false negative rate β_1. Indeed, recall that $1-\beta_1 = \alpha_1$ corresponds to pure guessing. By a standard reduction argument, any distinguisher from X to Q can be turned into a distinguisher from ρ to π that makes at most σ queries, where σ is the total number of blocks encrypted with counter mode. Hence, by Theorem 3, every distinguisher from X to Q with false positive rate α_2 and false negative rate β_2 satisfies $1-\beta_2 \le g(\alpha_2)$, where

$$g(\alpha_2) = \frac{\alpha_2}{1-\frac{\sigma(\sigma-1)}{2N}}\,.$$

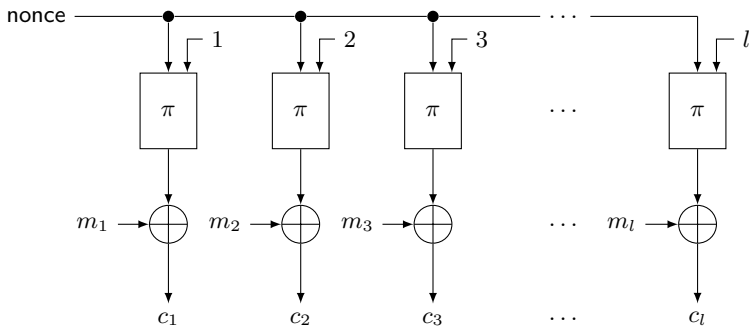

Fig. 2. Nonce-based counter mode using a uniform random permutation π.

By Theorem 4, the bounds from P to X and from X to Q can be combined to obtain the following bound for any distinguisher from P to Q:

$$1 - \beta \leq g(f(\alpha)) = \frac{\alpha}{1 - \frac{\sigma(\sigma-1)}{2N}} ,$$

where α is the false positive rate and β the false negative rate. ▷

4.2 Excluding bad events

Let P be a distribution on $\mathcal{T}$ and E an event with nonzero probability. We denote the distribution of P conditioned on E by P_E. That is,

$$P_E(A) = \frac{P(A \cap E)}{P(E)} ,$$

for all subsets A of $\mathcal{T}$. For distinguishers from P to P_E, we have the following result. In Theorem 5, the event E is denoted by $\mathcal{T} \setminus B$. This refers to the fact that, when proving power bounds, B is a 'bad event' that we would like to exclude from consideration.

Theorem 5. *Let P be a distribution on $\mathcal{T}$ and let $B \subseteq \mathcal{T}$ be some event with $P(B) \leq \varepsilon$. For all distinguishers from P to $P_{\mathcal{T} \setminus B}$,*

$$1 - \beta \leq \frac{\alpha}{1 - \varepsilon} ,$$

with α the false positive rate and β the false negative rate.

Proof. Let $\mathcal{R}$ be the critical region of the distinguisher. By definition,

$$1 - \beta = P_{\mathcal{T} \setminus B}(\mathcal{R}) = \frac{P(\mathcal{R} \setminus B)}{P(\mathcal{T} \setminus B)} \leq \frac{P(\mathcal{R})}{1 - P(B)} \leq \frac{\alpha}{1 - \varepsilon} ,$$

where we use $\alpha = P(\mathcal{R})$ and $P(B) \leq \varepsilon$ in the last inequality. □

Let us stress again that the choice of P as the null hypothesis in Theorem 5 is essential. Swapping P and $P_{\mathcal{T}\setminus B}$ would result in a significantly worse bound of the form $1 - \beta \leq \alpha + \varepsilon$. Despite its simplicity, Theorem 5 often simplifies proofs considerably. We illustrate this by giving a significantly shorter proof of Theorem 3. The proof in Example 3 is more intuitive, and more similar to typical proofs of advantage bounds for distinguishers between uniform random functions and uniform random permutations.

Example 3. As in Theorem 3, let ρ be a uniform random function and π a uniform random permutation. Assume that q distinct queries are made, and let $\mathcal{T}$ denote the set of all possible transcripts of length q. Denote the distribution of transcripts for ρ by P and the distribution of transcripts for π by Q. Intuitively, the only difference between ρ and π is that collisions may occur between the outputs of ρ but not between the outputs of π. Formally, if B is the set of all transcripts that contain a collision, then $Q = P_{\mathcal{T}\setminus B}$. Since every pair of outputs of ρ collides with probability $1/N$, the union bound gives

$$P(B) \leq \frac{1}{N}\binom{q}{2} = \frac{q(q-1)}{2N}\,.$$

Hence, by Theorem 5, for all distinguishers from P to Q,

$$1 - \beta \leq \frac{\alpha}{1 - \frac{q(q-1)}{2N}}\,,$$

with α the false positive rate and β the false negative rate. ▷

4.3 H-coefficient technique

Many security proofs in symmetric-key cryptography rely on the *H-coefficient* technique of Patarin [21]. In this section, we develop asymmetric variants of this result that can be used to prove power bounds.

First, we recall the classical H-coefficient theorem as described by Chen and Steinberger [6]. In the following, $\mathcal{T} = \mathcal{T}_{\mathsf{g}} \sqcup \mathcal{T}_{\mathsf{b}}$ denotes a partition of $\mathcal{T}$ into subsets $\mathcal{T}_{\mathsf{g}}$ and $\mathcal{T}_{\mathsf{b}}$. These are typically called 'good' and 'bad' sets of transcripts. To simplify the statement, we reformulate it in terms of distributions rather than probability mass functions.

Theorem 6 (H-coefficient technique). *Let P and Q be distributions on $\mathcal{T} = \mathcal{T}_{\mathsf{g}} \sqcup \mathcal{T}_{\mathsf{b}}$. If $Q(\mathcal{T}_{\mathsf{b}}) \leq \varepsilon_{\mathsf{b}}$ and $P(E) \geq (1 - \varepsilon_{\mathsf{g}})Q(E)$ for all $E \subseteq \mathcal{T}_{\mathsf{g}}$, then*

$$\Delta\big(P\,;\,Q\big) \leq \varepsilon_{\mathsf{g}} + \varepsilon_{\mathsf{b}}\,.$$

Proof. Since our description is slightly more general than that of Chen and Steinberger [6], we give a proof for completeness. The partition of $\mathcal{T}$ induces a partition of any critical region $\mathcal{R} = \mathcal{R}_{\mathsf{g}} \sqcup \mathcal{R}_{\mathsf{b}}$ with $\mathcal{R}_{\mathsf{g}} = \mathcal{R} \cap \mathcal{T}_{\mathsf{g}}$ and $\mathcal{R}_{\mathsf{b}} = \mathcal{R} \cap \mathcal{T}_{\mathsf{b}}$. The advantage is equal to

$$\Delta\big(P\,;\,Q\big) \leq \max_{\mathcal{R}} Q(\mathcal{R}) - P(\mathcal{R}) = \max_{\mathcal{R}} Q(\mathcal{R}_{\mathsf{b}}) + Q(\mathcal{R}_{\mathsf{g}}) - P(\mathcal{R}_{\mathsf{g}}) - P(\mathcal{R}_{\mathsf{b}})\,,$$

where the maximum is over all $\mathcal{R}$ such that $Q(\mathcal{R}) \geq P(\mathcal{R})$. Using the fact that $Q(\mathcal{R}_{\mathsf{b}}) \leq \varepsilon_{\mathsf{b}}$ and $P(\mathcal{R}_{\mathsf{b}}) \geq 0$, one gets

$$\Delta\big(P\,;\,Q\big) \leq \varepsilon_{\mathsf{b}} + \max_{\mathcal{R}} Q(\mathcal{R}_{\mathsf{g}}) - P(\mathcal{R}_{\mathsf{g}}) \leq \varepsilon_{\mathsf{b}} + \varepsilon_{\mathsf{g}} \max_{\mathcal{R}} Q(\mathcal{R}_{\mathsf{g}})\,,$$

The result follows from the fact that $Q(\mathcal{R}_{\mathsf{g}}) \leq 1$. □

In most applications of Theorem 6, the lower bound $P(E) \geq (1-\varepsilon_{\mathsf{g}})Q(E)$ for subsets E of $\mathcal{T}_{\mathsf{g}}$ is obtained by lower bounding the ratio

$$\frac{p(\tau)}{q(\tau)} \geq 1 - \varepsilon_{\mathsf{g}}\,,$$

for all transcripts τ in $\mathcal{T}_{\mathsf{g}}$ with $q(\tau) \neq 0$. More generally, however, it is sufficient to bound the expectation of $\min\{1, p(\boldsymbol{\tau})/q(\boldsymbol{\tau})\}$ for a random variable $\boldsymbol{\tau}$ with distribution Q. This is the special case described by Chen and Steinberger [6] and is known as the expectation method.

