Finding collisions using differentials

Yann Rotella
séminaire CASYS, Grenoble

27 juin 2019

European Research Council
Established by the European Commission

Radboud University

IN·DEI·NOMINE·FELICITER

Structure of this Talk

Question

Let (M, M') be a pair of public messages, then

$$\Pr[F(M) = F(M')]?$$

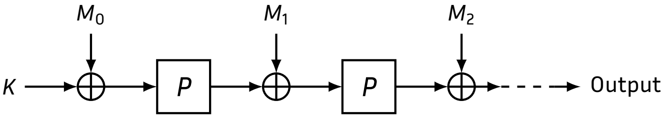
Question

Let (M, M') be a pair of public messages, then

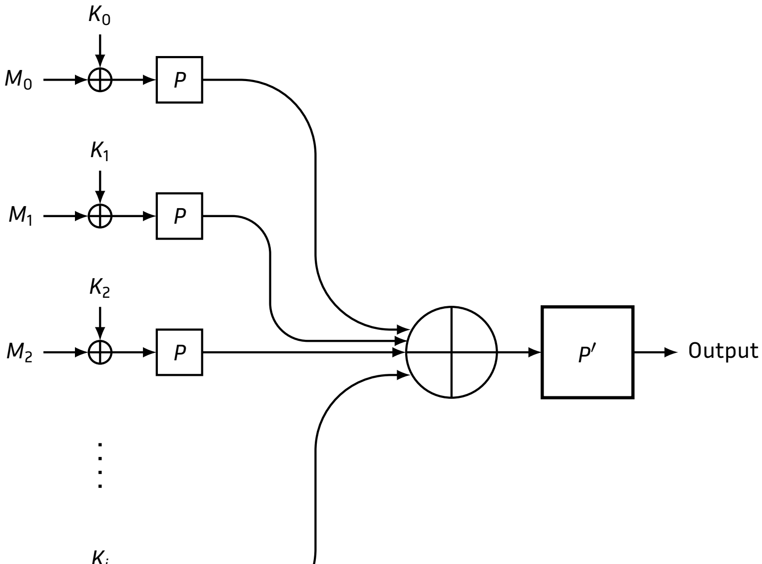
$$\Pr[F(M) = F(M')]?$$

$$\Pr[F(M) = F(M')|M + M' = \Delta]?$$

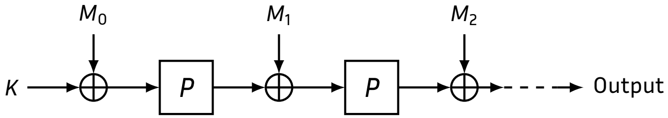
Serial Construction

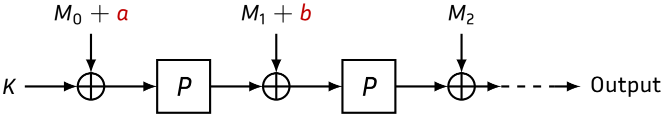


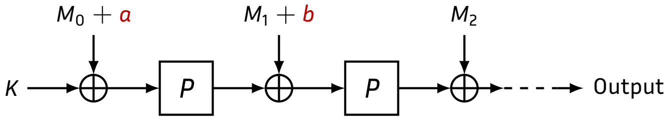
Parallel Construction



Plan of this Section


M_0
M_1
M_2
K
P
P
Output


$M_0 + a$
$M_1 + b$
M_2
K
P
P
Output



$$\Pr[Collision] = \mathsf{DP}(a, b)$$

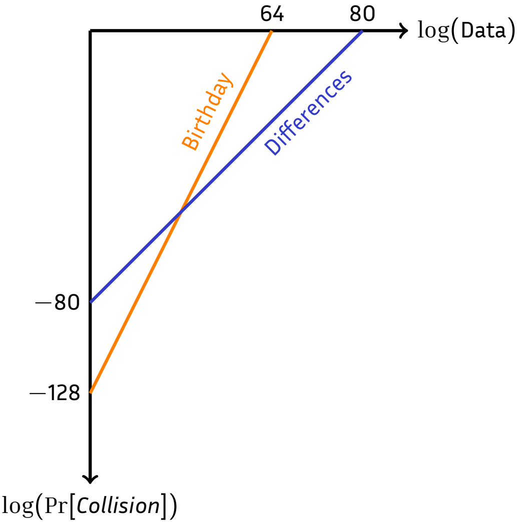
Framework

- P is "easier" than $P \circ P$.
- $\mathrm{DP}(a, b)$ is known for all $a, b \in \mathbb{F}_2^n$.

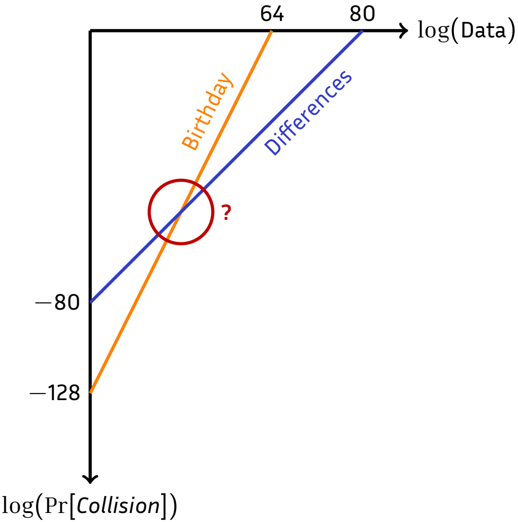
Goal

Find a collision in the output.

