The Indistinguishability of the XOR of k Permutations

B. Cogliati, R. Lampe, J. Patarin

University of Versailles

March 4, 2014

We will use the following notations:

- I_n is the set of $n-$bit strings,
- F_n is the set of functions from I_n to I_n,
- B_n is the set of permutations of I_n,
- $\tilde{b}$ is the mean of b.

$$f = f_1 \oplus \ldots \oplus f_k$$

$$f_1, \ldots, f_k \in_R B_n$$

$$F \in_R F_n$$

The advantage $\mathrm{adv}_{A,f}$ of an adversary A trying to distinguish the XOR f of k permutations from a truly random function F in less than q queries is:

$$\mathrm{adv}_{A,f,q} = |\mathbb{P}(A(f) = 1) - \mathbb{P}(A(F) = 1)|.$$

Our goal is to upper bound the maximal advantage adv_q any adversary can get.

Theorem

Let $k, n \geq 1$, $f_1, \ldots, f_k \in_R B_n$ and $q \leq 2^{n-1}/k$ be the number of queries the adversary can ask. Then the advantage to distinguish $f = f_1 \oplus \ldots \oplus f_k$ from a uniformly random function using q queries satisfies:

$$\mathrm{adv}_q \leq 2^{-k(n-1)} * \sum_{0 \leq i \leq q} i^k = O\left(\frac{q^{k+1}}{2^{kn}}\right).$$

The best known attacks for the XOR of k permutations give the following bounds:

- $\mathrm{adv}_q \geq \mathcal{O}\left(\frac{q(q-1)}{2^{kn}}\right)$ if $q \ll 2^{\frac{n}{2}}$,
- $\mathrm{adv}_q \geq \mathcal{O}\left(\frac{q}{2^{(k-\frac{1}{2})n}}\right)$ if $2^{\frac{n}{2}} \ll q \ll 2^n$.

Theorem

Let $n \geq 1$, $f_1, f_2 \in_R B_n$ and $q \ll 2^n$ be the umber of queries asked by the adversary. Then the advantage when trying to distinguish $f = f_1 \oplus f_2$ from a uniformly random function in less than q queries satisfies:

$$\mathrm{adv}_q \leq \mathcal{O}\left(\frac{q}{2^n}\right).$$

Let a, b be two sequences of q n-bit strings. $H_q(a, b)$ corresponds to the number of $(f_1, \ldots, f_k) \in B_n^k$ such that

$$\forall i,\ 1 \leq i \leq q,\ (f_1 \oplus \ldots \oplus f_k)(a_i) = b_i.$$

Theorem

Let α, β be two positive real numbers. Let $E \subset I_n^q$ such that $|E| \geq (1-\beta)2^{nq}$. Suppose that for every sequences $(a_i)_{1 \leq i \leq q}$, $(b_i)_{1 \leq i \leq q}$ of pairwise distincts n-bit queries such that $(b_i)_{1 \leq i \leq m} \in E$, one has:

$$H_q(a, b) \geq (1-\alpha)\tilde{H}_q.$$

Then

$$\mathrm{adv}_q \leq \alpha + \beta.$$

Let a, b be two sequences of q n-bit strings. $H_q(a, b)$ corresponds to the number of $(f_1, \dots, f_k) \in B_n^k$ such that

$$\forall i, 1 \le i \le q, (f_1 \oplus \dots \oplus f_k)(a_i) = b_i.$$

Theorem

Let α, β be two positive real numbers. Let $E \subset I_n^q$ such that $|E| \ge (1 - \beta)2^{nq}$. Suppose that for every sequences $(a_i)_{1 \le i \le q}$, $(b_i)_{1 \le i \le q}$ of pairwise distincts n-bit queries such that $(b_i)_{1 \le i \le m} \in E$, one has:

$$H_q(a, b) \ge (1 - \alpha)\tilde{H}_q.$$

Then

$$\mathrm{adv}_q \le \alpha + \beta.$$

$H_q(a,b)$ is the number of $(f_1,\dots,f_k) \in B_n^k$ such that:

$$\begin{cases} f_1(a_1) \oplus f_2(a_1) \oplus \dots \oplus f_{k-1}(a_1) \oplus f_k(a_1) = b_1 \\ \vdots \\ f_1(a_q) \oplus f_2(a_q) \oplus \dots \oplus f_{k-1}(a_q) \oplus f_k(a_q) = b_q \end{cases}$$