The following theorem extends the H-coefficient technique to our setting, so that it can be used to obtain bounds of the form $1-\beta \leq f(\alpha)$. The proof is only slightly more complicated.

Theorem 7. *Let P and Q be distributions on $\mathcal{T} = \mathcal{T}_{\mathsf{g}} \sqcup \mathcal{T}_{\mathsf{b}}$. If $Q(\mathcal{T}_{\mathsf{b}}) \leq \varepsilon_{\mathsf{b}}$ and $P(E) \geq (1-\varepsilon_{\mathsf{g}})Q(E)$ for all $E \subseteq \mathcal{T}_{\mathsf{g}}$, then for every distinguisher from P to Q,*

$$1 - \beta \leq \varepsilon_{\mathsf{b}} + \frac{\alpha}{1-\varepsilon_{\mathsf{g}}}\,,$$

with α the false positive rate and β the false negative rate.

Proof. Let $\mathcal{R}$ be the critical region of the distinguisher. As in the proof of Theorem 6, the partition of $\mathcal{T}$ induces a partition $\mathcal{R} = \mathcal{R}_{\mathsf{g}} \sqcup \mathcal{R}_{\mathsf{b}}$. By definition,

$$1 - \beta = Q(\mathcal{R}) = Q(\mathcal{R}_{\mathsf{g}}) + Q(\mathcal{R}_{\mathsf{b}}) \leq Q(\mathcal{R}_{\mathsf{g}}) + \varepsilon_{\mathsf{b}}\,.$$

To deal with the term $Q(\mathcal{R}_{\mathsf{g}})$, we will divide by $\alpha = P(\mathcal{R})$. It can be assumed that $\alpha \neq 0$, since if $P(\mathcal{R}) = 0$ then also $Q(\mathcal{R}_{\mathsf{g}}) = 0$ so that the bound holds. Dividing by α gives

$$Q(\mathcal{R}_{\mathsf{g}}) = \alpha\,\frac{Q(\mathcal{R}_{\mathsf{g}})}{P(\mathcal{R})} = \alpha\,\frac{Q(\mathcal{R}_{\mathsf{g}})}{P(\mathcal{R}_{\mathsf{g}}) + P(\mathcal{R}_{\mathsf{b}})} \leq \alpha\,\frac{Q(\mathcal{R}_{\mathsf{g}})}{(1-\varepsilon_{\mathsf{g}})Q(\mathcal{R}_{\mathsf{g}}) + P(\mathcal{R}_{\mathsf{b}})}\,,$$

where we have used $P(\mathcal{R}_{\mathsf{g}}) \geq (1-\varepsilon_{\mathsf{g}})Q(\mathcal{R}_{\mathsf{g}})$. The result follows from the obvious lower bound $P(\mathcal{R}_{\mathsf{g}}) \geq 0$, since the result is trivial for $Q(\mathcal{R}_{\mathsf{g}}) = 0$ and

$$Q(\mathcal{R}_{\mathsf{g}}) \leq \alpha \frac{Q(\mathcal{R}_{\mathsf{g}})}{(1-\varepsilon_{\mathsf{g}})Q(\mathcal{R}_{\mathsf{b}})} \leq \frac{\alpha}{1-\varepsilon_{\mathsf{g}}}\,.$$

Plugging this into the bound $1 - \beta \leq Q(\mathcal{R}_{\mathsf{g}}) + \varepsilon_{\mathsf{b}}$ yields the result. □

There is an important difference in how we will apply Theorem 7 compared to typical uses of the H-coefficient theorem. In proofs of advantage bounds, Q is often chosen as the ideal world transcript distribution. This makes it easier to obtain the bound $Q(\mathcal{T}_{\mathsf{b}}) \leq \varepsilon_{\mathsf{b}}$, but it is not a good way to deal with bad events in the asymmetric setting. Instead, as we will see in Section 5, is is often easier to take P as the ideal world transcript distribution and exclude bad events using a hybrid argument.

More generally, it is our view that Theorem 7 should be used sparingly and with great care when proving power bounds. In particular, one should avoid using the term ε_{b} to deal with bad events that could have been excluded using a hybrid argument. Indeed, this is an additive term that worsens the multi-user bound. This contrasts with proofs of advantage bounds, where is it not uncommon to use Theorem 6 to exclude bad events from the ideal world that are just as likely to happen in the real world. In Section 5, we will encounter an example of this in the context of the Even-Mansour construction.

4.4 Hybrid with the H-coefficient technique

In this section, we provide a solution to the problem of the constant term ε_{b} introduced in Theorem 7. This approach works in general when the bad events only happen in one of the two worlds, but this can be achieved in many cases by redefining the events. Indeed, suppose $1-\beta \leq f_1(\alpha) \leq \alpha/(1-\varepsilon_{\mathsf{g}})$ for distinguishing from the ideal world *without bad events*. If $1-\beta \leq f_2(\alpha)$ for distinguishing the ideal world without bad events from the actual ideal world, then one gets

$$1-\beta \leq \frac{f_2(\alpha)}{1-\varepsilon_{\mathsf{g}}}\,.$$

For most of the practical constructions, if such bad events exist, then these usually only appear in the ideal world. However, our approach below can also handle the case when the bad events only appear in the real world.

Theorem 8. *Let P and Q be distributions on $\mathcal{T} = \mathcal{T}_{\mathsf{g}} \sqcup \mathcal{T}_{\mathsf{b}} \sqcup \mathcal{T}'_{\mathsf{b}}$. If $P(\mathcal{T}_{\mathsf{b}}) \leq \varepsilon_{\mathsf{b}}$, $Q(\mathcal{T}_{\mathsf{b}}) = 0$, $Q(\mathcal{T}'_{\mathsf{b}}) \leq \varepsilon'_{\mathsf{b}}$, and $P(E) \geq (1-\varepsilon_{\mathsf{g}})Q(E)$ for all $E \subseteq \mathcal{T}_{\mathsf{g}}$, then for every distinguisher from P to Q,*

$$1-\beta \leq \frac{\alpha}{1-\varepsilon_{\mathsf{b}}-\varepsilon_{\mathsf{g}}} + \varepsilon'_{\mathsf{b}}\,,$$

where α is the false positive rate and β the false negative rate.

Note that result is the general case that can be applied to most practical constructions. The goal in practice is to define the dominant bad events so that $\varepsilon'_{\mathsf{b}} = 0$ or at least very small.

