A short introduction to provable security in symmetric cryptography

Pierre Karpman

November 8, 2023

1 Statistical distinguishers

A *distinguisher* for two discrete* probability distributions $\mathfrak{U}$ and $\mathfrak{R}$ over a finite set $\mathcal{S}$ is a (possibly randomised) algorithm A whose input is a random variable X sampled either from $\mathfrak{U}$ or $\mathfrak{R}$, each with probability one half, and whose output is a unique bit. The *advantage* of A in distinguishing $\mathfrak{U}$ from $\mathfrak{R}$ is defined as:

$$\mathrm{Adv}_A^{\mathfrak{U},\mathfrak{R}} := |\Pr[A(X) = 1 : X \sim \mathfrak{U}] - \Pr[A(X) = 1 : X \sim \mathfrak{R}]| \tag{1}$$

where the probabilities are computed over the sampling of X and the coins of A (if any).

The *total variation distance* (sometimes called "the" statistical distance) $d_{tv}(\mathfrak{U}, \mathfrak{R})$ between $\mathfrak{U}$ and $\mathfrak{R}$ is defined as:

$$d_{tv}(\mathfrak{U}, \mathfrak{R}) = \frac{1}{2} \sum_{x \in \mathcal{S}} |\mathfrak{U}(x) - \mathfrak{R}(x)| \tag{2}$$

where for $\mathfrak{D}$ a probability distribution we use the notation $\mathfrak{D}(x) := \Pr[X = x : X \sim \mathfrak{D}]$. If we denote by $\mathcal{S}^+$ (resp. $\mathcal{S}^-$) the subset of $\mathcal{S}$ for which $\mathfrak{U}(x) > \mathfrak{R}(x)$ (resp. $\mathfrak{U}(x) \leq \mathfrak{R}(x)$), then we have an alternative definition for d_{tv} as:

$$d_{tv}(\mathfrak{U}, \mathfrak{R}) = \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) - \mathfrak{R}(x) \tag{3}$$

This simply follows from:

$$\begin{aligned}
\sum_{x \in \mathcal{S}} |\mathfrak{U}(x) - \mathfrak{R}(x)| &= \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) - \mathfrak{R}(x) \quad + \quad \sum_{x \in \mathcal{S}^-} \mathfrak{R}(x) - \mathfrak{U}(x) \\
&= \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) - \left(1 - \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x)\right) + \left(1 - \sum_{x \in \mathcal{S}^+} \mathfrak{R}(x)\right) - \sum_{x \in \mathcal{S}^+} \mathfrak{R}(x) \\
&= 2 \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) - \mathfrak{R}(x)
\end{aligned}$$

We then have:

Theorem 1. *Let $\mathfrak{U}$, $\mathfrak{R}$ be two probability distributions over a finite set $\mathcal{S}$ and A a distinguisher for $\mathfrak{U}$ and $\mathfrak{R}$, then:*

$$\mathrm{Adv}_A^{\mathfrak{U},\mathfrak{R}} \leq d_{tv}(\mathfrak{U}, \mathfrak{R})$$

*This qualifier will from now on be implicit, since all the distributions we will encounter are discrete.

Proof. For any distinguisher A, let $\alpha : \mathcal{S} \to [0,1]$ be its "decision" function defined from $\alpha(x) = \Pr[A(x) = 1]$. Then by definition $\mathrm{Adv}_A^{\mathfrak{U},\mathfrak{R}} = \left|\sum_{x\in\mathcal{S}} \alpha(x)\,\mathfrak{U}(x) - \sum_{x\in\mathcal{S}} \alpha(x)\,\mathfrak{R}(x)\right|$. Reusing the above notation for $\mathcal{S}^+$, $\mathcal{S}^-$, we can rearrange the terms of these sums as:

$$\left|\sum_{x\in\mathcal{S}^+} \alpha(x)(\mathfrak{U}(x) - \mathfrak{R}(x)) + \sum_{x\in\mathcal{S}^-} \alpha(x)(\mathfrak{U}(x) - \mathfrak{R}(x))\right|$$

Then since all the terms of the left (resp. right) sum are non-negative (resp. negative), the overall value is maximised when $\alpha(x)$ is 1 for $x \in \mathcal{S}^+$ and 0 for $x \in \mathcal{S}^-$ (or the converse, by symmetry). In that case the expression simplifies to $\sum_{x\in\mathcal{S}^+} \mathfrak{U}(x) - \mathfrak{R}(x) = d_{tv}(\mathfrak{U}, \mathfrak{R})$. □

Remark that this proof is constructive, since we have shown in passing that the distinguisher that answers 1 on input x if and only if $\mathfrak{U}(x) > \mathfrak{R}(x)$ reaches the upper-bound.[†] However this requires a complete knowledge of the two distributions, which will seldom be the case in our applications (additionally, the running time and memory cost of this distinguisher may also be exponential in the size of the elements of $\mathcal{S}$).

A major interest of Theorem 1 in the context of cryptography is that it allows to prove the security of a construction (usually under some assumptions) or to craft a security reduction by upper-bounding the advantage of any distinguisher thanks to a bound on the total variation distance between two well-chosen (families of) distributions. Such a bound that only relies on statistical properties of the construction and ignores the actual (possibly high) time and memory cost of a distinguisher is often said to be in the "information theory model".[‡]

The main steps to implement this strategy are: 1) defining the distributions in a way that makes them "compatible" with a relevant security definition (say, PRP); 2) computing a non-trivial upper-bound. While the first point is usually easy, the second often requires some (hard) work. Several frameworks have been developed to help with this, and we will give a short introduction to one of them: the "H-coefficient" technique of Patarin [Pat08] (see also Chen and Steinberger's exposition, on which our presentation is heavily based [CS14]).

Our starting point is to rewrite Equation (3) as $d_{tv}(\mathfrak{U}, \mathfrak{R}) = \sum_{x\in\mathcal{S}^+} \mathfrak{U}(x)(1-\mathfrak{R}(x)/\,\mathfrak{U}(x))$. For any partition of $\mathcal{S}^+$ into k pairwise-disjoint subsets $\mathcal{S}_1^+, \ldots, \mathcal{S}_k^+$, this can be further rewritten as:

$$\sum_{i=1}^{k} \sum_{x\in\mathcal{S}_i^+} \mathfrak{U}(x)(1 - \mathfrak{R}(x)/\,\mathfrak{U}(x)) \tag{4}$$

Now if one knows $\varepsilon_1, \ldots, \varepsilon_k$, all in $[0,1]$ s.t. for all $1 \le i \le k$, $x \in \mathcal{S}_i^+ \Rightarrow \mathfrak{R}(x)/\,\mathfrak{U}(x) \ge 1-\varepsilon_i$, then we immediately get the upper-bound:

$$(4) \le \sum_{i=1}^{k} \sum_{x\in\mathcal{S}_i^+} \mathfrak{U}(x)\varepsilon_i = \sum_{i=1}^{k} \Pr[X \in \mathcal{S}_i^+ : X \sim \mathfrak{U}]\varepsilon_i. \tag{5}$$

This method can still be applied even if one does not know a partition of $\mathcal{S}^+$ exactly, since "augmenting" the subsets $\mathcal{S}_i^+$ with some elements of $\mathcal{S}^-$ can only increase the bound (and in that case $\mathfrak{R}(x)/\,\mathfrak{U}(x) > 1$ so the lower bound by $1 - \varepsilon_i$ trivially holds); similarly, an overlap between some of the $\mathcal{S}_i^+$'s may only increase the bound.[§] We formalise this as the following:

[†] One may also answer 1 for any x s.t. $\mathfrak{U}(x) = \mathfrak{R}(x)$ without changing the advantage.