Birthday VS Difference



Birthday VS Difference



Not really Birthday

- (M_0, M_1),$(M_0 + a, M_1 + b)$,$(M'_0, M'_1), (M'_0 + a, M'_1 + b), \ldots$

Not really Birthday

- (M_0, M_1),$(M_0 + a, M_1 + b)$,(M'_0, M'_1), $(M'_0 + a, M'_1 + b)$, ...
- $\Pr[Collision] = \delta_P(a, b)$

Not really Birthday

- (M_0, M_1),$(M_0 + a, M_1 + b)$,(M'_0, M'_1), $(M'_0 + a, M'_1 + b), \ldots$
- $\Pr[\mathit{Collision}] = \delta_P(a, b)$
- (M_0, M_1),$(M_0 + a, M_1 + b)$, (M'_0, M'_1),$(M'_0 + a, M'_1 + b), \ldots$

Not really Birthday

- (M_0, M_1),$(M_0 + a, M_1 + b)$,(M'_0, M'_1), $(M'_0 + a, M'_1 + b), \ldots$
- $\Pr[Collision] = \delta_P(a, b)$
- (M_0, M_1),$(M_0 + a, M_1 + b)$, (M'_0, M'_1),$(M'_0 + a, M'_1 + b), \ldots$
- If $M_0 = M'_0$ or $M_0 = M'_0 + a$ then $\Pr = 0$, else $\Pr = 2^{-n}$.

Not really Birthday

- (M_0, M_1),$(M_0 + a, M_1 + b)$,(M'_0, M'_1), $(M'_0 + a, M'_1 + b), \ldots$
- $\Pr[\mathit{Collision}] = \delta_P(a, b)$
- (M_0, M_1),$(M_0 + a, M_1 + b)$, (M'_0, M'_1),$(M'_0 + a, M'_1 + b), \ldots$
- If $M_0 = M'_0$ or $M_0 = M'_0 + a$ then $\Pr = 0$, else $\Pr = 2^{-n}$.

Choose carefully the messages in a specific subspace...

Using Covering Vector spaces

$\langle (a_1, b_1), (a_2, b_2), \ldots, (a_v, b_v) \rangle = V$ such that

$$\mathrm{Moy}_V = \sum_{(a,b) \in V} \delta_{a,b} > \delta \,.$$

By making this strategy:

$$\begin{array}{c} M_0, M_1 \\ M_0 + a_1, M_1 + b_1 \\ M_0 + a_2, M_1 + b_2 \\ M_0 + a_1 + a_2, M_1 + b_1 + b_2 \\ \vdots \\ M_0 + \sum a_i, M_1 + \sum b_i \end{array}$$

$$\begin{array}{c} M'_0, M'_1 \\ M'_0 + a_1, M'_1 + b_1 \\ M'_0 + a_2, M'_1 + b_2 \\ M'_0 + a_1 + a_2, M'_1 + b_1 + b_2 \\ \vdots \\ M'_0 + \sum a_i, M'_1 + \sum b_i \end{array}$$

....

Using Covering Vector spaces

We get

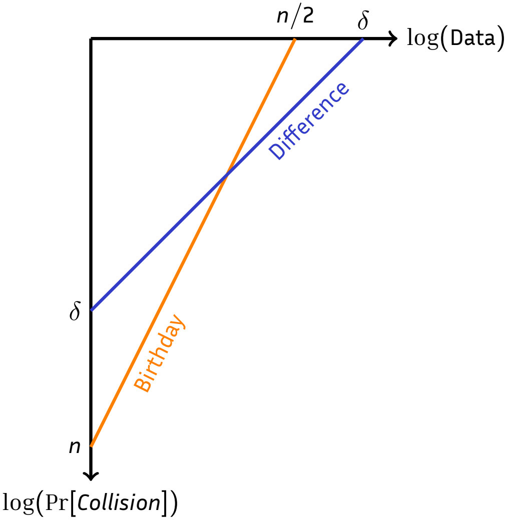
$$\Pr[Collision] = n_{block} \times \mathrm{Moy}_V \times 2^{v-1} + \frac{1}{2^n} \times \binom{n_{block}}{2} 2^{2v}$$

$$\Pr[Collision] = \frac{D}{2} \times \mathrm{Moy}_V + \frac{D^2}{2^n}$$

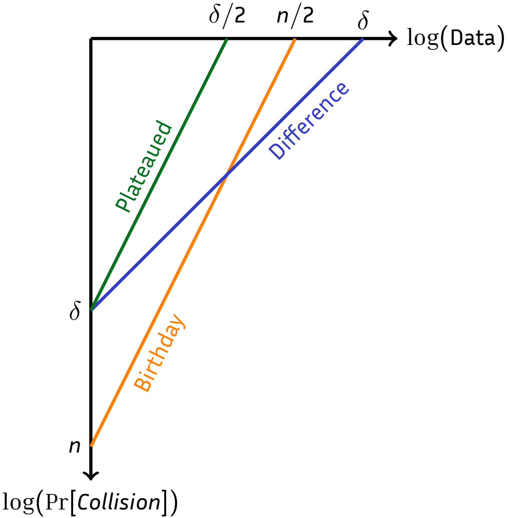
And

$$\mathrm{Moy}_V >> \delta\,??$$

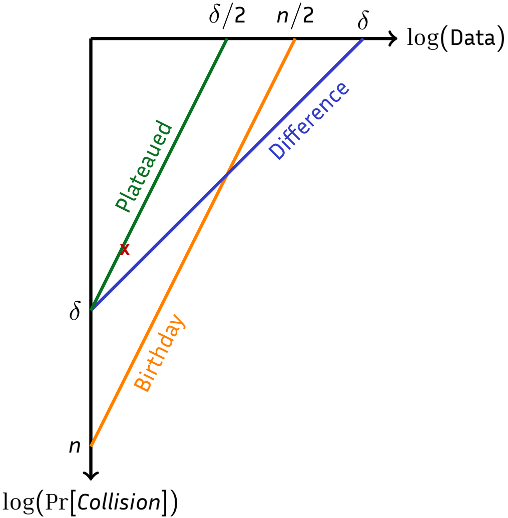
What could possibly be wrong?