Proof. Let $\mathcal{R}$ be the critical region of the distinguisher. As before, the partition of $\mathcal{T}$ induces a partition $\mathcal{R} = \mathcal{R}_{\mathsf{g}} \sqcup \mathcal{R}_{\mathsf{b}} \sqcup \mathcal{R}'_{\mathsf{b}}$. Define B as the transcripts where the 'bad events' occur, which is the set of transcripts that we would like to exclude

from consideration. If every distinguisher from P to $P_{\mathcal{T}\setminus B}$ satisfies $1-\beta_1 \leq f_1(\alpha_1)$ and every distinguisher from $P_{\mathcal{T}\setminus B}$ to Q satisfies $1-\beta_2 \leq f_2(\alpha_2)$, then based on Theorem 4 every distinguisher from P to Q satisfies

$$1-\beta \leq f_2(f_1(\alpha))\,.$$

Hence, it is sufficient to find f_1 for distributions P and $P_{\mathcal{T}\setminus B}$ and f_2 for distributions $P_{\mathcal{T}\setminus B}$ and Q. We will first start with f_1. By Theorem 5, we obtain

$$1-\beta_1 = P_{\mathcal{T}\setminus B}(\mathcal{R}) \leq \frac{\alpha_1}{1-\varepsilon_{\mathsf{b}}}\,.$$

Now it remains to find f_2, note that taking $B=\mathcal{T}_{\mathsf{b}}$, we have

$$1-\beta_2 = Q(\mathcal{R}) = Q(\mathcal{R}_{\mathsf{g}}) + Q(\mathcal{R}'_{\mathsf{b}}) \leq Q(\mathcal{R}_{\mathsf{g}}) + \varepsilon'_{\mathsf{b}}\,.$$

To deal with the term $Q(\mathcal{R}_{\mathsf{g}})$, we will divide by $\alpha_2 = P_{\mathcal{T}\setminus B}(\mathcal{R}\setminus\mathcal{T}_{\mathsf{b}})$. It can be assumed that $\alpha_2 \neq 0$, since if $P(\mathcal{R}\setminus\mathcal{T}_{\mathsf{b}})=0$ then also $Q(\mathcal{R}_{\mathsf{g}})=0$ so that the bound holds. Dividing by α_2 gives

$$Q(\mathcal{R}_{\mathsf{g}}) = \alpha_2\,\frac{Q(\mathcal{R}_{\mathsf{g}})}{P_{\mathcal{T}\setminus B}(\mathcal{R}\setminus\mathcal{T}_{\mathsf{b}})} = \alpha_2\,\frac{Q(\mathcal{R}_{\mathsf{g}})\cdot P(\mathcal{T}\setminus\mathcal{T}_{\mathsf{b}})}{P(\mathcal{R}\setminus\mathcal{T}_{\mathsf{b}})} \leq \alpha_2\,\frac{Q(\mathcal{R}_{\mathsf{g}})}{Q(\mathcal{R}_{\mathsf{g}})(1-\varepsilon_{\mathsf{g}})} = \frac{\alpha_2}{1-\varepsilon_{\mathsf{g}}}\,,$$

where we have used $P(\mathcal{R}\setminus\mathcal{T}_{\mathsf{b}}) = P(\mathcal{R}_{\mathsf{g}}\sqcup\mathcal{R}'_{\mathsf{b}}) \geq P(\mathcal{R}_{\mathsf{g}}) \geq (1-\varepsilon_{\mathsf{g}})Q(\mathcal{R}_{\mathsf{g}})$ and $P(\mathcal{T}\setminus\mathcal{T}_{\mathsf{b}})\leq 1$. Plugging f_1 and f_2 into Theorem 4, yields

$$1-\beta \leq \frac{\alpha}{(1-\varepsilon_{\mathsf{b}})(1-\varepsilon_{\mathsf{g}})} + \varepsilon'_{\mathsf{b}} \leq \frac{\alpha}{1-\varepsilon_{\mathsf{b}}-\varepsilon_{\mathsf{g}}} + \varepsilon'_{\mathsf{b}}\,,$$

using the fact that $1/(1-x)\times 1/(1-y) \leq 1/(1-x-y)$. □

4.5 Converting power bounds to advantage bounds

In Section 3.2, it was already explained how power bounds can be turned into advantage bounds in general. If the methods from Sections 4.2 to 4.4 are used, then one often obtains a power bound for distinguishers from P to Q of the form

$$1-\beta \leq \frac{\alpha}{1-\varepsilon_1} + \varepsilon_2\,, \tag{1}$$

with $\varepsilon_2 = 0$ in the best case. Since the bounds we obtain in Sections 5 and 6 are all of this form, it is useful to show here that (1) implies

$$\Delta\big(P\,;\,Q\big) \leq \varepsilon_1 + \varepsilon_2\,.$$

To see this, note that the power is at most one. Hence, by applying the general principle from Section 3.2 to (1), we get

$$\Delta\big(P\,;\,Q\big) \leq \max_{\alpha\,\in\,[0,1]} \min\left\{\frac{\alpha}{1-\varepsilon_1}+\varepsilon_2-\alpha,\ 1\right\} \leq \max_{\alpha\,\in\,[0,1]} \min\left\{\frac{\alpha}{1-\varepsilon_1}-\alpha,\ 1\right\}+\varepsilon_2\,,$$

Since $1-\varepsilon_1 \leq 1$, the maximum is achieved at $\alpha = 1-\varepsilon_1$. Hence, the advantage is upper bounded by $1-(1-\varepsilon_1)+\varepsilon_2 = \varepsilon_1+\varepsilon_2$.

5 Applications

In this section, we apply the proof techniques from Section 4 to more complicated constructions. More precisely, in Section 5.1 we derive power bounds for the Even-Mansour construction [11] using a simple hybrid argument. Compared to traditional proofs of advantage bounds for this construction such as [19], our bad events are defined in a more careful way so that the H-coefficient method is not necessary. In Section 5.2, we use the more powerful H-coefficient technique together with a hybrid argument to obtain power bounds for the sum-of-permutations construction [4, 15].

5.1 Even-Mansour construction

The Even-Mansour EM construction [11] is used to build a secure pseudorandom permutation based on a public permutation. We consider the version that only uses a single key, but our analysis carries over to the case with two keys.

Let n be a non-negative integer and π a permutation on $\{0,1\}^n$. The Even-Mansour construction is a family of permutations $\mathsf{EM}_K[\pi] : \{0,1\}^n \to \{0,1\}^n$ indexed by a key K in $\{0,1\}^n$ and based on the permutation π. Given a plaintext M, the corresponding ciphertext is equal to

$$\mathsf{EM}_K[\pi](M) = \pi(M \oplus K) \oplus K\,,$$

with $\oplus$ the exclusive-or operation on bitstrings of length n. Theorem 9 gives an upper bound on the power of any distinguisher for the EM construction, with EM as the alternative hypothesis. The security model for the Even-Mansour construction is the standard sprp security notion with oracle access to the public permutation π, so we do not describe it in detail here. Queries to made to π are called primitive queries, the others are construction queries.

Theorem 9. *Let n be a non-negative integer, $N = 2^n$, K a uniform random variable on $\{0,1\}^n$ and π a uniform random permutation on $\{0,1\}^n$. For every distinguisher (with oracle access to π) from a uniform random permutation to $\mathsf{EM}_K[\pi]$ that makes q construction queries and p primitive queries,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{2pq}{N}}\,,$$

where α is the false positive rate and β the false negative rate.

Compares to the bound obtained from the advantage [11, 16], which corresponds to

$$1 - \beta \leq \alpha + \frac{2pq}{N}\,.$$

As the case of PRP-PRF switching lemma, the gap between both bounds is proportional to $1 - \alpha$. As before, it is considerably more difficult to achieve a low error probability α than one might expect from the advantage bound. While for our case, low α means that the power cannot be too high.

Proof of Theorem 9. Since the resulting bound is non-decreasing in q and p, we can assume that the distinguisher never queries duplicate inputs. Every transcript contains the result of at most q construction queries and at most p primitive queries, both bidirectional and suitably ordered. We also include the key in the transcript; this can only improve the distinguisher. In terms of the security game, the keys are disclosed after the distinguisher finishes its interaction with the oracles but before it outputs its decision. In the ideal world the key is a uniform random dummy.

Let P denote the distribution of transcripts for the uniform random permutation and Q the distribution of transcripts for Even-Mansour. Intuitively, the property that distinguishes Even-Mansour from a uniform random permutation is that if the input of a construction query collides with the input of a primitive query, then the output of this construction query must collide with the output of the primitive query. One has a similar property for the outputs.