[‡] Note that an information-theoretic bound may sometimes still be used to reason about computational security definitions.

[§] It is also worth noting that everything here is symmetric, which is useful if it is rather a good partition of $\mathcal{S}^-$ that is known.

Lemma 2 (H-coefficients bound)**.** *Let $\mathfrak{U}$, $\mathfrak{R}$ be two probability distributions over a finite set $\mathcal{S}$, $\mathcal{S}^+ = \{x \in \mathcal{S} : \mathfrak{U}(x) \geq \mathfrak{R}(x)\}$, $\mathcal{S}_1, \ldots \mathcal{S}_k$ be k subsets of $\mathcal{S}$ s.t. $\cup_{i=1}^k \mathcal{S}_i \supseteq \mathcal{S}^+$, $\varepsilon_1, \ldots, \varepsilon_k \in [0,1]$ be s.t. for all $1 \leq i \leq k$, $x \in \mathcal{S}_i \Rightarrow \mathfrak{R}(x)/\mathfrak{U}(x) \geq 1 - \varepsilon_i$. Then $d_{tv}(\mathfrak{U}, \mathfrak{R}) \leq \sum_{i=1}^k \Pr[X \in \mathcal{S}_i : X \sim \mathfrak{U}]\varepsilon_i$.*

Lemma 2 is often used in a "simple" case where there are only two subsets in the partition: one set $\mathcal{S}_1$ that includes the "good" cases for which ε_1 is "small", meaning that $\mathfrak{U}$ and $\mathfrak{R}$ are roughly identical for the outcomes $x \in \mathcal{S}_1$, and one set $\mathcal{S}_2$ of "bad" cases where ε_2 may be much closer to 1. In this case one may simplify the bound from Lemma 2 without "losing too much" as:

$$d_{tv}(\mathfrak{U}, \mathfrak{R}) \leq \Pr[X \in \mathcal{S}_1 : \mathfrak{U}]\varepsilon_1 + \Pr[X \in \mathcal{S}_2 : \mathfrak{U}]\varepsilon_2 \leq \varepsilon_1 + \Pr[X \in \mathcal{S}_2 : \mathfrak{U}] \tag{6}$$

It may be useful to informally think of this last variant of the bound as a way of being "nice" to the distinguishers: by upper-bounding $\Pr[X \in \mathcal{S}_2 : \mathfrak{U}]\varepsilon_2$ by $\Pr[X \in \mathcal{S}_2 : \mathfrak{U}]$, we assume that there is a distinguisher that is always able to tell the difference between $\mathfrak{U}$ and $\mathfrak{R}$ whenever x is "bad". Said otherwise, we "give up pretending that $\mathfrak{R}$ is $\mathfrak{U}$" on bad events. This interpretation may sometimes help the process of constructing $\mathcal{S}_1$ and $\mathcal{S}_2$.

Remark 3. Equation (6) may still be useful even if one only knows an upper bound on $\Pr[X \in \mathcal{S}_2 : \mathfrak{U}]$ (this is possible without change since $\Pr[X \in \mathcal{S}_1 : \mathfrak{U}]$ has already been simplified away as '1') and if $\mathcal{S}_1$ or $\mathcal{S}_2$ are not subsets of $\mathcal{S}^+$ (since again in this case we only "add to" Equation (3)). This latter case in fact often occurs in practice, since one usually uses a partition of $\mathcal{S}$ rather than $\mathcal{S}^+$ in applications of Lemma 2.

Remark 4. In applications of Equation (6), one usually determines $\mathcal{S}_1$, $\mathcal{S}_2$ and the distribution for which to compute $p_{\text{bad}} := \Pr[X \in \mathcal{S}_2]$ independently. This may lead to somewhat degenerate cases where either ε_1 or p_{bad} is equal to zero (both will happen in our later examples). The approximations that lead to Equation (6) are still valid in this case, but one may then also directly derive the bound as following.

If $\varepsilon_1 = 0$, this means that $x \in \mathcal{S}_1 \Rightarrow \mathfrak{R}(x) \geq \mathfrak{U}(x)$, so $\mathcal{S}^+ \subseteq \mathcal{S}_2$ and $\Pr_{\mathfrak{U}}[X \in \mathcal{S}_2] = \sum_{x \in \mathcal{S}_2} \mathfrak{U}(x) \geq \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) \geq \sum_{x \in \mathcal{S}^+} \mathfrak{U}(x) - \mathfrak{R}(x) = d_{tv}(\mathfrak{U}, \mathfrak{R})$.

If $p_{\text{bad}} = 0$, this means that $\mathcal{S}^+ \subseteq \mathcal{S}_1$, and one is then in fact applying Lemma 2 on a trivial covering of $\mathcal{S}^+$, and the result thus remains valid.

2 Oracles and transcripts

This section is partially based on Chen and Steinberger's presentation of the H-coefficients method [CS14].

2.1 Definitions and first results

Our application to cryptography of the results from the previous section will be for distributions induced by algorithms interacting with one or several *oracles*. A key notion in the analysis will be the one of *transcripts*, and so we devolve this section to defining oracles, transcripts, and proving some useful results about them.

Definition 5 (Random oracle)**.** Let $\mathcal{F}$ be a space of functions from a finite set $\mathcal{S}$ to a finite set $\mathcal{S}'$. A *random oracle for $\mathcal{F}$ with distribution $\mathfrak{D}$* is the realisation of a random variable $X \sim \mathfrak{D}$, where $\mathfrak{D}$ is a distribution over $\mathcal{F}$.

Concrete instantiations of (the rather abstract) Definition 5 often have $\mathcal{F}$ either the permutations over $\mathcal{S}$ (written $\text{Perms}(\mathcal{S})$) or all the functions from $\mathcal{S}$ to $\mathcal{S}'$ (written $\text{Funcs}(\mathcal{S}, \mathcal{S}')$, or simply $\text{Funcs}(\mathcal{S})$ when the two sets are equal), with a uniform distribution.

Let $[\mathbb{O}]_i$ denote an ordered list of oracles indexed by i (index that we will freely drop from our notation when not needed), and $[\mathbb{O}]_i \twoheadleftarrow [\mathfrak{D}]_i$ the fact that they are the realisation of random variables sampled from the distributions in the similarly-defined $[\mathfrak{D}]_i$; we further write $A^{[\mathbb{O}]_i}$ for an algorithm that has oracle access to $[\mathbb{O}]_i$.