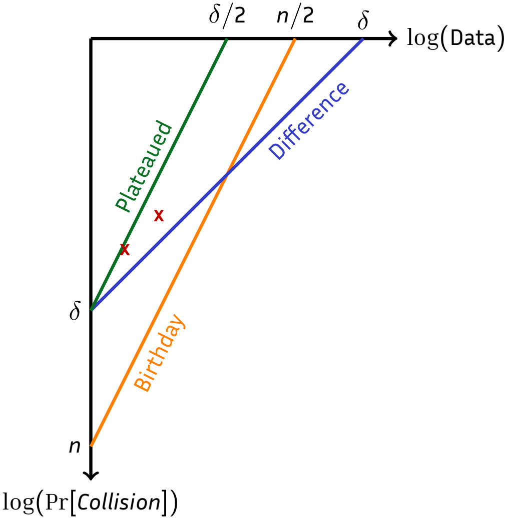
What could possibly be wrong?



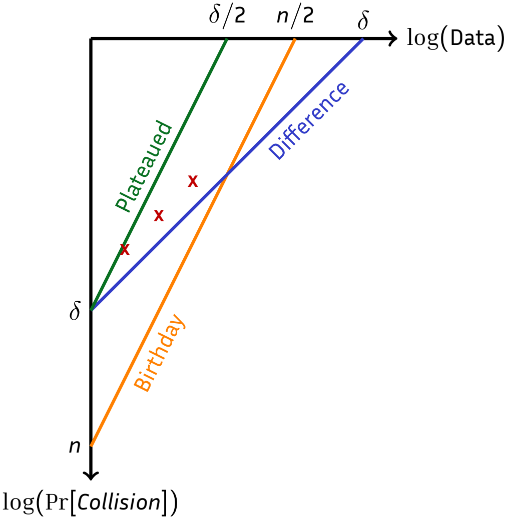
What could possibly be wrong?



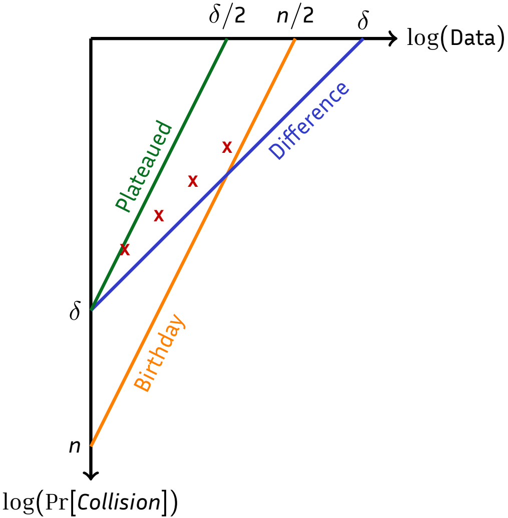
What could possibly be wrong?



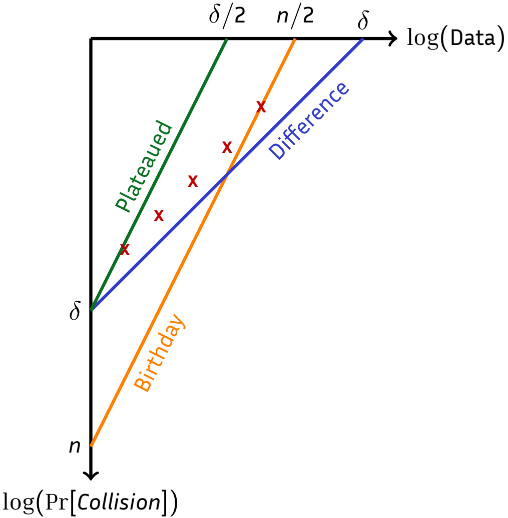
What could possibly be wrong?



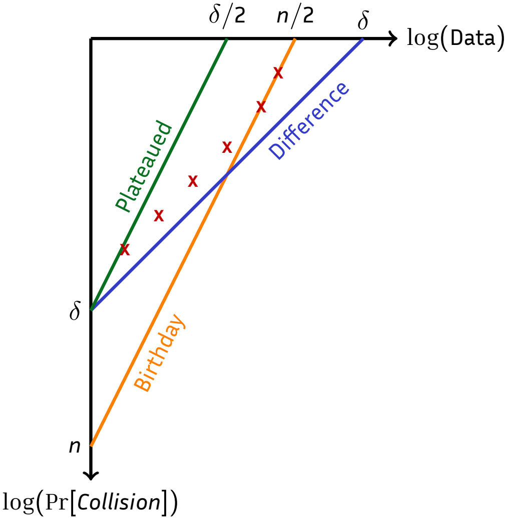
What could possibly be wrong?