Formally, let B be the set of all transcripts containing a construction query (M, C) and a primitive query (u, v) so that one of the following conditions holds:

$$\mathsf{BAD}_1 \colon M \oplus u = K \text{ and } C \oplus v \neq K\,,$$
$$\mathsf{BAD}_2 \colon M \oplus u \neq K \text{ and } C \oplus v = K\,.$$

Readers familiar with the proof of Mouha and Luykx [19] or other proofs in the advantage setting will note that our bad events are defined differently, and this is important. The goal of our bad events is to bring the ideal world closer to the real world. Hence, we only want to exclude events that do not happen in the real world. In particular, we use a hybrid argument with $P_{\mathcal{T}\setminus B}$ as the intermediate transcript distribution.

To apply Theorem 5, we need to bound $P(B)$. Since the dummy key is uniform random and independent of all queries, there are at most pq keys satisfying BAD_1. A similar argument can be made for BAD_2. Hence, by the union bound,

$$P(B) \leq \frac{2pq}{N}\,.$$

Hence, by Theorem 5, for all distinguishers from P to $P_{\mathcal{T}\setminus B}$,

$$1 - \beta \leq \frac{\alpha}{1 - \frac{2pq}{N}} := f(\alpha)\,,$$

with α the false positive rate and β the false negative rate. To complete the proof, we still need to verify that $P_{\mathcal{T}\setminus B} = Q$. First, it is easy to see that $P_{\mathcal{T}\setminus B}$ and Q have the same support. The result then follows from the observation that all transcripts in the support have the same probability. The theorem follows by by applying Theorem 4 with $g(\alpha) = \alpha$ and f as above. □

5.2 Sum-of-permutations construction

The sum-of-permutations construction was first proposed by Bellare et al. [4] and by Hall et al. [15]. This construction is used to build a pseudorandom function

from a block cipher. In this section, we consider the variant based on a single permutation with domain separation. The easier variant with two independent block ciphers is analyzed in Appendix A

Let n be a non-negative integer and π a permutation on n bits. The sum-of-permutations construction is a family of functions $\mathsf{SoP}[\pi] : \{0,1\}^{n-1} \to \{0,1\}^n$ based on the (secret) permutation π. For an input M in $\{0,1\}^{n-1}$, it returns

$$\mathsf{SoP}[\pi](M) = \pi(0\|M) \oplus \pi(1\|M)\,.$$

In practice, π is instantiated with a block cipher and one relies on its prp security to replace it by a uniform random permutation using a hybrid argument.

The security model for the sum-of-permutations construction is the standard prf security security notion, i.e. the construction is compared to a uniform random function on $\{0,1\}^{n-1}$. Theorem 10 gives an upper bound on the power of any test that distinguishes SoP from a uniform random function.

Theorem 10. *Let $N = 2^n$ with $n > 12$. For every distinguisher from a uniform random function to the $\mathsf{SoP}[\pi]$ construction instantiated with a uniform random permutation π, making $q \leq N/4n$ queries,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{q}{N} - \frac{25q^2}{N^2} - \frac{49(n+1)^2}{N}} + \left(\frac{2q}{N}\right)^n\,,$$

where α is the false positive rate and β the false negative rate.

Note that the term $(q/N)^n$ in the security bound is due to the fact that a variant of the mirror theory result of Choi et al. [8] is used in the proof below for the good transcript analysis. This term might be avoided by using a stronger variant of mirror theory, but this remains a conjecture. In any case, the term is negligible and does not pose significant issues for our multi-user result in Section 6.

Proof. The transcript consists of q input-output pairs. Since the resulting bound is non-decreasing in q, we can assume that all the inputs are distinct. Let P denote the distribution of transcripts for a uniform random function ρ, and Q the distribution for $\mathsf{SoP}[\pi]$.

Intuitively, the only difference between ρ and $\mathsf{SoP}[\pi]$ is that the zero bitstring 0^n may be output by ρ but not by $\mathsf{SoP}[\pi]$. Formally, let B be the set of all transcripts that contain a zero bitstring. Importantly, since $Q(B) = 0$, Theorem 8 can be applied with $\mathcal{T}_{\mathsf{b}} = B$. Since every output of ρ appears with probability $1/N$, we have

$$P(B) \leq \frac{q}{N}\,.$$

To obtain a lower bound on the ratio, we will rely on the result[2] of Choi et al. [8]. However, due to the fact that a special variant of the mirror theory is

[2] Choi et al. prove the result for a modified ideal world case, where the ideal world never returns a zero output. However, their result can be straightforwardly extended to the general situation.

used in the work of Choi et al. [8], an additional technical bad event is needed in order to apply their result. More precisely, let B' be the set of all transcripts that contain outputs $C_1, \ldots, C_n$ (not necessarily consecutive) such that

$$C_1 = \cdots = C_n.$$

Hence, we will apply Theorem 8 with $\mathcal{T}_{\mathsf{b}} = B$, $\mathcal{T}'_{\mathsf{b}} = B'$ and $\mathcal{T}_{\mathsf{g}} = \mathcal{T} \setminus (B \cup B')$. For this, we need to bound B' in the *real world*. We have ‘

$$Q(B') \leq \frac{\binom{q}{n}}{(2^n - q)^{n-1}} \leq \frac{q^n}{n!(2^{n-1})^{n-1}} \leq \frac{q^n}{2^{(n-1)n}} = \left(\frac{2q}{N}\right)^n ,$$

because $q \leq 2^{n-1}$, $n! \geq 2^{n+1}$ and $2^{n+1} \cdot (2^{n-1})^{n-1} \geq (2^{n-1})^n$. Focusing on the first term in the proof of Theorem 10 of [8] with $u = 1$, and using that $n > 12$ and $q \leq N/4n$,

$$\varepsilon_{\mathsf{g}} \leq \frac{25q^2}{N^2} + \frac{49(n+1)^2}{N} .$$

Hence, with $\varepsilon_{\mathsf{b}} = q/N$ and $\varepsilon'_{\mathsf{b}} = (2q/N)^n$, Theorem 8 gives

$$1 - \beta \leq \frac{\alpha}{1 - \varepsilon_{\mathsf{b}} - \varepsilon_{\mathsf{g}}} + \varepsilon'_{\mathsf{b}} = \frac{\alpha}{1 - \frac{q}{N} - \frac{25q^2}{N^2} - \frac{49(n+1)^2}{N}} + \left(\frac{2q}{N}\right)^n$$

This completes the proof. □

6 Multi-user security

In this section, we propose two methods to convert single-user power bounds into multi-user power bounds. The first method is based on the idea of point-wise proximity due to Hoang and Tessaro [16], and is described in Section 6.1. However, one limitation of this method is that the power bound must satisfy some nontrivial conditions. Therefore, in Section 6.2, we propose another method that requires users to be independent but that avoids the limitations of point-wise proximity.

6.1 From point-wise proximity

In this section, we present a first method to convert single-user power bounds into multi-user power bounds. It based on the point-wise proximity method of Hoang and Tessaro [16]. Unlike the method that we propose in Section 6.2, it is also applicable when there is some dependency between different users, such as a shared primitive. However, it has limitations of its own: the power bound must be linear and the dependence on the number of queries must satisfy some nontrivial conditions. In addition, the result in Section 6.2 yields a slightly better bound.

Let us recall the definition of point-wise proximity [16, Definition 1]. Two probability mass functions p and q on a set $\mathcal{T}$ satisfy ε-point-wise proximity if, for all x in $\mathcal{T}$,

$$q(x) - p(x) \leq \varepsilon\, q(x)\,.$$

Note that this is equivalent to a ratio bound of the form $p(x)/q(x) \geq 1 - \varepsilon$, the same as in the H-coefficient method (see Section 4.3). The following result shows that *in some cases* a power bound implies point-wise proximity. The converse is also true: point-wise proximity always implies a strong power bound. This follows from Theorem 7 with $\varepsilon_{\mathsf{b}} = 0$. However, direct point-wise proximity estimates are more difficult to obtain because this requires analyzing the probabilities of individual transcripts.