From now on, our typical setting will be an algorithm $A^{[\mathbb{O}]}$ that tries to distinguish the two cases $[\mathbb{O}]_i \twoheadleftarrow [\mathfrak{D}]_i$ and $[\mathbb{O}]_i \twoheadleftarrow [\mathfrak{D}']_i$, where each case happens with probability one half. This is similar to the scenario from the previous section, except now that A may make an arbitrary number of arbitrary queries to its oracles, and so there may be many distributions to consider. We consequently modify our definition of advantage to:

$$\mathrm{Adv}^{[\mathfrak{D}],[\mathfrak{D}']}_{A^{[\mathbb{O}]}_{q_1,\ldots,q_n}} := \left|\Pr\left[A^{[\mathbb{O}]}_{q_1,\ldots,q_n} = 1 : [\mathbb{O}] \twoheadleftarrow [\mathfrak{D}]\right] - \Pr\left[A^{[\mathbb{O}]}_{q_1,\ldots,q_n} = 1 : [\mathbb{O}] \twoheadleftarrow [\mathfrak{D}']\right]\right| \tag{7}$$

where we assume $n \geq 1$ oracles, $A^{[\mathbb{O}]}_{q_1,\ldots,q_n}$ denotes an algorithm that makes q_i queries to $\mathbb{O}_i$, and the probabilities are computed over the samplings of the oracle and the randomness of A (if any). It is worth noting that even though the oracles are randomly sampled from distributions that depend from an initial uniform bit, they are "fixed" once at the beginning of the experiment, and so from the perspective of A they behave deterministically; in the following we will refer to this by saying that the oracles are deterministic.

As previously, given some $[\mathfrak{D}]$, $[\mathfrak{D}']$, our main objective will be to find upper-bounds for the advantage of *any* distinguisher, usually in function of $q_1, \ldots, q_n$. To that end and as a first simplification, we will from now on restrict ourselves to *deterministic* distinguishers that do not make repeated queries. The latter is clearly without loss of generality, since we do not restrict the running time or the memory (and so there is no benefit in making twice the same query to the same oracle (which we recall is deterministic from the point of view of the distinguisher)). To argue that the former is also w.l.o.g., we observe that any terminating randomised algorithm with bounded input size (both conditions being fulfilled in our case since we force the algorithm to terminate and have just forbidden repeated queries) can be made deterministic by fixing a value for all its (finite number of) coins. Hence, for any such randomised algorithm A, there is a deterministic algorithm A' that performs as well as A (in our case, meaning that it has maximum advantage among the ones reachable by A).

Now in order to handle the many possible distributions induced by A's queries to its oracles, we define the central notion of *transcript*.

Definition 6 (Transcript)**.** Let $A^{[\mathbb{O}]}$ be an algorithm with oracle access to $[\mathbb{O}]$. The *transcript* τ produced by $A^{[\mathbb{O}]}$ is the ordered list $[(j, x_i, \mathbb{O}_j(x_i))]_i$ of all the (ordered) queries made by $A^{[\mathbb{O}]}$ to its oracles, along with their answers.[¶]

For deterministic distinguishers and oracles, for a fixed distinguisher, the transcript produced by $A^{[\mathbb{O}]}$ is fully determined by the values of the oracles $[\mathbb{O}]$; in other words, there is a well-defined function $T : A^{[\mathbb{O}]} \mapsto \tau$, where τ is the transcript produced by $A^{[\mathbb{O}]}$. Since the output of $A^{[\mathbb{O}]}$ is obviously a deterministic function of the transcript it produces, we also have that once a distinguisher is fixed, its output $A^{[\mathbb{O}]}$ is fully determined by $[\mathbb{O}]$.

Even though we have just defined transcripts as being ordered, we will soon show the important fact that for deterministic distinguishers and oracles, the best advantage achievable for a given transcript τ does *not* depend on the order in which the queries have been made, but solely on the set of queried values and their answers. To prepare for that, we give:

[¶]Here we emphasise the identity "j" of the oracle to which the ith query is addressed, but may again freely drop this information when it is clear from the context.

Definition 7 (Transcript compatibility)**.** Let τ be a transcript, the oracles for $[\mathcal{F}]$ *compatible with* τ, written $\mathcal{C}_{[\mathcal{F}]}(\tau)$, is the set of all ordered lists $[\mathbb{O}]$ of oracles defined for $[\mathcal{F}]$ s.t. $(j, x_i, y_i) \in \tau \Rightarrow \mathbb{O}_j(x_i) = y_i$.[‖]

Example 8. Let $\tau = [(1, 0, 0)]$, then:

- For $\mathcal{X} := [\mathrm{Perms}([\![0, N-1]\!])]$, $\mathcal{C}_{\mathcal{X}}(\tau)$ is the set of the $(N-1)!$ permutations over $[\![0, N-1]\!]$ that have 0 as a fixed point.
- For $\mathcal{X} := [\mathrm{Funcs}([\![0, N-1]\!])]$, $\mathcal{C}_{\mathcal{X}}(\tau)$ is the set of the N^{N-1} functions over $[\![0, N-1]\!]$ that have 0 as a fixed point.
- For $\mathcal{X} := [\mathrm{Perms}([\![0, N-1]\!]), \mathrm{Perms}([\![0, N-1]\!])]$, $\mathcal{C}_{\mathcal{X}}(\tau)$ is the set of the $(N-1)!N!$ pairs of permutations over $[\![0, N-1]\!]$ s.t. the first has 0 as a fixed point.
- For $\mathcal{X} := [\mathrm{Der}([\![0, N-1]\!])]$, where $\mathrm{Der}(\mathcal{S})$ is the set of derangements over $\mathcal{S}$, $\mathcal{C}_{\mathcal{X}}(\tau)$ is empty.
- For $\mathcal{X} := [\mathrm{Perms}([\![1, N]\!])]$, $\mathcal{C}_{\mathcal{X}}(\tau)$ is empty.

An immediate consequence of Definition 7 and the discussion preceding it is that for deterministic distinguishers and oracles, we have:

$$\left(T\left(A^{[\mathbb{O}]}\right) = \tau\right) \Rightarrow \left(\forall [\mathcal{F}], \forall [\mathbb{O}'] \in \mathcal{C}_{[\mathcal{F}]}(\tau),\, T\left(A^{[\mathbb{O}']}\right) = \tau\right) \tag{8}$$

Here we voluntarily do not specify the defining set of $[\mathcal{F}]$ to convey the fact that $[\mathcal{F}]$ may be "arbitrarily" different from the function spaces over which $[\mathbb{O}]$ is defined, and that from the perspective of A this does not matter.