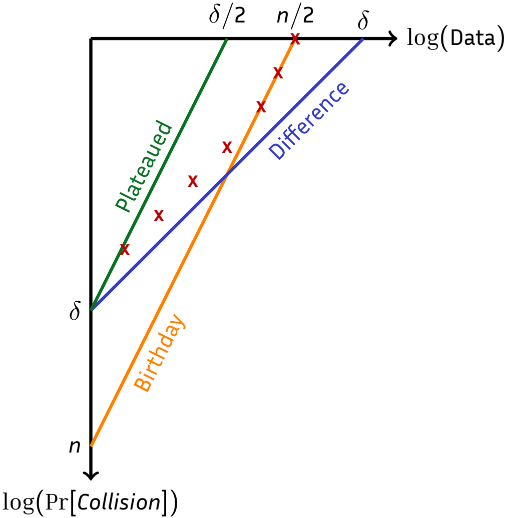
What could possibly be wrong?



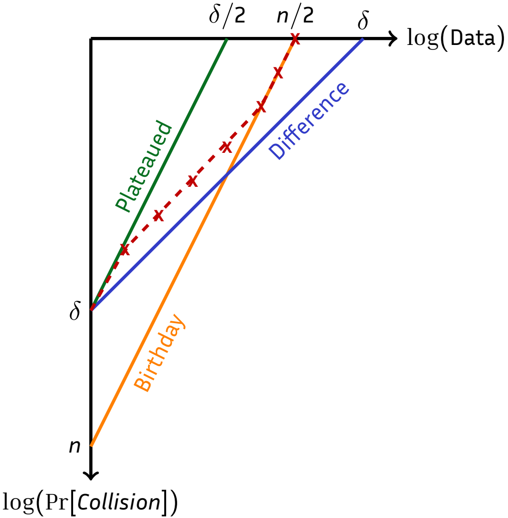
What could possibly be wrong?



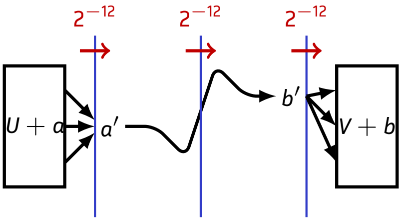
What could possibly be wrong?



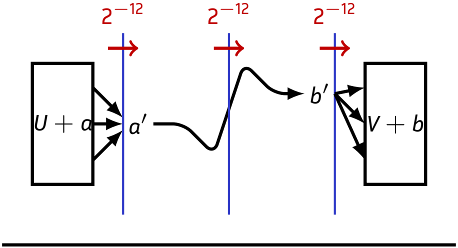
What could possibly be wrong?



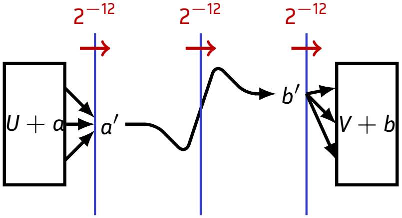
In Practice: XooDoo

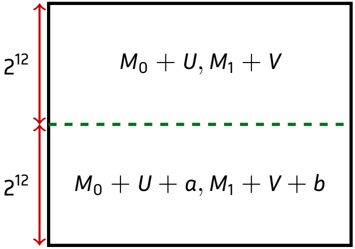


In Practice: XooDoo

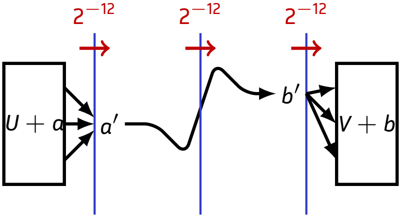


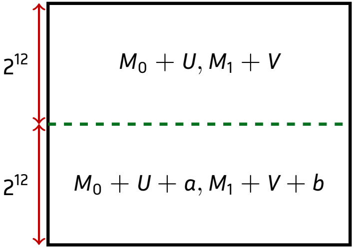
In Practice: XooDoo





In Practice: XooDoo

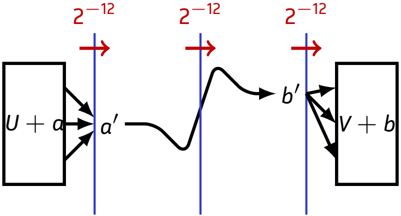




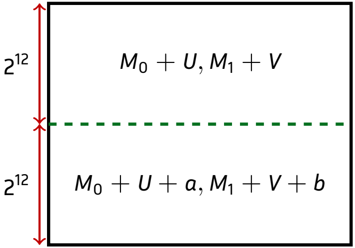
$$\text{Moy} \approx 2^{-24}$$

$$\Pr[Collision] = 2^{12} \times 2^{-24}$$

In Practice: XooDoo



$$\text{Moy} \approx 2^{-24}$$



$$\Pr[Collision] = 2^{12} \times 2^{-24}$$
$$>> 2^{13} \times 2^{-36}$$

Experiments on 3-rounds XooDoo

- a' touch 6 different S-boxes;
- Apply techniques to a subspace of dimension $6 \times 3 = 18$.

Experiments on 3-rounds XooDoo

- a' touch 6 different S-boxes;
- Apply techniques to a subspace of dimension $6 \times 3 = 18$.

But,

If we obtain a collision in a set, we obtain 8 collisions.

Experiments on 3-rounds XooDoo

On the choice of the subspace of dimension 12, 3 millions random sets of size 2^{13}.

0	1	2	3	4	5	6	7
*	200	65	17	8	0	1	0

Experiments on 3-rounds XooDoo

On the choice of the subspace of dimension 12, 3 millions random sets of size 2^{13}.