Theorem 11. *If every distinguisher from P to Q with false positive and negative rates α and β satisfies $1 - \beta \leq \alpha/(1-\varepsilon)$, then the probability mass functions of P and Q satisfy ε-point-wise proximity.*

Proof. By the given power bound, we have $Q(\mathcal{R}) \leq P(\mathcal{R})/(1-\varepsilon)$ for all critical regions $\mathcal{R}$, including $\mathcal{R} = \{x\}$. Hence, if p and q are the probability mass functions of P and Q respectively, then

$$(1-\varepsilon)\, q(x) \leq p(x)\,.$$

Equivalently, $q(x) - p(x) \leq \varepsilon\, q(x)$. □

From the point of view of Theorem 11, power bounds may be used as an efficient way to prove point-wise proximity estimates. For example, a typical point-wise proximity estimate for the PRP-PRF switching problem is not to different from the proof of Theorem 3 – computing the probability of every transcript. However, to derive a power bound, we can simply exclude collisions using a hybrid argument as in Example 3.

In this section, our main interest in Theorem 11 is that Hoang and Tessaro [16, §3.3] have shown that a single-user point-wise proximity estimate can be turned into a multi-user point-wise proximity estimate.

Multi-user security of EM. To illustrate how this works, consider the multi-user security of the Even-Mansour construction. The bound follows immediately from the single-user bound, since the proof of Theorem 9 already established point-wise proximity (by Theorem 11). This is formalized in Theorem 12.

Theorem 12. *Let n be a non-negative integer, $N = 2^n$, K a uniform random variable on $\{0,1\}^n$ and π a uniform random permutation on $\{0,1\}^n$. For every distinguisher (with oracle access to π) from a uniform random permutation to $\mathsf{EM}_{K_i}[\pi]$ construction that makes in total q construction queries to its u user for $i = 1, \ldots, u$ and p primitive queries,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{4q(p+q)}{N}}\,.$$

where α is the false positive rate and β the false negative rate.

Proof. Let us recall from the proof of Theorem 9 that P is the distribution of transcripts for ideal world (π_I, π) and Q is the distribution of transcripts for real world $(\mathsf{EM}_K[\pi], \pi)$. From Theorem 9 and Theorem 11, the probability mass functions of P and Q satisfy $\epsilon(p,q)$-point-wise proximity with $\epsilon(p,q) = \frac{2qp}{N}$. Since this bound satisfies (i) $\epsilon(x,y) + \epsilon(x,z) \leq \epsilon(x, y+z)$, for all non-negative integers x, y, z, and (ii) $\epsilon(\cdot, z)$ and $\epsilon(z, \cdot)$ are non-decreasing functions for every non-negative integer z, [16, Lemma 3] give us that

$$q(x) - p(x) \leq 2 \cdot \frac{2q(p+q)}{N}\, q(x)\,.$$

The result follows since point-wise proximity implies a strong power bound. □

We note that this bound is essentially the same as the one from Theorem 9, with an additional factor two and the additive term q. This additive term plays a significant role for the case of the EM construction. The $O(q^2/N)$ term takes into account collisions on the keys across multiple users, which allows to easily distinguish and is therefore tight.

We would like to stress once more the importance of giving security using power bounds, as opposed to advantage bounds. The bound obtained from the dedicated analysis by Mouha and Luykx (ML) [19] corresponds to

$$1 - \beta \leq \alpha + \frac{2q(p+q)}{N}\,,$$

and the bound obtained by naively using the hybrid argument in the classical advantage setting [16] on the single-user result corresponds to

$$1 - \beta \leq \alpha + \frac{2uq(p+q)}{N}\,.$$

The gap between these bounds and ours is proportional to $1 - \alpha$. The lower the value of α, the more important the power bound becomes compared to the advantage bounds.

6.2 Independent users

In this section, we provide an alternative way to deduce multi-user power bounds from single-user power bounds. It requires the assumption that users are independent, but it gives a tighter bound than point-wise proximity and more importantly does not make any assumptions about how the single-user bound depends on the number of queries. It is, however, necessary to assume that the single-user power bound is linear in the false positive rate. That is, $f(\alpha) \leq a(q)\alpha$, with a an arbitrary function of the number of queries q. This means that our result in this section can also be understood as a way to obtain point-wise proximity bounds without assumptions on $\varepsilon(q) = 1 - 1/a(q)$.

In the multi-user setting with u users, it can be assumed that P and Q are distributions on a set $\mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2 \times \cdots \times \mathcal{T}_u$. Although this implies that queries

for different users are kept in separate parts of the transcript, this does not mean that we neglect their order. Indeed, the queries may be numbered. If the u users are independent, then P and Q are related to the marginal single user distributions $P_1, \ldots, P_u$ and $Q_1, \ldots, Q_u$ as follows:

$$P(E_1 \times E_2 \times \cdots \times E_u) = P_1(E_1)P_2(E_2)\cdots P_u(E_u),$$
$$Q(E_1 \times E_2 \times \cdots \times E_u) = Q_1(E_1)Q_2(E_2)\cdots Q_u(E_u).$$

This will be denoted by $P = P_1 \otimes P_2 \otimes \cdots \otimes P_u$ and $Q = Q_1 \otimes Q_2 \otimes \cdots \otimes Q_u$.

In the multi-user setting, it important to be precise about the dependence of single-user bound depends on the number of queries. For this reason, in the following results, we always mention the number of queries made by the distinguisher and write $1 - \beta \leq f(\alpha; q)$ if that number is q. From the point of view of probability theory, this is simply a restriction on the critical region of the hypothesis test. The precise definition of this restriction does not actually matter for the theorems that we prove.

In the remainder of this section, we prove Theorem 13. Despite the independence assumption, which is rather strong from a probability theory point of view, the proof is subtle. There are two issues that we have to address:

(i) Although users are independent, the critical region is *not* a product of subsets of $\mathcal{T}_1, \ldots, \mathcal{T}_u$. Indeed, from the Neyman-Pearson lemma (Theorem 1), we see that there is no reason to expect this for the optimal test. This corresponds to the fact that the adversary can, of course, use the result of a query for one user as the input to a query for another user.

(ii) Although the total number of queries is fixed to q, we do not know the number of queries made to each user individually. This means that the critical region can contain, *for every user*, transcripts with up to q queries. In fact, this is precisely why a naive hybrid argument leads to a loss proportional to the number of users.

The main idea of the proof is to establish the result first under the assumption that every transcript in the critical region contains exactly q_i queries to user i. This is easy, and essentially follows the point-wise proximity proof. To deal with the second problem above, we partition the critical region into subsets that each contain only transcripts with a specific number of queries to every user. The new idea here is that power bounds apply to distributions, which is different from the perspective of probability mass functions that is adopted when using point-wise proximity. This allows us to keep track of the probability of all the sets in the partition, and is the reason why our result does not require any assumptions (such as super-linearity) on the query-dependence.

Theorem 13. *For $i = 1, \ldots, u$, let P_i and Q_i be distributions on the same domain. If every distinguisher from P_i to Q_i making q queries satisfies $1 - \beta_i \leq a_i(q)\,\alpha_i$ for false positive and negative rates α_i and β_i, then for every distinguisher from $P_1 \otimes P_2 \otimes \cdots \otimes P_u$ to $Q_1 \otimes Q_2 \otimes \cdots \otimes Q_u$ making q queries,*

it holds that

$$1 - \beta \leq \alpha \max_{q_1 + \cdots + q_u = q} \prod_{i=1}^{q} a_i(q_i)\,,$$

with α the false positive rate and β the false negative rate.

The following lemma establishes the first part of our claim, namely a power bound for distinguishers that make a predetermined number of queries to every user.

Lemma 1. *For $i = 1, \ldots, u$, let P_i and Q_i be distributions with the same domain. If every distinguisher from P_i to Q_i making q queries satisfies $1 - \beta_i \leq a_i(q)\,\alpha_i$ for false positive and negative rates α_i and β_i, then for every distinguisher from $P_1 \otimes P_2 \otimes \cdots \otimes P_u$ to $Q_1 \otimes Q_2 \otimes \cdots \otimes Q_u$ that makes q_i queries to user i,*

$$1 - \beta \leq \alpha \prod_{i=1}^{u} a_i(q_i)\,,$$

with α the false positive rate and β the false negative rate.