Since our permutations are fixed on only q queries, what actually matters is the number $h_q(b)$ of solutions of the following system:

$$\begin{cases} P_1^1 \oplus P_1^2 \oplus \dots \oplus P_1^{k-1} \oplus P_1^k = b_1 \\ \vdots \\ P_q^1 \oplus P_q^2 \oplus \dots \oplus P_q^{k-1} \oplus P_q^k = b_q \\ P_i^1 \neq P_j^1 \text{ if } i \neq j \\ \vdots \\ P_i^k \neq P_j^k \text{ if } i \neq j \end{cases}$$

$H_q(a,b)$ is the number of $(f_1,\dots,f_k) \in B_n^k$ such that:

$$\begin{cases} f_1(a_1) \oplus f_2(a_1) \oplus \dots \oplus f_{k-1}(a_1) \oplus f_k(a_1) = b_1 \\ \vdots \\ f_1(a_q) \oplus f_2(a_q) \oplus \dots \oplus f_{k-1}(a_q) \oplus f_k(a_q) = b_q \end{cases}$$

Since our permutations are fixed on only q queries, what actually matters is the number $h_q(b)$ of solutions of the following system:

$$\begin{cases} P_1^1 \oplus P_1^2 \oplus \dots \oplus P_1^{k-1} \oplus P_1^k = b_1 \\ \vdots \\ P_q^1 \oplus P_q^2 \oplus \dots \oplus P_q^{k-1} \oplus P_q^k = b_q \\ P_i^1 \neq P_j^1 \text{ if } i \neq j \\ \vdots \\ P_i^k \neq P_j^k \text{ if } i \neq j \end{cases}$$

Lemma

Then for $a, b \in I_n^q$*:*

$$H_q(a,b) = h_q(b)\left(\frac{|B_n|}{2^n \times \cdots \times (2^n - q + 1)}\right)^k .$$

We want to compute $\frac{H_q}{\tilde{H}_q} = \frac{h_q}{\tilde{h}_q}$.

It is done recursively : we find t such that

$$\frac{h_{\alpha+1}}{\tilde{h}_{\alpha+1}} \geq \frac{h_\alpha}{\tilde{h}_\alpha}(1-t).$$

Hence

$$\frac{h_q}{\tilde{h}_q} \geq (1-t)^q \geq 1 - qt.$$

Then, using the relationship between h_q and the advantage,

$$\mathrm{adv}_q \leq qt.$$

We want to compute $\frac{H_q}{\tilde{H}_q} = \frac{h_q}{\tilde{h}_q}$.
It is done recursively : we find t such that

$$\frac{h_{\alpha+1}}{\tilde{h}_{\alpha+1}} \geq \frac{h_\alpha}{\tilde{h}_\alpha}(1-t).$$

Hence

$$\frac{h_q}{\tilde{h}_q} \geq (1-t)^q \geq 1 - qt.$$

Then, using the relationship between h_q and the advantage,

$$\mathrm{adv}_q \leq qt.$$

We want to compute $\frac{H_q}{\tilde{H}_q} = \frac{h_q}{\tilde{h}_q}$.
It is done recursively : we find t such that

$$\frac{h_{\alpha+1}}{\tilde{h}_{\alpha+1}} \geq \frac{h_\alpha}{\tilde{h}_\alpha}(1-t).$$

Hence

$$\frac{h_q}{\tilde{h}_q} \geq (1-t)^q \geq 1-qt.$$

Then, using the relationship between h_q and the advantage,

$$\mathrm{adv}_q \leq qt.$$

We want to compute $\frac{H_q}{\tilde{H}_q} = \frac{h_q}{\tilde{h}_q}$.
It is done recursively : we find t such that

$$\frac{h_{\alpha+1}}{\tilde{h}_{\alpha+1}} \geq \frac{h_\alpha}{\tilde{h}_\alpha}(1-t).$$