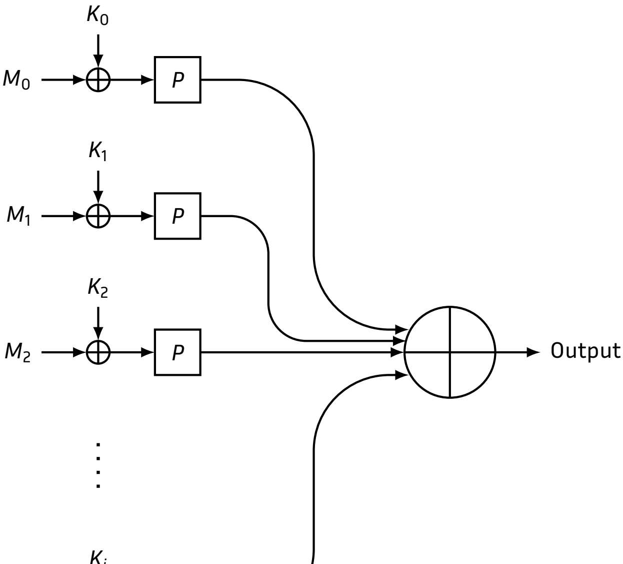
0	1	2	3	4	5	6	7
*	200	65	17	8	0	1	0

The probability of getting A collision is then smaller than expected...

Where this behaviour comes from?

Plan of this Section

New Criteria: Squared pseudo-Walsh Coefficient



Framework

- P operates on n-bit words;
- indepent keys with $|K_i| = |M_i|$;
- $\delta_P(a, b)$ known for all $a, b \in \mathbb{F}_2^n$.

Framework

- P operates on n-bit words;
- indepent keys with $|K_i| = |M_i|$;
- $\delta_P(a, b)$ known for all $a, b \in \mathbb{F}_2^n$.

Goal

Find a collision in the output

New criteria

Let $M = (M_0, M_1, \ldots, M_i)$ and $M' = (M'_0, M'_1, \ldots, M'_j)$.

New criteria

Let $M = (M_0, M_1, \ldots, M_i)$ and $M' = (M'_0, M'_1, \ldots, M'_j)$.

$$\sum_{\alpha=0}^{i} P(M_\alpha + K_\alpha) + \sum_{\beta=0}^{j} P(M'_\beta + K_\beta) = F_j(M, M') + F_{i,j}(M)$$

where

$$F_j(M, M') = \sum_{\alpha=0}^{j} (P(M_\alpha + K_\alpha) + P(M'_\alpha + K_\alpha))$$

and

$$F_{i,j}(M) = \sum_{\beta=j+1}^{i} P(M_\beta + K_\beta)\,.$$

New criteria

Let $M = (M_0, M_1, \ldots, M_i)$ and $M' = (M'_0, M'_1, \ldots, M'_j)$.

$$\sum_{\alpha=0}^{i} P(M_\alpha + K_\alpha) + \sum_{\beta=0}^{j} P(M'_\beta + K_\beta) = F_j(M, M') + F_{i,j}(M)$$

where

$$F_j(M, M') = \sum_{\alpha=0}^{j} \left(P(M_\alpha + K_\alpha) + P(M'_\alpha + K_\alpha)\right)$$

and

$$F_{i,j}(M) = \sum_{\beta=j+1}^{i} P(M_\beta + K_\beta)\,.$$

Goal

$$p = \Pr\left[\sum_{\alpha=0}^{i} P(M_\alpha + K_\alpha) = \sum_{\beta=0}^{j} P(M'_\beta + K_\beta)\right]$$

New criteria

$$p = \Pr\left[F_j(M, M') + F_{i,j}(M) = 0\right]$$

New criteria

$$p = \Pr\left[F_j(M, M') + F_{i,j}(M) = 0\right]$$

Independent keys implies

$$\Pr\left[F_{i,j}(M) = A\right] = 2^{-n},$$

$$p = 2^{-n} \sum_{A \in \mathbb{F}_2^n} \Pr\left[F_j(M, M') = A\right].$$

New criteria

$$p = \Pr\left[F_j(M, M') + F_{i,j}(M) = 0\right]$$

Independent keys implies

$$\Pr\left[F_{i,j}(M) = A\right] = 2^{-n},$$

$$p = 2^{-n} \sum_{A \in \mathbb{F}_2^n} \Pr\left[F_j(M, M') = A\right].$$

$$p = 2^{-n} \sum_{A, b_1, \ldots, b_j \in \mathbb{F}_2^n} \mathrm{DP}(a_0, A + b_1 + \cdots + b_j)\mathrm{DP}(a_1, b_1)\mathrm{DP}(a_2, b_2) \cdots \mathrm{DP}(a_j, b_j).$$

Pseudo-Walsh Transform

We pose, for $a \in \mathbb{F}_2^n$,

$$\mathcal{W}_P^a(\mu) = \sum_{b \in \mathbb{F}_2^n} (-1)^{b \cdot \mu} \mathrm{DP}(a, b) .$$

Then, we have

$$\mathrm{DP}(a, b_0) = \frac{1}{2^n} \sum_{\mu \in \mathbb{F}_2^n} (-1)^{b_0 \cdot \mu} \mathcal{W}_P^a(\mu) .$$

Pseudo-Walsh Transform

We pose, for $a \in \mathbb{F}_2^n$,

$$\mathcal{W}_P^a(\mu) = \sum_{b \in \mathbb{F}_2^n} (-1)^{b \cdot \mu} \mathrm{DP}(a, b) \, .$$

Then, we have

$$\mathrm{DP}(a, b_0) = \frac{1}{2^n} \sum_{\mu \in \mathbb{F}_2^n} (-1)^{b_0 \cdot \mu} \mathcal{W}_P^a(\mu) \, .$$

$$(\mathrm{DP}(a_0) * \mathrm{DP}(a_1)(b_0) = \sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b) \mathrm{DP}(a_1, b + b_0)$$