Proof. Theorem 11 shows that the bounds $1 - \beta_i \leq a_i(q)\alpha_i$ establish point-wise proximity between P_i and Q_i. Put more simply, for every event E_i,

$$Q_i(E_i) \leq a_i(q_i)\, P_i(E_i)\,.$$

Since $P = P_1 \otimes \cdots \otimes P_u$ and $Q = Q_1 \otimes \cdots \otimes Q_u$, multiplying these bounds gives

$$Q(E_1 \times \cdots \times E_u) \leq \prod_{i=1}^{u} a_i(q_i)\, P(E_1 \times \cdots \times E_u)$$

Since this bound works when the events E_i are singletons, so it works for every transcript. Hence, if $\mathcal{R}$ is the critical region, then also $Q(\mathcal{R}) \leq \alpha\, P(\mathcal{R})$. □

The first part of the proof is now complete. Using Lemma 1, we can now proceed with the partition argument.

Proof of Theorem 13. The idea of the theorem is to apply Lemma 1. However, we must first partition the critical region $\mathcal{R}$ according to the number of queries that were made for each user. Let $\mathcal{R}_{q_1,\ldots,q_u}$ be the subset of $\mathcal{R}$ containing all transcripts with q_i queries for user i. This gives the following partition of $\mathcal{R}$:

$$\mathcal{R} = \bigsqcup_{q_1 + \cdots + q_u = q} \mathcal{R}_{q_1,\ldots,q_u}\,.$$

In terms of probabilities, we have

$$1 - \beta = Q(\mathcal{R}) = \sum_{q_1 + \cdots + q_u = q} Q(\mathcal{R}_{q_1,\ldots,q_u})\,,$$

where $Q = Q_1 \otimes \cdots \otimes Q_u$. Lemma 1 shows that

$$Q(\mathcal{R}_{q_1,\ldots,q_u}) \le P(\mathcal{R}_{q_1,\ldots,q_u}) \prod_{i=1}^{u} a_i(q_i)\,.$$

By taking the maximum and using linearity, we obtain

$$1-\beta \le \left(\max_{q_1+\cdots+q_u=q} \prod_{i=1}^{u} a_i(q_i) \right) \sum_{q_1+\cdots+q_u=q} P(\mathcal{R}_{q_1,\ldots,q_u}) \le \alpha \max_{q_1+\cdots+q_u=q} \prod_{i=1}^{u} a_i(q_i)\,.$$

This the desired result. □

Note that our result achieves a better point-wise proximity estimate compared to the results of Hoang and Tessaro [16]. Indeed, in our case, the function $a(q)$ can be an arbitrary function, while the $\varepsilon(q)$ function in [16] needs to be super-linear. The main reason is because we maximize over all $q_1 + \cdots + q_u = q$ to pull $\prod_{i=1}^{u} a_i(q_i)$ out of the summation in the last step of our proof, which is natural from the point of view of power bounds. Indeed, in the case when $\varepsilon(q) = q/N + c$ with c some small constant, then Hoang and Tessaro's result does not hold any more since $q_1/N + c + q_2/N + c > (q_1 + q_2)/N + c$, whereas our result still works.

To illustrate Theorem 13, we prove a multi-user variant of the PRP-PRF switching lemma. The power bound we prove in the following example implies the strongest possible advantage bound. This is unlike the pointwise proximity method of Hoang and Tessaro [16], which loses a factor of two.

Example 4. In the multi-user variant of the PRP-PRF switching lemma, we have independent uniform random permutations $\pi_1, \ldots, \pi_u$ and independent uniform random functions $\rho_1, \ldots, \rho_u$. In Theorem 3 (and in Example 3) we showed that for all distinguishers from ρ_i to π_i making at most q_i queries,

$$1 - \beta_i \le \frac{\alpha_i}{1 - \frac{q_i(q_i-1)}{2N}}\,,$$

where α_i is the false positive rate and β_i the false negative rate. By Theorem 13, we get the overall power bound

$$1 - \beta \le \alpha \max_{q_1+\cdots+q_u \le q} \prod_{i=1}^{q} \frac{1}{1 - \frac{q_i(q_i-1)}{2N}} \le \max_{q_1+\cdots+q_u \le q} \frac{\alpha}{1 - \sum_{i=1}^{u} \frac{q_i(q_i-1)}{2N}}\,.$$

Here we have used the fact that $1/(1-x) \times 1/(1-y) \le 1/(1-x-y)$. Clearly, $\sum_{i=1}^{u} q_i(q_i-1) \le q(q-1)$. Hence, we get the bound

$$1 - \beta \le \frac{\alpha}{1 - \frac{q(q-1)}{2N}}\,,$$

which is exactly the same as the single-user bound. ▷

6.3 Multi-user security of SoP.

Finally, we consider the multi-user security of SoP. The easier variant with two independent block ciphers is analyzed in Appendix B. In the multi-user variant of the SoP construction, we have SoP based on independent uniform random permutations $\mathsf{SoP}[\pi_1], \ldots, \mathsf{SoP}[\pi_u]$ and independent uniform random functions $\rho_1, \ldots, \rho_u$. In Theorem 10 we showed that for all distinguishers from ρ_i to $\mathsf{SoP}[\pi_i]$ making at most q_i queries,

$$1 - \beta_i \leq \frac{\alpha_i}{1 - \frac{q_i}{N} - \frac{25q_i^2}{N^2} - \frac{49(n+1)^2}{N}} + \left(\frac{2q}{N}\right)^n ,$$

where α_i is the false positive rate and β_i the false negative rate. In order to be able to apply Theorem 13, we must avoid the constant term above. This is possible by excluding the relevant (technical) bad event in the multi-user setting. This leads to a factor of u loss for this term, but this does matter as the term is comparatively small.

Let $\mathcal{R} = \mathcal{R}_\mathsf{g} \sqcup \mathcal{R}_\mathsf{b}$ be the critical region of the multi-user distinguisher, then based on Theorem 13 (which applies with arbitrary restrictions on the critical region) we have

$$1 - \beta \leq Q(\mathcal{R}_\mathsf{g}) + Q(\mathcal{R}_\mathsf{b}) \leq P(\mathcal{R}_\mathsf{g}) \max_{q_1 + \cdots + q_u = q} \prod_{i=1}^{u} a_i(q_i) + Q(\mathcal{R}_\mathsf{b}) .$$

Since $P(\mathcal{R}_\mathsf{g}) \leq P(\mathcal{R}) = \alpha$ and using the fact that the $a_i(q_i)$'s are non-decreasing functions, the above equation becomes

$$1 - \beta \leq \alpha \max_{q_1 + \cdots + q_u = q} \prod_{i=1}^{u} a_i(q_i) + Q(\mathcal{R}_\mathsf{b}) .$$

Recall that B' is the set of all transcripts that contain outputs $C_1, \ldots, C_n$ such that $C_1 = \cdots = C_n$, as defined in the proof of Theorem 10. Let $\mathcal{T}_i$ be the set of all possible transcripts of user i and $\mathcal{T}$ the set of all possible transcripts of all users, then $\mathcal{R}_\mathsf{b} = \mathcal{T} \setminus \prod_{i=1}^{u} (\mathcal{T}_i \setminus B')$. Hence

$$Q(\mathcal{R}_\mathsf{b}) \leq u \left(\frac{2q}{N}\right)^n ,$$

From Theorem 10, but now focusing on the term multiplied by α, we obtain

$$\begin{aligned}
&\alpha \max_{q_1 + \cdots + q_u \leq q} \prod_{i=1}^{u} \frac{1}{1 - \frac{q_i}{N} - \frac{25q_i^2}{N^2} - \frac{49(n+1)^2}{N}} \\
&\qquad \leq \max_{q_1 + \cdots + q_u \leq q} \frac{\alpha}{1 - \sum_{i=1}^{u} \left(\frac{q_i}{N} - \frac{25q_i^2}{N^2} - \frac{49(n+1)^2}{N}\right)} \\
&\qquad \leq \frac{\alpha}{1 - \frac{q}{N} - \frac{25q^2}{N^2} - \frac{49u(n+1)^2}{N}} ,
\end{aligned}$$

using the fact that $1/(1-x) \times 1/(1-y) \leq 1/(1-x-y)$. Therefore, we obtain the following theorem.