We introduce yet another notation on transcripts: by $\mathcal{Q}(\tau)$, we denote the (unordered) *set* of queries that appear in τ (*i.e.*, we forget about the answers to the oracles' queries and the order in which they were made), and we similarly write $\mathcal{Q}(A^{[\mathbb{O}]})$ for the queries appearing in the transcript $T(A^{[\mathbb{O}]})$ produced by $A^{[\mathbb{O}]}$. If we wish to remember the order of the queries (only forgetting the answers), we write *e.g.* $\mathcal{Q}[A^{[\mathbb{O}]}]$. We are now ready to prove:

Lemma 9. *Let τ be a transcript, A a deterministic distinguisher, and $[\mathfrak{D}]$ a list of uniform distributions*[♯] *over a list of function spaces $[\mathcal{F}]$. If $\exists\, [\mathbb{O}'] \in [\mathcal{F}]$ s.t. $T(A^{[\mathbb{O}']}) = \tau$, then:*

$$\Pr\left[T\left(A^{[\mathbb{O}]}\right) = \tau : [\mathbb{O}] \twoheadleftarrow [\mathfrak{D}]\right] = \frac{\#\,\mathcal{C}_{[\mathcal{F}]}(\tau)}{\#[\mathcal{F}]}$$

Proof. We write $[\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau$ as a shorthand to denote the fact that the values of the oracles in $[\mathbb{O}]$ on the points appearing in $\mathcal{Q}(\tau)$ are the same as the corresponding answers in τ. By definition $([\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau) \Leftrightarrow ([\mathbb{O}] \in \mathcal{C}_{[\mathcal{F}]}(\tau))$, and so $\Pr[[\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau] = \#\,\mathcal{C}_{[\mathcal{F}]}(\tau)/\#[\mathcal{F}]$ (we drop the mention of the underlying probability space for conciseness). Then again by definition, $\Pr[T(A^{[\mathbb{O}]}) = \tau] = \Pr[\mathcal{Q}[A^{[\mathbb{O}]}] = \mathcal{Q}[\tau] \wedge [\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau]$, which in turn is equal to $\Pr[\mathcal{Q}[A^{[\mathbb{O}]}] = \mathcal{Q}[\tau] : [\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau] \times \Pr[[\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau]$. Now from the fact that there exists some $[\mathbb{O}']$ s.t. $T(A^{[\mathbb{O}']}) = \tau$ and that A is deterministic, it follows from Equation (8) that $\Pr[\mathcal{Q}[A^{[\mathbb{O}]}] = \mathcal{Q}[\tau] : [\mathbb{O} \,|\, \mathcal{Q}(\tau)] \equiv \tau] = 1$, which allows to conclude. □

In natural language, this means that the probability that a deterministic distinguisher produces a given transcript is either zero, or the probability that its oracles are compatible with this transcript. Crucially, this latter probability does not depend on the order of the

[‖] We assume here that the size of the lists are at least as big as the number of different oracles queried in τ.

[♯] The generalisation to arbitrary distributions is obvious.

queries in the transcript; this for instance means that if two distinguishers with the same oracle distributions produce (with non-zero probability) transcripts with the same set of queries and answers, possibly made in a different order, then they do so with the same probability. In particular, this means that one does not need (at this point) to "think about" the possible *adaptivity* of the distinguishers (by which we mean the fact that they may define their i^{th} query in function of the answer to the $i-1$ previous ones), since only the set of queries eventually matters, and not the internal process leading to it.

We conclude this section by drawing a link between Lemmas 2 and 9 and our main objective of upper-bounding distinguishing advantages. We will first do so for *non-adaptive* distinguishers, and then address the (usually harder) case of *adaptive* ones.

2.2 Upper-bounding the advantage of non-adaptive distinguishers

We (informally) say that a distinguisher is *non-adaptive* if it always makes the same queries to its oracles. From the above (and in particular Lemma 9), the probability of a transcript being produced does not depend on the order of the queries, so we may assume w.l.o.g. that a non-adaptive distinguisher always makes its queries in the same order.[♭] Conceptually, a (deterministic) non-adaptive distinguisher may then be thought of as fixing a set $\mathcal{R}$ of queries, sending all of those together to its oracles, receiving all of the answers together, and making its decision based on these.

Let now $[\mathcal{F}]$, $[\mathcal{F}']$ be two lists of n function spaces with corresponding uniform distributions $[\mathfrak{D}]$ and $[\mathfrak{D}']$. Fix a set $\mathcal{R}$ of queries to the n oracles, and let $\mathcal{T}^{\mathcal{R}}$ (or simply $\mathcal{T}$ if $\mathcal{R}$ is clear from the context) be the set of all possible transcripts that can be produced in either *world* $[\mathfrak{D}]$ or $[\mathfrak{D}']$ when querying the oracles on $\mathcal{R}$, and *where the order of the queries has been removed from the transcripts*. Each distribution $[\mathfrak{D}]$ and $[\mathfrak{D}']$ induces a *transcript distribution* over $\mathcal{T}$ as $\Pr_{[\mathfrak{D}]}[X = \tau] = \#\mathcal{C}_{[\mathcal{F}]}(\tau)/\#[\mathcal{F}] =: p_\tau^{[\mathfrak{D}]}$ and $\Pr_{[\mathfrak{D}']}[X = \tau] = \#\mathcal{C}_{[\mathcal{F}']}(\tau)/\#[\mathcal{F}'] =: p_\tau^{[\mathfrak{D}']}$ respectively. We can check that those are indeed distributions, since (say) $\sum_{\tau\in\mathcal{T}} \#\mathcal{C}_{[\mathcal{F}]}(\tau) = \#[\mathcal{F}]$ (all oracles produce some transcript,[♮] and since we have fixed the queries and forgotten about their order, no $[\mathbb{O}]$ may be compatible with two distinct transcripts).

Now if $\mathcal{T}^{\mathcal{R}}$ can be partitioned into two disjoint sets $\mathcal{T}^{\mathcal{R}}_{\text{good}}$ and $\mathcal{T}^{\mathcal{R}}_{\text{bad}}$ s.t. $\tau \in \mathcal{T}^{\mathcal{R}}_{\text{good}} \Rightarrow p_\tau^{[\mathfrak{D}]}/p_\tau^{[\mathfrak{D}']} \geq 1-\varepsilon$, then by Lemma 2 an upper-bound on the distance between the transcript distributions over $\mathcal{T}^{\mathcal{R}}$ is $d_{tv}(\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}}, \mathcal{T}^{\mathcal{R}}_{\mathfrak{D}'}) \leq \varepsilon + \Pr_{[\mathfrak{D}']}[X \in \mathcal{T}^{\mathcal{R}}_{\text{bad}}]$.

On the other hand, by definition, a non-adaptive distinguisher with queries $\mathcal{R}$ samples its transcripts from $\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}}$ or $\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}'}$. In other words, if it produces the transcript τ with non-zero probability, it follows from Lemma 9 that this probability is either $p_\tau^{[\mathfrak{D}]}$ or $p_\tau^{[\mathfrak{D}']}$ depending on the world it is interacting with. Consequently, the advantage of a distinguisher that makes the queries $\mathcal{R}$ to its oracles is upper-bounded by $d_{tv}(\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}}, \mathcal{T}^{\mathcal{R}}_{\mathfrak{D}'})$. It then follows that for non-adaptive distinguishers A:

$$\mathrm{Adv}^{[\mathfrak{D}],[\mathfrak{D}']}_{A^{[\mathbb{O}]}_{q_1,\ldots,q_n}} \leq \max_{\mathcal{R}\in\langle q_1,\ldots,q_n\rangle} d_{tv}(\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}}, \mathcal{T}^{\mathcal{R}}_{\mathfrak{D}'}) \tag{9}$$

(where $\langle q_1, \ldots, q_n\rangle$ denotes the set of queries containing q_i queries for its ith oracle).