Pseudo-Walsh Transform

We pose, for $a \in \mathbb{F}_2^n$,

$$\mathcal{W}_P^a(\mu) = \sum_{b \in \mathbb{F}_2^n} (-1)^{b \cdot \mu} \mathrm{DP}(a, b) \, .$$

Then, we have

$$\mathrm{DP}(a, b_0) = \frac{1}{2^n} \sum_{\mu \in \mathbb{F}_2^n} (-1)^{b_0 \cdot \mu} \mathcal{W}_P^a(\mu) \, .$$

$$(\mathrm{DP}(a_0) * \mathrm{DP}(a_1)(b_0) = \sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b)\mathrm{DP}(a_1, b + b_0)$$

- Associative
- Commutative
- Bilinear

Pseudo-Walsh Tranform

$$p = \mathcal{W}_P^{-1}\left(\mathcal{W}_P[\mathrm{DP}(a_0)](\mu)\mathcal{W}_P[\mathrm{DP}(a_1)](\mu)\cdots\mathcal{W}_P[\mathrm{DP}(a_j)](\mu)\mathcal{W}_P[\mathrm{DP}(uni)](\mu)\right)(0)$$

which can be expressed with

$$p = \frac{1}{2^n}\sum_{\mu\in\mathbb{F}_2^n}\mathcal{W}_P[\mathrm{DP}(a_1)](\mu)\cdots\mathcal{W}_P[\mathrm{DP}(a_j)](\mu)\mathcal{W}_P[\mathrm{DP}(uni)](\mu)$$

$$\sum_b \mathrm{DP}(a_i, b) = 1\,.$$

This means then exactly that for all $\mu \in \mathbb{F}_2^n$ and for all a_i, $|\mathcal{W}_P[\mathrm{DP}(a_i)](\mu)| \leq 1$.

Pseudo-Walsh Transform

Moreover, as $\mathrm{DP}(uni) = 2^{-n}$, we have $\mathcal{W}_P[\mathrm{DP}(uni)](\mu) = 0$ for all $\mu \neq 0$ and $\mathcal{W}_P[\mathrm{DP}(uni)](0) = 1$. This $\mathcal{W}_P[\mathrm{DP}(uni)]$ appears if and only if the size of the messages are different. If this is the case, we obtain the probability

$$p = \frac{1}{2^n}\mathcal{W}_P[\mathrm{DP}(a_1)](0)\mathcal{W}_P[\mathrm{DP}(a_2)](0)\cdots\mathcal{W}_P[\mathrm{DP}(a_j)](0)$$

but for all 2^n differential vector,

$$\mathcal{W}_P[\mathrm{DP}(a_2)](0) = \sum_{b\in\mathbb{F}_2^n} \mathrm{DP}(a, b) = 1$$

Hence, when two messages are of different size, the probability that guetting a collision is exactly 2^{-n}.

Best choice is when $a_1 = a_2$

$$\sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b)\mathrm{DP}(a_1, b) = \frac{1}{2}\left(\sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b)^2 + \mathrm{DP}(a_1, b)^2\right)$$

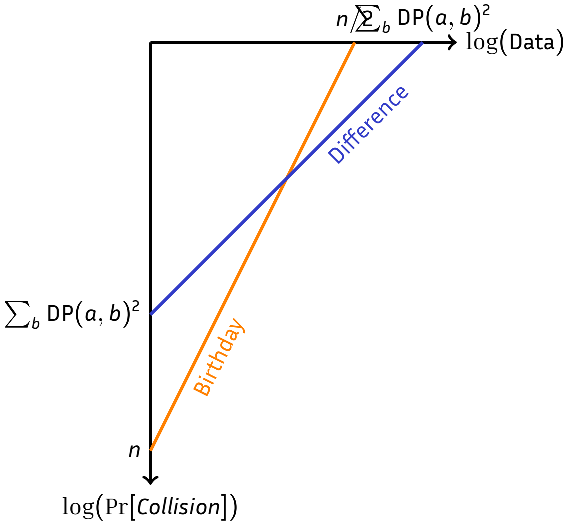
Best choice is when $a_1 = a_2$

$$\sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b)\mathrm{DP}(a_1, b) = \frac{1}{2}\left(\sum_{b \in \mathbb{F}_2^n} \mathrm{DP}(a_0, b)^2 + \mathrm{DP}(a_1, b)^2\right)$$