Theorem 14. *Let $N = 2^n$ with $n > 12$. For every distinguisher from a uniform random function to the* SoP$[\pi]$ *construction instantiated with a uniform random permutation π, making in total $q \leq N/4n$ to its u user oracles,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{q}{N} - \frac{25q^2}{N^2} - \frac{49u(n+1)^2}{N}} + u \left(\frac{2q}{N}\right)^n .$$

where α is the false positive rate and β the false negative rate.

Assuming that $q \leq N/4n$, the term $u\ (2q/N)^n$ becomes negligible. Therefore, this bound is essentially the same as the one from Theorem 10, except that the term $(n+1)^2/N$ in the single-user case now becomes $u(n+1)^2/N$. This additional u related term is probably an artifact of the proof technique used to obtain the single-user bound rather than reflecting a true security loss when using power bound, and may be avoided by improving the single-user security bound of SoP.

To our knowledge, this is the first time a multi-user security of the SoP construction is proven, where the dominating terms do not depend on both q and u. This is a significant improvement over the previous result by Choi et al. [8], which corresponds to

$$1 - \beta \leq \alpha + \frac{26u^{1/2}q^2}{N^2} + \frac{49u^{1/2}(n+1)^2}{N} ,$$

and the bound obtained by naively using the hybrid argument on the single-user results obtained by Dutta et al. [10]:

$$1 - \beta \leq \alpha + \frac{uq}{N} .$$

It is worth highlighting that there is currently no general solution for the multi-user security of the SoP construction using other methods. As Chen et al. [7] have already pointed out, this is because the fact that the advantage bound of SoP is dominated by the bad event that the ideal world returns a zero output, while this cannot happen in the real world. This event gives a term uq/N for the multi-user security using the squared-ratio method [7]. Choi et al.'s result is based on a different security model where the ideal world is forbidden to return 0^n as output. It is not certain whether this assumption about the model can have a major impact on practical applications. When using power bounds, no additional assumptions need to be made.

References

1. Bellare, M., Boldyreva, A., Micali, S.: Public-key encryption in a multi-user setting: Security proofs and improvements. In: EUROCRYPT 2000. LNCS, vol. 1807, pp. 259–274
2. Bellare, M., Desai, A., Jokipii, E., Rogaway, P.: A concrete security treatment of symmetric encryption. In: 38th FOCS. pp. 394–403

3. Bellare, M., Kilian, J., Rogaway, P.: The security of cipher block chaining. In: CRYPTO'94. LNCS, vol. 839, pp. 341–358
4. Bellare, M., Krovetz, T., Rogaway, P.: Luby-Rackoff backwards: Increasing security by making block ciphers non-invertible. In: EUROCRYPT'98. LNCS, vol. 1403, pp. 266–280
5. Bellare, M., Rogaway, P.: The security of triple encryption and a framework for code-based game-playing proofs. In: EUROCRYPT 2006. LNCS, vol. 4004, pp. 409–426
6. Chen, S., Steinberger, J.P.: Tight security bounds for key-alternating ciphers. In: EUROCRYPT 2014. LNCS, vol. 8441, pp. 327–350
7. Chen, Y.L., Choi, W., Lee, C.: Improved multi-user security using the squared-ratio method. In: CRYPTO 2023, Part II. LNCS, vol. 14082, pp. 694–724
8. Choi, W., Hhan, M., Wei, Y., Zikas, V.: Fine-tuning ideal worlds for the xor of two permutation outputs. IACR Cryptol. ePrint Arch. p. 1704
9. Dai, W., Hoang, V.T., Tessaro, S.: Information-theoretic indistinguishability via the chi-squared method. In: CRYPTO 2017, Part III. LNCS, vol. 10403, pp. 497–523
10. Dutta, A., Nandi, M., Saha, A.: Proof of mirror theory for $\xi_{\max} = 2$. IEEE Trans. Inf. Theory **68**(9), 6218–6232
11. Even, S., Mansour, Y.: A construction of a cipher from a single pseudorandom permutation. In: ASIACRYPT'91. LNCS, vol. 739, pp. 210–224
12. Ghoshal, A., Jaeger, J., Tessaro, S.: The memory-tightness of authenticated encryption. In: CRYPTO 2020, Part I. LNCS, vol. 12170, pp. 127–156
13. Ghoshal, A., Tessaro, S.: The query-complexity of preprocessing attacks. In: CRYPTO 2023, Part II. LNCS, vol. 14082, pp. 482–513
14. Goldwasser, S., Micali, S.: Probabilistic encryption and how to play mental poker keeping secret all partial information. In: 14th ACM STOC. pp. 365–377
15. Hall, C., Wagner, D., Kelsey, J., Schneier, B.: Building PRFs from PRPs. In: CRYPTO'98. LNCS, vol. 1462, pp. 370–389
16. Hoang, V.T., Tessaro, S.: Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In: CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 3–32
17. Hoang, V.T., Tessaro, S.: The multi-user security of double encryption. In: EUROCRYPT 2017, Part II. LNCS, vol. 10211, pp. 381–411
18. Impagliazzo, R., Rudich, S.: Limits on the provable consequences of one-way permutations. In: CRYPTO'88. LNCS, vol. 403, pp. 8–26
19. Mouha, N., Luykx, A.: Multi-key security: The Even-Mansour construction revisited. In: CRYPTO 2015, Part I. LNCS, vol. 9215, pp. 209–223
20. Neyman, J., Pearson, E.S.: IX. On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **231**(694-706), 289–337
21. Patarin, J.: The "coefficients H" technique (invited talk). In: SAC 2008. LNCS, vol. 5381, pp. 328–345
22. Rogaway, P., Shrimpton, T.: A provable-security treatment of the key-wrap problem. In: EUROCRYPT 2006. LNCS, vol. 4004, pp. 373–390

Supplementary Martial

A The SoP2 construction

In this section, we consider the SoP2 construction mentioned in Section 5.2. Here, in particular, we will consider a version that involves two independent permutations.

Let n a non-negative integer and π_1, π_2 permutations on n bits. The sum-of-two-independent-permutations construction is a family of functions $\mathsf{SoP2}[\pi_1, \pi_2] : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ based on the (secret) permutations π_1, π_2. For an input M in $\{0,1\}^n$, it returns

$$\mathsf{SoP2}[\pi_1, \pi_2](M) = \pi_1(M) + \pi_2(M)\,.$$

In practice, π_1, π_2 are instantiated with a block cipher and one relies on their prp security to replace these by uniform random permutations using a hybrid argument.

Theorem 15 gives an upper bound on the power of any test that distinguishes SoP2 from a uniform random function.

Theorem 15. *Let $N = 2^n$ with $n > 17$. For every distinguisher from a uniform random function to the* SoP2 *construction instantiated with a uniform random permutations π_1, π_2, making $q \leq N/17$ queries,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{19q^2}{N^2} - \frac{8n^3}{N^2}}\,.$$

where α is the false positive rate and β the false negative rate.

Proof. The transcript consists of q input-output pairs. Since the resulting bound is non-decreasing in q, we can assume that all the inputs are distinct. Let P denote the distribution of transcripts for a uniform random function ρ, and Q the distribution for $\mathsf{SoP2}[\pi_1, \pi_2]$.