The sudden apparition of a "max" term in Equation (9) may seem worrisome if one is to compute an actual bound, and indeed one cannot remove it in general.[•] However,

[♭] In fact, in the deterministic case, it would often be hard to define a distinguisher that always makes the same set of queries but in potentially different orders in function of the realisation of its oracles.

[♮] If necessary, one may augment each function's domain to "everything" and its co-domain with a special failure symbol '$\perp$' that may be returned on an input outside of the original domain.

[•] Think for instance of two worlds where $\mathbb{O}(0) = 0$ in one world and $\mathbb{O}(0) = 1$ in the other, but all other points follow the same distribution in both worlds. It is clear that the two worlds can be distinguished with advantage 1 if the queries include 0, and zero otherwise (regardless of their number).

in many cases the oracles will be highly "symmetric", and so $d_{tv}(\mathcal{T}^{\mathcal{R}}_{\mathfrak{D}}, \mathcal{T}^{\mathcal{R}}_{\mathfrak{D}'})$ will only be a function of the number of queries to each oracle.

2.3 Upper-bounding the advantage of adaptive distinguishers

In the more general case of *adaptive* distinguishers, nothing prevents one to make different queries depending on the previous answers from the oracles. In that case, there may not be any single set $\mathcal{R}$ any more s.t. the transcript distributions implied by a distinguisher are over $\mathcal{T}^{\mathcal{R}}$. Said otherwise, Equation (9) is not valid any more when considering such adaptive distinguishers. We illustrate this with the following:

Example 10. Consider a single oracle $\mathbb{O}$ over $\{0,1,2\} \to \{0,1\}$ which in the "ideal" world is sampled uniformly from all the 2^3 possible functions, and in the other "real" world is sampled uniformly from the reduced set that maps 0, 1, 2 (in that order) to either of 0, 0, 1; 0, 1, 1; 1, 1, 0; 1, 1, 1. The "absolute" distance between the two oracle distributions may be computed from the transcript distributions $\mathcal{T}^{\mathcal{R}}$ with $\mathcal{R} = \{0,1,2\}$, and it is equal to $1/2$. One may also compute the distance for any fixed $\mathcal{R}$ of size 2, which is $1/4$ in all the three cases.

Now consider the adaptive two-query distinguisher that first queries $y_0 := \mathbb{O}(0)$, and then $y_1 := \mathbb{O}(2)$ if $y_0 = 0$ and $y_1 := \mathbb{O}(1)$ if $y_0 = 1$.** The transcripts that may be produced in the real world are $[(0,0),(2,1)]$ and $[(0,1),(1,1)]$, each with probability $1/2$. On the other hand, the ideal world produces these transcripts with probability only $1/4$, and may additionally produce the transcripts $[(0,0),(2,0)]$ and $[(0,1),(1,0)]$, again with probability $1/4$. It follows that the distance between the two transcripts distributions is $1/2$, which is better than what can be achieved with fixed queries.

In order to handle adaptive distinguishers, one may (partially) rely on the fact that from Lemma 9, if a distinguisher produces a transcript τ with non-zero probability in at least one world, (using notation from Section 2.2) it does so with probability $p_\tau^{[\mathfrak{D}]}$ (resp. $p_\tau^{[\mathfrak{D}']}$) in the world $[\mathfrak{D}]$ (resp. $[\mathfrak{D}']$).†† Crucially, we insist again on the fact that these probabilities do not depend on the order in which the queries appear in τ, and also do not depend on the other transcripts that may be produced by this same distinguisher. Said otherwise, if we write $\mathcal{T}^{q_1,\ldots,q_n}$ the set of all possible unordered transcripts that may be produced in either world by any distinguisher that makes $q_1, \ldots q_n$ queries to its n oracles, it follows that the set $\mathcal{T}^A \subseteq \mathcal{T}^{q_1,\ldots,q_n}$ of transcripts produced by a distinguisher A needs to satisfy the two conditions: 1) $\sum_{\tau \in \mathcal{T}^A} p_\tau^{[\mathfrak{D}]} = \sum_{\tau \in \mathcal{T}^A} p_\tau^{[\mathfrak{D}']} = 1$; 2) if τ is produced by A with non-zero probability in at least one world, then it is produced with probability $p_\tau^{[\mathfrak{D}]}$ and $p_\tau^{[\mathfrak{D}']}$ respectively in the two worlds. We may then remark that by definition (and using again notation from Section 2.2), A is non-adaptive if and only if there is no $\mathcal{R} \in \langle q_1, \ldots, q_n \rangle$ s.t. $\mathcal{T}^A = \mathcal{T}^{\mathcal{R}}$.

Now an important observation is that if one partitions $\mathcal{T}^{q_1,\ldots,q_n}$ into two disjoint subsets $\mathcal{T}^{q_1,\ldots,q_n}_{\text{good}}$ and $\mathcal{T}^{q_1,\ldots,q_n}_{\text{bad}}$ with the property that there is an ε s.t. $\tau \in \mathcal{T}^{q_1,\ldots,q_n}_{\text{good}} \Rightarrow p_\tau^{[\mathfrak{D}]}/p_\tau^{[\mathfrak{D}']} \geq 1-\varepsilon$, then defining $\mathcal{T}^A_{\text{good}}$ as $\mathcal{T}^A \cap \mathcal{T}^{q_1,\ldots,q_n}_{\text{good}}$, one again has that $\tau \in \mathcal{T}^A_{\text{good}} \Rightarrow p_\tau^{[\mathfrak{D}]}/p_\tau^{[\mathfrak{D}']} \geq 1-\varepsilon$ with the same ε. In other words, the partition into "good" and "bad" sets and the computation of ε may be done "once" for $\mathcal{T}^{q_1,\ldots,q_n}$ (which does not depend on any particular distinguisher) and still be used for any distinguisher even if one "does not know" $\mathcal{T}^A$. Similarly defining $\mathcal{T}^A_{\text{bad}}$ as $\mathcal{T}^A \cap \mathcal{T}^{q_1,\ldots,q_n}_{\text{bad}}$, $\mathcal{T}^A_{\text{good}}$ and $\mathcal{T}^A_{\text{bad}}$ then form a partition of $\mathcal{T}^A$, and if one is able to compute an upper-bound on the probability p_{bad} that a transcript

**We may further specify the distinguisher by defining its answer to be 1 if $y_1 = 1$, and 0 otherwise. However this does not really matter for our purpose, which is only concerned with the transcripts produced by the distinguisher (hence its "potential" advantage), and not its actual advantage.

††Note that one of those two probabilities may be zero, but not both.

produced by A in the world $\mathfrak{D}'$ belongs to $\mathcal{T}^A_{\text{bad}}$, then by Lemma 2 the advantage of A will be upper-bounded by $\varepsilon + p_{\text{bad}}$. Since p_{bad} now *does* (in principle) depend on $\mathcal{T}^A$, it follows that our ability to upper-bound the advantage of any adaptive distinguisher will directly depend on our ability to upper-bound this probability independently of A. (This will be the case for our two examples of application developped in the next sections.)