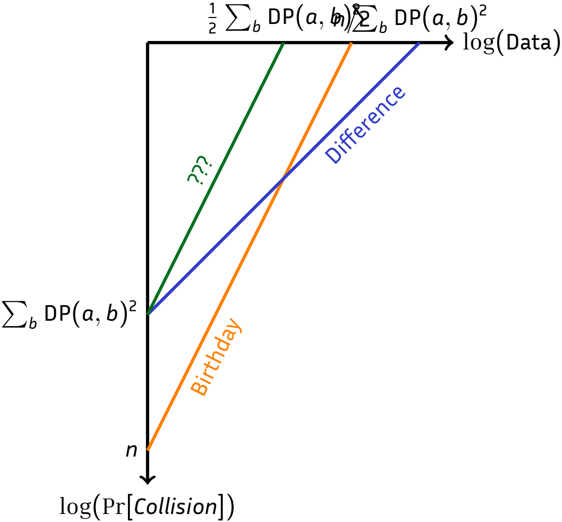
So we focus on

$$\max_{a \in \mathbb{F}_2^n} \sum_{\in \mathbb{F}_2^n} \mathrm{DP}(a, b)^2$$

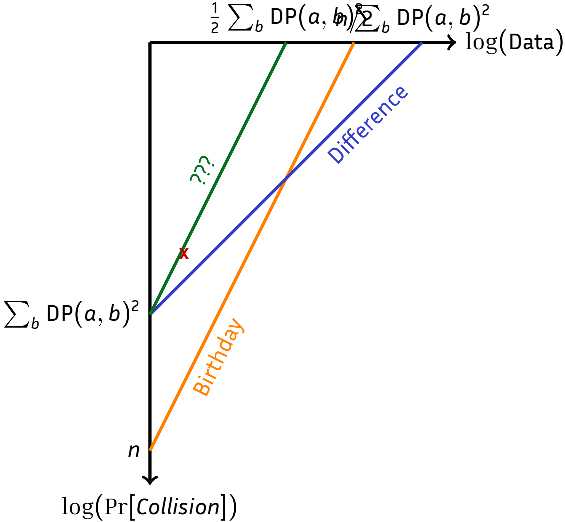
Tight or not tight?



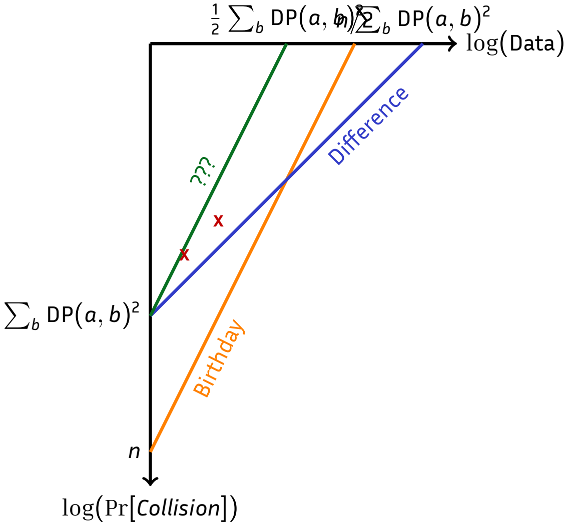
Tight or not tight?



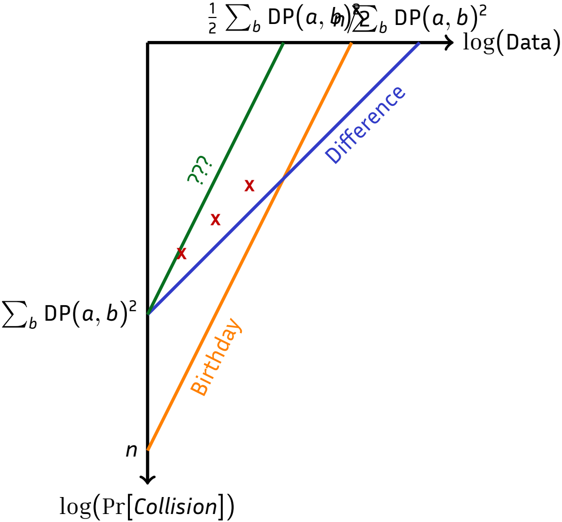
Tight or not tight?



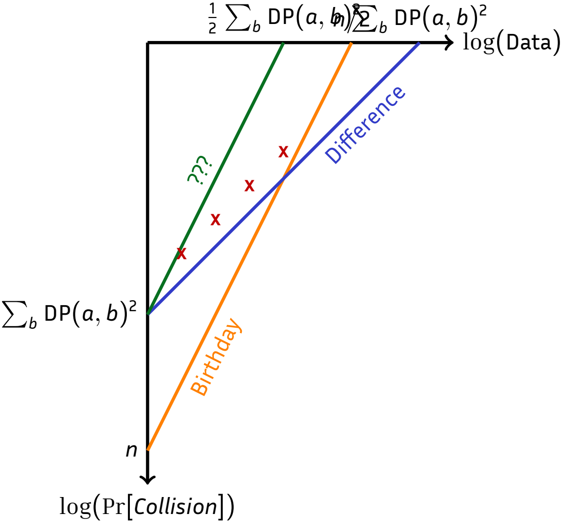
Tight or not tight?



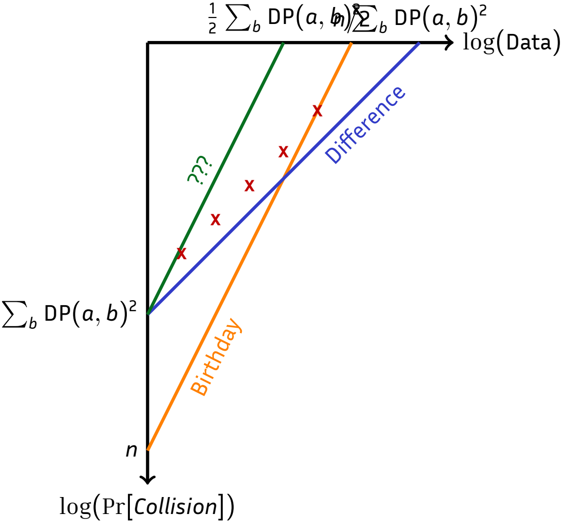
Tight or not tight?



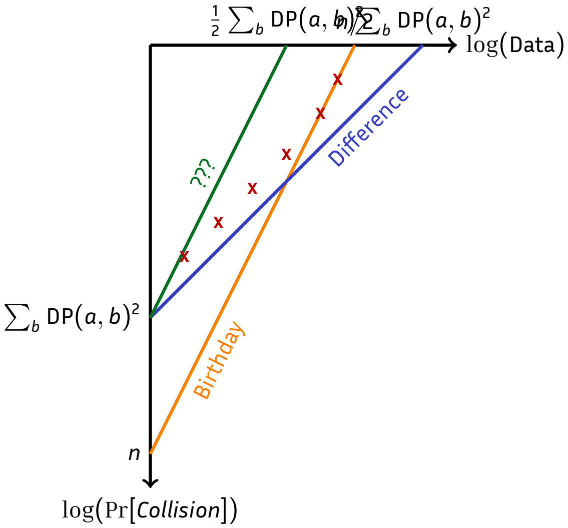
Tight or not tight?