Hence

$$\frac{h_q}{\tilde{h}_q} \geq (1-t)^q \geq 1-qt.$$

Then, using the relationship between h_q and the advantage,

$$\mathrm{adv}_q \leq qt.$$

Our goal is to compute $h_{\alpha+1}$ from h_α, i.e. the number of $(P_i^j)_{1\leq i\leq m, 1\leq j\leq k}$ such that:

$$\begin{array}{ccccccccccc}
P^1_{\alpha+1} & \oplus & P^2_{\alpha+1} & \oplus & \dots & \oplus & P^{k-1}_{\alpha+1} & \oplus & P^k_{\alpha+1} & = & b_{\alpha+1} \\
P^1_{\alpha} & \oplus & P^2_{\alpha} & \oplus & \dots & \oplus & P^{k-1}_{\alpha} & \oplus & P^k_{\alpha} & = & b_{\alpha} \\
\vdots & & \vdots & & & & & \vdots & \vdots & & \vdots \\
P^1_{1} & \oplus & P^2_{1} & \oplus & \dots & \oplus & P^{k-1}_{1} & \oplus & P^k_{1} & = & b_{1}
\end{array}$$

Pairwise distinct messages

$$
\begin{array}{ccccccccccc}
P^1_{\alpha+1} & \oplus & P^2_{\alpha+1} & \oplus & \dots & \oplus & P^{k-1}_{\alpha+1} & \oplus & P^k_{\alpha+1} & = & b_{\alpha+1} \\
\boxed{\begin{array}{c} P^1_\alpha \\ \vdots \\ P^1_1 \end{array}} & \begin{array}{c} \oplus \\ \\ \oplus \end{array} & \boxed{\begin{array}{c} P^2_\alpha \\ \vdots \\ P^2_1 \end{array}} & \begin{array}{c} \oplus \\ \\ \oplus \end{array} & \begin{array}{c} \dots \\ \\ \dots \end{array} & \begin{array}{c} \oplus \\ \\ \oplus \end{array} & \boxed{\begin{array}{c} P^{k-1}_\alpha \\ \\ P^{k-1}_1 \end{array}} & \begin{array}{c} \oplus \\ \vdots \\ \oplus \end{array} & \boxed{\begin{array}{c} P^k_\alpha \\ \vdots \\ P^k_1 \end{array}} & \begin{array}{c} = \\ \\ = \end{array} & \begin{array}{c} b_\alpha \\ \vdots \\ b_1 \end{array}
\end{array}
$$

Pairwise distinct messages

Theorem

If $q < \frac{2^n}{12}$ and $k \geq 3$,

$$\begin{aligned} \mathrm{adv} &\leq \frac{kq^2.2^n}{(2^n - q)^k} + 12\frac{q^{k+2}}{(2^n - 3q)(2^n - q)^k} && (1) \\ &\leq \frac{kq^2}{2^{(k-1)n}(1 - k\frac{q}{2^n})} + 12\frac{q^{k+2}}{2^{(k+1)n}(1 - (k+3)\frac{q}{2^n})}. && (2) \end{aligned}$$

Theorem

Let α, β be two positive real numbers. Let $E \subset I_n^q$ such that $|E| \geq (1-\beta)2^{nq}$. Suppose that for every sequence $(a_i)_{1 \leq i \leq q}$, $(b_i)_{1 \leq i \leq q}$ of pairwise distinct messages, $(b_i)_{1 \leq i \leq m} \in E$, we have:

$$H(a, b) \geq (1-\alpha)\tilde{H}_q.$$

Then

$$\mathrm{adv}_q \leq \alpha + \beta.$$

Using this theorem and the Bienaymé-Tchebitchev's inequality, we get:

$$\begin{aligned} \mathrm{adv}_q &\leq 2\left(\frac{\mathsf{V}\left[H_q(a)\right]}{\tilde{H}_q(a)^2}\right)^{1/3} = 2\left(\frac{\mathsf{V}\left[h_q\right]}{\tilde{h}_q^2}\right)^{1/3} \\ &\leq 2\left(\frac{\lambda_q}{U_q} - 1\right)^{1/3} , \end{aligned}$$

where $U_q := 2^{nq}\tilde{h_q}^2$ and λ_q is the number of sequences $P^1, P^2, \ldots, P^{2k}$ of q pairwise distinct messages such that $P^1 \oplus \ldots \oplus P^{2k} = 0$