3 PRP/PRF switching

We will now illustrate the techniques presented in the previous sections by proving a nice (and useful) result, *viz.* the *PRP/PRF switching lemma.*

We first recall the definitions of PRP and PRF advantage in distinguishing a family of mappings $E : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ from a uniform permutation and function respectively:

Definition 11 (PRP advantage)**.** The *PRP advantage* of $E : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ is defined as:

$$\begin{aligned}\text{Adv}^{\text{PRP}}_E(q,t) := \max_{A_{q,t}} \Big| &\Pr\left[A^{\mathbb{O}}_{q,t}(\cdot) = 1 : \mathbb{O} \twoheadleftarrow \text{Perms}(\mathcal{M})\right] - \\ &\Pr\left[A^{\mathbb{O}}_{q,t}(\cdot) = 1 : \mathbb{O} = E(K,\cdot), K \twoheadleftarrow \mathcal{K}\right]\Big|\end{aligned}$$

Definition 12 (PRF advantage)**.** The *PRF advantage* of $E : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ is defined as:

$$\begin{aligned}\text{Adv}^{\text{PRF}}_E(q,t) := \max_{A_{q,t}} \Big| &\Pr\left[A^{\mathbb{O}}_{q,t}(\cdot) = 1 : \mathbb{O} \twoheadleftarrow \text{Funcs}(\mathcal{M})\right] - \\ &\Pr\left[A^{\mathbb{O}}_{q,t}(\cdot) = 1 : \mathbb{O} = E(K,\cdot), K \twoheadleftarrow \mathcal{K}\right]\Big|\end{aligned}$$

Here $A^{\mathbb{O}}_{q,t}$ denotes a (deterministic, which is w.l.o.g. as per the previous section) algorithm that makes q pairwise-distinct queries to its oracle $\mathbb{O}$, runs in time t (something that we will mostly ignore, given our information-theoretic proof), and returns a unique bit, and $\text{Perms}(\mathcal{S})$ (resp. $\text{Funcs}(\mathcal{S})$) denotes the set of all permutations over the finite set $\mathcal{S}$ (resp. functions from $\mathcal{S}$ to itself). The "PRP/PRF switching problem" can then be stated as asking to upper-bound the PRF advantage of E in function of its PRP advantage.

Let $A^E_{q,t}$, $A^P_{q,t}$, $A^F_{q,t}$ be shorthands for $A^{\mathbb{O}}_{q,t}$ when $\mathbb{O}$ is respectively $E(k,\cdot)$ with a uniform k, a uniform permutation and a uniform function. Then from the triangular inequality:

$$\begin{aligned}&\Big|\Pr\left[A^F_{q,t}() = 1\right] - \Pr\left[A^E_{q,t}() = 1\right]\Big| = \\ &\Big|\Pr\left[A^F_{q,t}() = 1\right] - \Pr\left[A^P_{q,t}() = 1\right] + \Pr\left[A^P_{q,t}() = 1\right] - \Pr\left[A^E_{q,t}() = 1\right]\Big| \leq \\ &\Big|\Pr\left[A^F_{q,t}() = 1\right] - \Pr\left[A^P_{q,t}() = 1\right]\Big| + \Big|\Pr\left[A^P_{q,t}() = 1\right] - \Pr\left[A^E_{q,t}() = 1\right]\Big|\end{aligned}$$

Consequently, letting $\text{Adv}^{\text{FP}}(q,t) := \max_{A_{q,t}} \Big|\Pr\left[A^F_{q,t}() = 1\right] - \Pr\left[A^P_{q,t}() = 1\right]\Big|$, we get:

$$\text{Adv}^{\text{PRF}}_E(q,t) \leq \text{Adv}^{\text{PRP}}_E(q,t) + \text{Adv}^{\text{FP}}(q,t) \tag{10}$$

which has the form that we are looking for. Informally, what we have done here is "paying" $\text{Adv}^{\text{PRP}}_E$ to pass E as a uniform permutation, so that we can then use an E-independent term for the advantage in distinguishing a uniform permutation from a uniform function. Remark that this approach is in part made possible by the fact that Definitions 11 and 12 use exactly the same definition for the "real" oracle E, and by the fact that they are expressed in terms of a "distance" between the two oracles.

We now address the main task of computing (an upper-bound on) $\text{Adv}^{\text{FP}}(q,t)$. Following Sections 1 and 2, we will do so by computing an upper-bound (function of q) on the

total variation distance between the transcript distributions induced by an oracle access to P and F respectively. We write $\mathcal{T}^q$ for the set of all possible (unordered, w.l.o.g.) transcripts with q queries to either of P or F. Then following Section 2.3, we wish to define a partition of $\mathcal{T}^q$ into "good" and "bad" subsets and find the corresponding terms ε and p_{bad}, where we are "free" to choose the distribution w.r.t. which computing the upper-bound for p_{bad} (that following Section 2.3, we would like to be independent from any distinguisher).

The good set should contain all the transcripts that have a roughly equal probability to be obtained when making q queries to P or F, while the bad set should contain the (hopefully few) ones that are much likelier to occur in one of the two cases. Since it is clear that collisions are only possible for F, an obvious possibility for the bad set is to define it as $\mathcal{T}_{\text{col}}$, the set of transcripts for which $\exists i, j \neq i$ s.t. $y_i = y_j$; the set of good transcripts $\mathcal{T}_{\text{unq}}$ is then simply taken to be the complementary $\mathcal{T}^q \backslash \mathcal{T}_{\text{col}}$. Now we immediately have that for any $\tau \in \mathcal{T}_{\text{col}}$, $p_\tau^P = 0$ (where by an abuse of notation we use 'P' and 'F' to denote the uniform distribution over $\text{Perms}(\mathcal{M})$ and $\text{Funcs}(\mathcal{M})$ respectively), so in particular for any A, $\tau \in \mathcal{T}_{\text{col}}^A \Rightarrow p_\tau^P = 0$, and it follows that if computed w.r.t. P, p_{bad} is simply zero. It remains to compute ε s.t. for all $\tau \in \mathcal{T}_{\text{unq}}$, $1 - \varepsilon \leq p_\tau^F / p_\tau^P =: p$, or equivalently $1 - p \leq \varepsilon$. If we let $N = \#\mathcal{M}$, we have by definition of the oracles that $p = (1/N^q)/(1/\prod_{i=0}^{q-1}(N-i)) = \prod_{i=0}^{q-1}(N-i)/N^q$, and it becomes clear that p is also equal to the probability that there is no collision in q queries to F; equivalently $1 - p$ is the probability that there is at least one collision, which we can upper-bound by $q(q-1)/2N$.[‡‡] Taking this bound for ε then finally leads to $\text{Adv}^{\text{FP}}(q,t) \leq \varepsilon + p_{\text{bad}} = q(q-1)/2N$[§§] and:

Lemma 13 (PRP/PRF switching)**.** *Let $E : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$, $N = \#\mathcal{M}$, then:*

$$\text{Adv}_E^{PRF}(q,t) \leq \text{Adv}_E^{PRP}(q,t) + q(q-1)/2N$$

Remark 14. A more direct approach in upper-bounding $\text{Adv}^{\text{FP}}(q,t)$ is to remark from the above that $p_\tau^P > p_\tau^F$ iff. $\tau \in \mathcal{T}_{\text{unq}}$, *i.e.* "$\mathcal{S}^+$" is equal to $\mathcal{T}_{\text{unq}}$. From the proof of Theorem 1, the advantage of an optimal distinguisher is then $1 - p$. Note that this also provides us with an explicit optimal distinguisher and proves that our expression for $\text{Adv}^{\text{FP}}(q,t)$ is only an approximation in so far as we approximate $1 - p$.