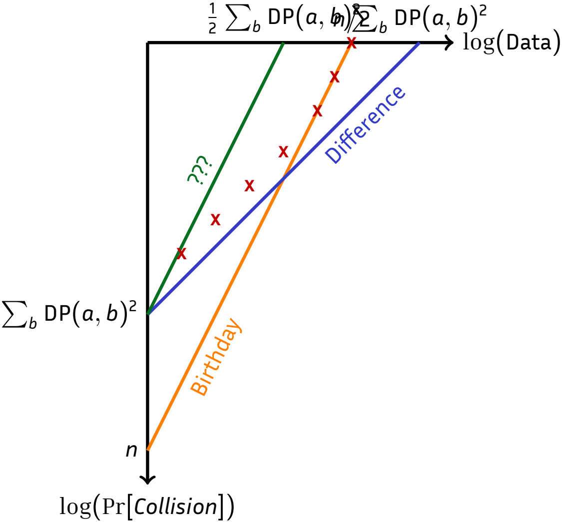
Tight or not tight?



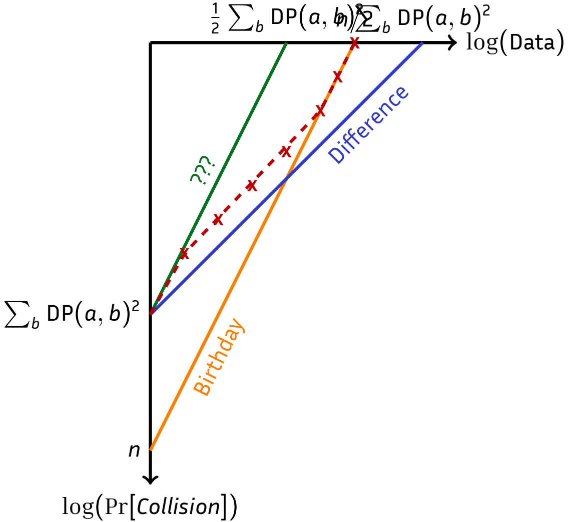
Tight or not tight?



Tight or not tight?



Tight or not tight?



Tight in number of queries

Let Δ such that $\sum_b \mathrm{DP}(\Delta, b)^2$ is maxmal.

Tight in number of queries

Let Δ such that $\sum_b \mathrm{DP}(\Delta, b)^2$ is maxmal.

$M_0 + \Delta$	$M_0 + \Delta$	$M_0 + \Delta$	$M_0 + \Delta$	$M_0 + \Delta$	$M_0 + \Delta$	$M_0 + \Delta$
$M_1 + \Delta$	M_1	M_1	M_1	M_1	M_1	M_1
M_2	$M_2 + \Delta$	M_2	M_2	M_2	M_2	M_2
M_3	M_3	$M_3 + \Delta$	M_3	M_3	M_3	M_3
M_4	M_4	M_4	$M_4 + \Delta$	M_4	M_4	M_4
M_5	M_5	M_5	M_5	$M_5 + \Delta$	M_5	M_5
M_6	M_6	M_6	M_6	M_6	$M_6 + \Delta$	M_6
M_7	M_7	M_7	M_7	M_7	M_7	$M_7 + \Delta$
M_8	M_8	M_8	M_8	M_8	M_8	M_8
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Tight in number of blocks?

- The same technique as the parallel one can be applied (win one round)
- Find a vector space... You know the rest

Using the average and not the max

$$M_0 + \Delta_0, M_1 + \Delta_0$$
$$M_0 + \Delta_1, M_1 + \Delta_1$$
$$M_0 + \Delta_2, M_1 + \Delta_2$$
$$M_0 + \Delta_3, M_1 + \Delta_3$$
$$M_0 + \Delta_4, M_1 + \Delta_4$$
$$M_0 + \Delta_5, M_1 + \Delta_5$$
$$M_0 + \Delta_6, M_1 + \Delta_6$$

$$\Pr[Collision] = \text{Moy}(\sum \delta^2) \times D^2$$

Using the average and not the max

$$M_0 + \Delta_0, M_1 + \Delta_0$$
$$M_0 + \Delta_1, M_1 + \Delta_1$$
$$M_0 + \Delta_2, M_1 + \Delta_2$$
$$M_0 + \Delta_3, M_1 + \Delta_3$$
$$M_0 + \Delta_4, M_1 + \Delta_4$$
$$M_0 + \Delta_5, M_1 + \Delta_5$$
$$M_0 + \Delta_6, M_1 + \Delta_6$$

$$\Pr[Collision] = \mathrm{Moy}(\sum \delta^2) \times D^2$$

$$\mathrm{Moy}(\sum \mathrm{DP}^2) > \frac{1}{2^n}$$

Plan of this Section

Comparisons

Parallel	Serial
2 blocks is "easier"	2 blocks is the best
DP	$\sum \mathrm{DP}^2$
bound	bound
not tight if cheat	tight if cheat
almost tight for D small	almost tight for D small

Questions

- Experiments for 3-rounds XooDoo parallel ?
- Find greater vector spaces ?
- DP is easy, but what about $\sum \mathrm{DP}^2$?