Unlike the single-keyed case, this time we can apply theorem 7, with P the distribution of transcripts for the SoP2 construction and Q the distribution of transcripts for the uniform random function. Hence, we have

$$1 - \beta \leq \frac{\alpha}{1 - \varepsilon_{\mathtt{g}}}\,.$$

since for SoP2 we have $\varepsilon_{\mathtt{b}} = 0$ (there are no bad events) and therefore we can just use the mirror theory result with lower bound of [10] (Corollary 2) to obtain $\varepsilon_{\mathtt{g}}$. We have

$$\varepsilon_{\mathtt{g}} = \frac{19q^2}{N^2} + \frac{8\log(N)^3}{N^2}\,.$$

for $n > 17$ and $q \leq N/17$. Hence, we have

$$1 - \beta \leq \frac{\alpha}{1 - \frac{19q^2}{N^2} - \frac{8\log(N)^3}{N^2}}\,.$$

This completes the proof. □

B Multi-user security of SoP2.

We consider the multi-user security of SoP2. In the multi-user variant of the SoP2 construction, we have SoP2 based on independent uniform random permutations $\mathsf{SoP}[\pi_1, \pi_2], \ldots, \mathsf{SoP}[\pi_{2u-1}, \pi_{2u}]$ and independent uniform random functions $\rho_1, \ldots, \rho_u$. In Theorem 15 we showed that for all distinguishers from ρ_i to $\mathsf{SoP}[\pi_i]$ making at most q_i queries,

$$1 - \beta_i \leq \frac{\alpha_i}{1 - \frac{19q_i^2}{N^2} - \frac{8n^3}{N^2}} ,$$

where α_i is the false positive rate and β_i the false negative rate. By Theorem 13, we get the overall power bound

$$\begin{aligned} \alpha \max_{q_1+\cdots+q_u \leq q} \prod_{i=1}^{u} \frac{1}{1 - \frac{19q_i^2}{N^2} - \frac{8n^3}{N^2}} \\ &\leq \max_{q_1+\cdots+q_u \leq q} \frac{1}{1 - \sum_{i=1}^{u} \left(\frac{19q_i^2}{N^2} - \frac{8n^3}{N^2} \right)} \\ &\leq \frac{\alpha}{1 - \frac{19q^2}{N^2} - \frac{8n^3}{N^2}} , \end{aligned}$$

using the fact that $1/(1-x) \times 1/(1-y) \leq 1/(1-x-y)$. Therefore, we obtain the following theorem.

Theorem 16. *Let $N = 2^n$ with $n > 17$. For every distinguisher from a uniform random function to the $\mathsf{SoP2}[\pi_1, \pi_2]$ construction instantiated with a uniform random permutation π, making in total $q \leq N/17$ to its u user oracles,*

$$1 - \beta \leq \frac{\alpha}{1 - \frac{19q^2}{N^2} - \frac{8n^3}{N^2}} .$$

where α is the false positive rate and β the false negative rate.

We compare our result to the previous result by Chen et al. [7], which corresponds to

$$1 - \beta \leq \alpha + \frac{10u^{1/2}q^2}{N^2} + \frac{17u^{1/2}(n+1)^2}{N} ,$$

and the bound obtained by naively using the hybrid argument on the single-user results obtained by Dutta et al. [10]:

$$1 - \beta \leq \alpha + \frac{19q^2}{N^2} + \frac{8n^3}{N^2} .$$

We see the improvement as these in the previous results

C Alternative proof of Theorem 13

The following two lemmas establish the first part of our claim, namely a power bound for distinguishers that make a predetermined number of queries to every user, in an alternative way. The proof uses more classical techniques – it is a hybrid proof – that we hope might be extended more easily to the case of dependent users.

The claim of the first lemma may seem obvious, since one might expect a reduction that simply simulates P to build a distinguisher from Q_1 to Q_2. However, this is a probabilistic reduction, and one needs to argue that it achieves the desired false positive rate and power for a complete proof.

Lemma 2. *Let P, Q_1 and Q_2 be distributions such that Q_1 and Q_2 have the same domain. If every distinguisher from Q_1 to Q_2 with false positive and negative rates α and β that makes q queries satisfies $1-\beta \leq a(q)\,\alpha$, then for all distinguishers between $P \otimes Q_1$ and $P \otimes Q_2$*

$$1-\beta' \leq a(q)\,\alpha\,,$$

with α' the false positive rate and β' the false negative rate.

Proof. Let $\mathcal{T}_1$ be the domain of P and $\mathcal{T}_2$ the shared domain of Q_1 and Q_2. Suppose the distinguisher from $P \otimes Q_1$ to $P \otimes Q_2$ has critical region $\mathcal{R} \subseteq \mathcal{T}_1 \times \mathcal{T}_2$. The fiber of x in $\mathcal{T}_1$ under the canonical projection $\mathcal{T}_1 \times \mathcal{T}_2 \to \mathcal{T}_1$ can be identified with $\mathcal{T}_2$ and defines a map $\mathscr{X}(x) = \{x_2 \in \mathcal{T}_2 \mid (x, x_2) \in \mathcal{R}\}$ from $\mathcal{T}_1$ to the set of subsets of $\mathcal{T}_2$. Let $\mathcal{R}_1$ be the set of first components of the elements of $\mathcal{R}$. By the definition of $\mathscr{X}$, we have

$$\mathcal{R} = \big\{(x, y) \mid x \in \mathcal{R}_1 \text{ and } y \in \mathscr{X}(x)\big\}\,.$$

Hence, if p is the probability mass function of P, then

$$1-\beta' = \big(P \otimes Q_2\big)(\mathcal{R}) = \sum_{x \in \mathcal{R}_1} p(x)\, Q_2\big(\mathscr{X}(x)\big)\,.$$

It is given that $Q_2(\mathscr{X}(x)) \leq a(q)\, Q_1(\mathscr{X}(x))$. Hence,

$$1-\beta' \leq a(q) \sum_{x \in \mathcal{R}_1} p(x)\, Q_1\big(\mathscr{X}(x)\big) = a(q)\,\alpha'\,.$$

The last equality depends on the fact that $\alpha' = \big(P \otimes Q_1\big)(\mathcal{R})$. □

Lemma 3. *For $i = 1, \ldots, u$, let P_i and Q_i be distributions with the same domain. If every distinguisher from P_i to Q_i making q queries satisfies $1-\beta_i \leq a_i(q)\,\alpha_i$ for false positive and negative rates α_i and β_i, then for every distinguisher from $P_1 \otimes P_2 \otimes \cdots \otimes P_u$ to $Q_1 \otimes Q_2 \otimes \cdots \otimes Q_u$ that makes q_i queries to user i,*

$$1-\beta \leq \alpha \prod_{i=1}^{u} a_i(q_i)\,,$$

with α the false positive rate and β the false negative rate.

Proof. The proof is based on a hybrid argument with u steps, where we replace P_i by Q_i in step $u+1-i$. That is, we have the following sequence of intermediate distributions:

$$\begin{aligned} &P_1 \otimes P_2 \otimes \cdots \otimes P_u \\ &P_1 \otimes P_2 \otimes \cdots \otimes Q_u \\ &\qquad\vdots \\ &P_1 \otimes Q_2 \otimes \cdots \otimes Q_u \\ &Q_1 \otimes Q_2 \otimes \cdots \otimes Q_u \,. \end{aligned}$$

For step $u + 1 - i$, we need a power bound for distinguishers from $P_1 \otimes \cdots \otimes P_i \otimes Q_{i+1} \otimes \cdots \otimes Q_u$ to $P_1 \otimes \cdots \otimes P_{i-1} \otimes Q_i \otimes \cdots \otimes Q_u$. Since the order is only a notational convention, this is the same as a distinguisher from $X \otimes P_i$ to $X \otimes Q_i$ with $X = P_1 \otimes \cdots \otimes P_{i-1} \otimes Q_{i+1} \otimes \cdots Q_u$. Hence, Lemma 2 gives $1 - \beta_i \le f_i(\alpha_i;\ q_i) = a_i(q_i)\,\alpha_i$ with false positive and negative rates α_i and β_i. These bounds can be combined using Theorem 4 to obtain

$$1 - \beta \le f_1\big(f_2(\cdots f_u(\alpha;\ q_u) \cdots ;\ q_2);\ q_1\big)\,.$$

This is the desired result. □