We conclude this section by sketching a direct application of PRP/PRF switching to the analysis of the IND-CPA security of a block cipher E in counter mode. It is clear that the IND-CPA advantage of the counter mode instantiated with a uniform function is zero, and since the IND-CPA and PRF definitions are "compatible" in their usage of E, we simply have $\text{Adv}_{\text{CTR}[E]}^{\text{IND-CPA}}(q,t) = \text{Adv}_E^{\text{PRF}}(q',t) \leq \text{Adv}_E^{\text{PRP}}(q',t) + q'(q'-1)/2N$ (where q' is the total number of queries to E implied by the q queries made by the adversary in the IND-CPA game).

4 A provably-secure SPRP

In this section, we present a block cipher construction from Even and Mansour [EM91, EM97] and prove in an "ideal" ("non-standard") model that "it is an SPRP up to the birthday bound".

Let $P : \mathcal{M} \to \mathcal{M}$ be a public (*i.e.* not secret) permutation, then one simply defines the "Even-Mansour cipher" built from P as $E = (k_1 || k_2, m) \mapsto P(m \oplus k_1) \oplus k_2$ (here we assume that $\mathcal{M}$ can be endowed with a group operation '$\oplus$'[¶¶]). We will also actually focus

[‡‡] This can be obtained *e.g.* from the union bound applied to the $q(q-1)/2$ pairs (y_i, y_j) that all have probability $1/N$ of forming a collision.

[§§] As promised, this bound does not depend on t.

[¶¶] In practice one usually takes $\mathcal{M} = \{0,1\}^n$ for some n and $\oplus$ the bitwise XOR (this corresponds to Even and Mansour's original description).

on the "single-key" variant that takes $k_1 = k_2$, since (up to a few nuances) its security is in fact the same as the two-key original version; we hereafter denote this construction by "*SEM*".

The "ideal" model in which we will study the security of *SEM* considers that $P \twoheadleftarrow \mathrm{Perms}(\mathcal{M})$ and that one may only do oracle (or "black-box") accesses to it and its inverse. This will allow to show that $\mathrm{Adv}^{\mathrm{SPRP}}_{SEM}(q,t) \leq 2qt/\#\mathcal{M}$, where q denotes as usual the number of queries to the block cipher under investigation (here, *SEM*) or its inverse (since we aim for SPRP security) and t now denotes *the number of oracle accesses* to P or P^{-1} (written $P^{\pm}$ for short).[‖ ‖] Informally such a proof (only) captures "generic" attacks, in that it guarantees that for a *SEM* instantiation with a "concrete" explicit permutation for P, any attack that does better than the bound *must* exploit some "structural" properties of this permutation. While this is a common feature of security proofs in cryptography, the fact that it is done here in an ideal model has for consequence that there is no explicit term to account for the non-idealness of P (which, since not defined, will then obviously be hard to evaluate☂). This is to be contrasted with proofs that reduce the security of a construction to a "standard" security definition (not implemented through oracle accesses); for instance we showed in the previous section that $\mathrm{Adv}^{\text{IND-CPA}}_{\mathrm{CTR}[E]} \leq$ *some structural term* + *some generic term*, and any candidate E for the instantiation may be evaluated independently w.r.t. the structural term (here, PRP security).

Upper-bounding the distinguishing advantage

We now wish to upper-bound $\mathrm{Adv}^{\mathrm{SPRP}}_{SEM}$, *i.e.* we want to upper-bound the advantage of any distinguisher which is given oracle access to $P^{\pm}$ and $E^{\pm}$, where in both worlds $P \twoheadleftarrow \mathrm{Perms}(\mathcal{M})$, in the *ideal world* $E \twoheadleftarrow \mathrm{Perms}(\mathcal{M})$ and in the *real world* $E = x \mapsto P(x \oplus k) \oplus k$, $k \twoheadleftarrow \mathcal{M}$. In order to simplify notation, we will also assume here that $\oplus$ is s.t. $a \oplus a = 0$.

Similarly as in Section 3, we start by defining transcripts that summarise the information obtained by a distinguisher from its interaction with its oracles. We will then use Lemma 2 to derive an upper-bound on the total variation distance between the transcript distributions from the real and the ideal world in function of the number of queries made to each oracle, which will allow us to conclude. A slight difference from Section 3 is that the two worlds differ not only by the definition of their oracles but also by the presence or not of one variable, *viz.* k; to make the computation of an upper-bound easier, we will include this variable k in the real-world transcript and a similar $k \twoheadleftarrow \mathcal{M}$ (that does not really correspond to anything) in the ideal world one. We may conceptualise this as giving away k to the distinguisher once it has made all the oracle queries it wanted, and clearly this cannot lower the advantage of an optimal distinguisher. We also assume that the distinguisher does not make repeated queries or a query to P^{-1} (resp. E^{-1}) on an answer from a previous query to P (resp. E), and vice-versa; again this is w.l.o.g. w.r.t. optimal distinguishers. To summarise, a transcript is then an ordered list $[(x_i, P(x_i))_i, (x'_i, P^{-1}(x'_i))_i, (x''_i, E(x''_i))_i, (x'''_i, E^{-1}(x'''_i))_i, k]$.

Now that we have defined our transcripts, the objective is to upper-bound the distance between real-world and ideal-world transcript distributions for adaptive distinguishers, which we will again do by following the approach described in Section 2.3.

Given $\mathcal{T}^{q,t}$ the set of (unordered, w.l.o.g.) transcripts that may be produced in either world when making q queries to $E^{\pm}$ and t queries to $P^{\pm}$, the first step is to define a partition of this set into "good" and "bad" transcripts. To do so, we may follow the intuition that it is unlikely that the ideal-world oracle E satisfies $E(x) = P(x \oplus k) \oplus k$ on

[‖ ‖] Notice that in line with a previous remark, this means that this "information theoretic" bound "does not care any more" about the physical time cost or the memory of a distinguisher.

a random x,[♯♯] and so any transcript that allows us to "check" if this equality holds will provide us with a strong distinguishing advantage. We thus define the bad transcripts $\mathcal{T}_{\text{bad}}$ as the ones that define both $P(x)$ (either because the query x appears for P, or because x is the answer to a query to P^{-1}) and $E(x \oplus k)$[♭♭] (again, either because $x \oplus k$ was queried for E, or because it was returned as an answer to a query to E^{-1}), along with the ones that define both $P^{-1}(x)$ and $E^{-1}(x \oplus k)$. Indeed, those are all the queries that allow "easy" consistency checks between P and E that are guaranteed to pass in the real world, and unlikely to do so in the ideal world.