The advantage any adversary can get with q queries, where $q \leq \frac{2^n}{2k}$, satisfies:

$$\mathrm{adv}_q \leq 2\left(\left(1+\frac{q2^n}{(2^n-q)^{2k}}+\frac{2kq^{2k+1}}{\left(1-\frac{2kq}{2^n}\right)2^n(2^n-q)^{2k}}\right)^q-1\right)^{1/3}.$$

i.e.

$$\mathrm{adv}_q \lesssim 2\left(\frac{q^2}{2^{(2k-1)n}(1-\frac{q}{2^n})^{2k}}+\frac{2kq^{2k+2}}{2^{(2k+1)n}(1-\frac{6kq}{2^n})}\right)^{1/3}.$$

technique	S. Lucks	H	H_σ
security bound	$O\left(\frac{q^{k+1}}{2^{kn}}\right)$	$O\left(\frac{q^{k+2}}{2^{(k+1)n}}\right)$	$O\left(\left(\frac{q^{2k+2}}{2^{(2k+1)n}}\right)^{1/3}\right)$

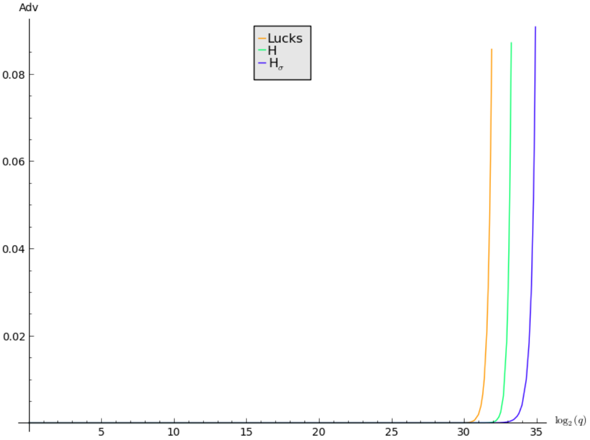


Figure : Upper bound for $n = 40$, $k = 5$

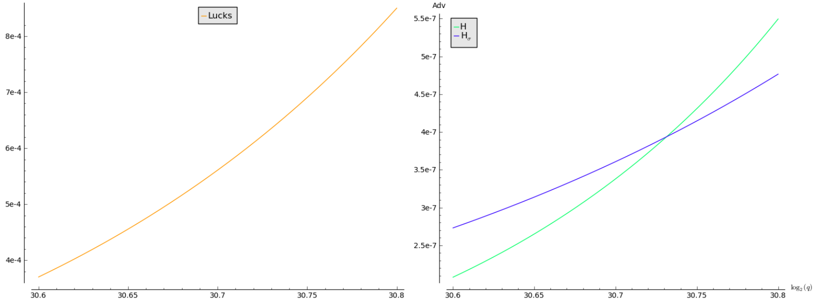
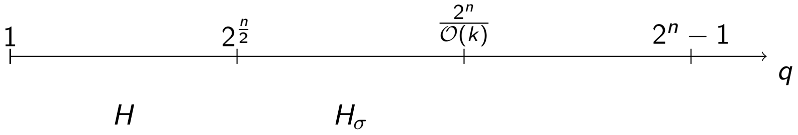


Figure : Upper bound for $n = 40$, $k = 5$

Our results can be further improved by using the techniques recursively, as in the original articles from J. Patarin.

These proof techniques (especially the H_σ coefficients) can be used on (both balanced and unbalanced) Feistel schemes.



Open problem: what happens in the third area?

Thank you for your attention.