Alternatively, one may interpret those bad transcripts as being the ones s.t. in the real world, either P is "overconstrained", or the transcript has zero probability. The overconstrained case comes from the observation that in this world, a query to $P^{\pm}$ (resp. $E^{\pm}$) defines a point for P except if it has already been defined by another query to $E^{\pm}$ (resp. $P^{\pm}$); thus the number of points of P defined in a transcript is $\leq t + q$, and bad transcripts with non-zero probability are exactly the ones for which this inequality is strict (this contrasts with the ideal world where $t + q$ queries always define $t + q$ points of the oracles P and E in total). We limit our discussion of the zero-probability case to an example, *viz.* a transcript s.t. $P(0) = 0$, $E(0) = 1$, $k = 0$.

Let now $\mathcal{T}_{SEM}^{A_{q,t}}$ and $\mathcal{T}_{\mathfrak{I}}^{A_{q,t}}$ respectively denote the transcript distributions from the real ("*SEM*") and ideal world, w.r.t. the transcripts produced by an arbitrary (possibly adaptive) distinguisher A that makes q queries to $E^{\pm}$ and t queries to $P^{\pm}$. Our goal is to compute un upper-bound on $p_{\text{bad}} := \Pr[X \in \mathcal{T}_{\text{bad}} : X \sim \mathcal{T}_{\mathfrak{I}}^{A_{q,t}}]$ valid for any $A_{q,t}$,[♮♮] and ε s.t. for all $\tau \in \mathcal{T}_{\text{good}}$, $p_{\tau}^{SEM}/p_{\tau}^{\mathfrak{I}} \geq 1 - \varepsilon$ (where we switch notation and drop the mention of $A_{q,t}$ to emphasise the fact that this bound will be computed directly from the "full" $\mathcal{T}^{q,t}$ and the structure of the oracles). From there Lemma 2 will give us $d_{tv}(\mathcal{T}_{SEM}^{A_{q,t}}, \mathcal{T}_{\mathfrak{I}}^{A_{q,t}}) \leq \varepsilon + p_{\text{bad}}$, which is also an upper-bound for $\text{Adv}_{SEM}^{\text{SPRP}}(q, t)$.

Upper-bounding p_{bad}

We start by observing that as far as oracle compatibility is concerned, it does not matter that a point $P(x) = y$ has been defined through a query x to P with answer y, or a query y to P^{-1} with answer x (and the same obviously goes for E). It is thus enough to think about a transcript in terms of the points (α, β) (resp. (α', β')) that have been defined for P (resp. E). Now a point (α, β) for (say) P leads to a transcript being bad iff. this latter includes a point (α', β') for E s.t. $\alpha' = \alpha \oplus k$ or $\beta' = \beta \oplus k$. Since k is drawn uniformly from $\mathcal{M}$ and independently from P and E, the probability (over any transcript distribution) that this happens for a fixed couple of points (α, β) and (α', β') is $\leq 2/\#\mathcal{M}$ (by the union bound applied to the two possible events); by the union bound again, the probability that (α, β) leads to badness is $\leq 2q/\#\mathcal{M}$ (since q is the number of queries to $E^{\pm}$ in the transcript); finally by the union bound again, the probability that a transcript is bad is $\leq 2qt/\#\mathcal{M}$ (since t is the number of queries to $P^{\pm}$). Note that as required, this upper-bound is the same for any distinguisher $A_{q,t}$: indeed, "badness" may only result from the outcome of the sampling of k, which is not controlled by the distinguisher.

[♯♯] The expected number of points x on which the equality holds is equal to the expected number of fixed points for a uniform permutation, which is 1.

[♭♭] Remember that k is part of the transcript, so this is a well-defined condition.

[♮♮] Following Remark 4, the choice of the distribution for which to compute p_{bad} is here determined by the one for which it is easiest; indeed, the independence of P, E and k in the ideal world will make this much more straightforward than in the real one.

Computing ε

To compute ε, we will explicitly compute p_τ^{SEM} and $p_\tau^{\mathfrak{I}}$ for good transcripts τ, and show that $p_\tau^{SEM} \geq p_\tau^{\mathfrak{I}}$; this will immediately lead to $\varepsilon = 0$.[••]

To compute $p_\tau^{\mathfrak{I}}$, it is enough to count the number of oracles compatible with τ in the ideal world, that is the number of oracles that agree with τ on: the t $P^\pm$ queries; the q $E^\pm$ queries; k (where we may call k an oracle by a slight abuse of definition). Letting $N = \#\mathcal{M}$, these probabilities are respectively equal to $(N-t)!$, $(N-q)!$ and 1, and so $p_\tau^{\mathfrak{I}} = ((N-t)! \times (N-q)!)/(N! \times N! \times N)$.

By definition of the real world, there is now a single permutation involved, and the number of oracles compatible with a transcript is given by how many agree with all the points defined through queries to $P^\pm$ and $E^\pm$. If the transcript is good, no point is defined more than once, and so there are $t+q$ of them and $(N-t-q)!$ compatible oracles. From the independence of the sampling of k, it follows that $p_\tau^{SEM} = (N-t-q)!/(N \times N!)$.

Now it only remains to show that $p_\tau^{SEM}/p_\tau^{\mathfrak{I}} \geq 1$. This ratio is equal to $((N-t-q)! \times N!)/((N-t)!(N-q)!)$, and letting $u = q$ and d (which we assume w.l.o.g. to be non-negative) s.t. $t = u - d$, it rewrites as $((N-2u+d)! \times N!)/(N-u+d)!(N-u)!$. Letting $(a)_b$ for $b < a$ denoting the falling factorial $a!/(a-b)!$, the ratio further rewrites as $(N)_u/(N-u+d)_u$, which is indeed at least 1. We conclude with the remark that the equality case is for $d = u$, meaning that the transcript only includes queries to E (in the present case, w.l.o.g.). This is consistent with the fact that in this case the two worlds behave *exactly* the same (as is clear by the oracles' definitions).

To conclude, all of the above gives us:

Theorem 15 (SPRP security of the SEM construction in the ideal permutation model)**.** *Using the above notation, one has:*

$$\mathrm{Adv}_{SEM}^{SPRP}(q,t) \leq 2qt/\#\mathcal{M}$$

References

[CS14] Shan Chen and John P. Steinberger. Tight security bounds for key-alternating ciphers. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology — EUROCRYPT 2014*, volume 8441 of *Lecture Notes in Computer Science*, pages 327–350. Springer, 2014.

[EM91] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. In Hideki Imai, Ronald L. Rivest, and Tsutomu Matsumoto, editors, *Advances in Cryptology — ASIACRYPT '91*, volume 739 of *Lecture Notes in Computer Science*, pages 210–224. Springer, 1991.

[EM97] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. *J. Cryptology*, 10(3):151–162, 1997.

[Pat08] Jacques Patarin. The "Coefficients H" technique. In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *Selected Areas in Cryptography — SAC 2008*, volume 5381 of *Lecture Notes in Computer Science*, pages 328–345. Springer, 2008.

[••]The exact computation of these probabilities is not always tractable in applications of the H-coefficient method. In that respect, our case is simple.