Qualcomm

Cryptanalysis of Ascon - An Information Theoretic Perspective - A Position Paper

Nicolas T. Courtois, Florian Caullery, Fred Amiel, William Whyte
QCT Security Architecture, Sophia Antipolis, Qualcomm France S.A.R.L.
Technical Standards, Qualcomm Technologies, MA, US

Agenda

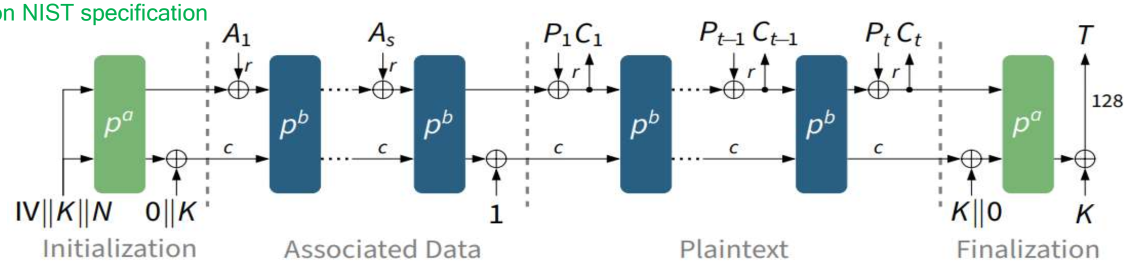
- New encryption standard!
- A novel approach to analyze the security of Ascon
- Strong and weak S-boxes: do we need two theories?
- How attacks can be mapped to basic properties
- Prediction of attacks and undesirable properties
- Contemplating the gap and combinatorial explosion
- Open problems

Background - Facts

Feb 2023	NIST selects ASCON to become a new encryption standard expected to be in use for many decades to come.
June 2023	NIST hosting the 6th LWC Workshop: NIST is soliciting research and discussion papers, surveys, presentations, panel proposals, case studies related to ASCON including Security results on the Ascon family + a call for public comments. Submission deadline = 1st May 2023.

- Through this presentation, our focus is to engage with NIST to support the community effort to develop the best possible encryption standard. We need to optimize the security and yet minimize HW implementation cost.
- A sensible analysis of security of Ascon should be:
  - Forward-looking: we cannot contemplate just some attacks already studied.
  - Robust: we should not just look for some rare and exceptional events (best case). We need methods to study understand what happens on average. We claim that there exists a ROBUST transparent way for evaluating a security of a cipher seen as a communications channel trying to maximize the "channel capacity".
  - Relevant: Several already known attacks CAN be modelled in terms of intersections of spaces of some "undesirable properties".
- Methodology: "transforming a constant into a variable"
  - replacing the S-box by several candidates, weak or strong,
  - showing how the attacks scale and showing that their existence can be reliably predicted from the following principles:
    - conditional entropy, mutual information and, discrete combinatorial events weighted by probabilities, which exist in small finite numbers because the S-box is tiny.

Ascon NIST specification



eprint.iacr.org 2016/490

Table 10: Summary of attacks on ASCON.

Type	Rounds	Time	Method
Key Recovery	6/12	2^{66}	Cube-like
Key Recovery	5/12	2^{35}	Cube-like
Key Recovery	5/12	2^{36}	Differential-Linear
Key Recovery	5/12	2^{58} or $2^{127.99}$	Truncated/Improbable
Key Recovery	4/12	2^{18}	Differential-Linear
Key Recovery	4/12	3^{48}	Truncated/Impossible
Forgery	4/12	2^{101}	Differential
Forgery	3/12	2^{33}	Differential

Qualcomm brought you foundational communications technologies.

Can information theory help cryptographers to design better ciphers?

1G 2G 3G 4G 5G

Analysis of Ascon

Agenda

Modelling Ascon as a Communications Channel

For 9 years Ascon was studied and seems very secure. All because of “strong diffusion”.
Any simple perturbation expand very quickly.
Game over = no hope to attack Ascon???

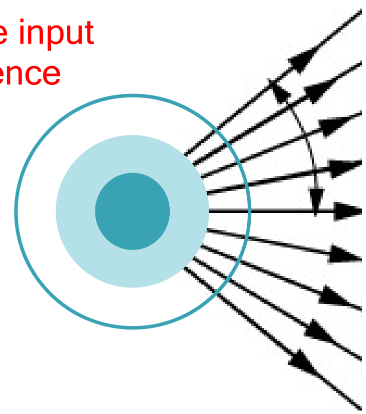
It should be critical to consider attacks that AGGREGRATE input perturbations.

[Tezcan 2014]

Table 2: Undisturbed Bits of ASCON’s S-box.

Input Difference	Output Difference	Input Difference	Output Difference
00001	?1???	10000	?10??
00010	1???1	10001	10??1
00011	???0?	10011	0???0
00100	??110	10100	0?1??
00101	1????	10101	????1
00110	????1	10110	1????
00111	0??1?	10111	????0
01000	??11?	11000	??1??
01011	???1?	11100	??0??
01100	??00?	11110	?1???
01110	?0???	11111	?0???
01111	?1?0?		

simple input difference

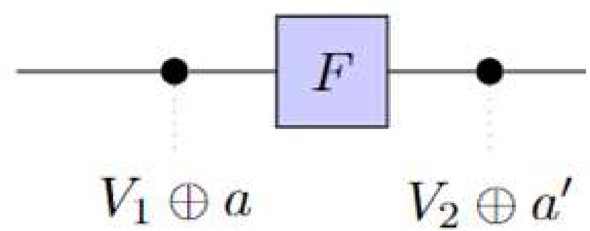


"diffusion cone"

nb. of active bits at output!

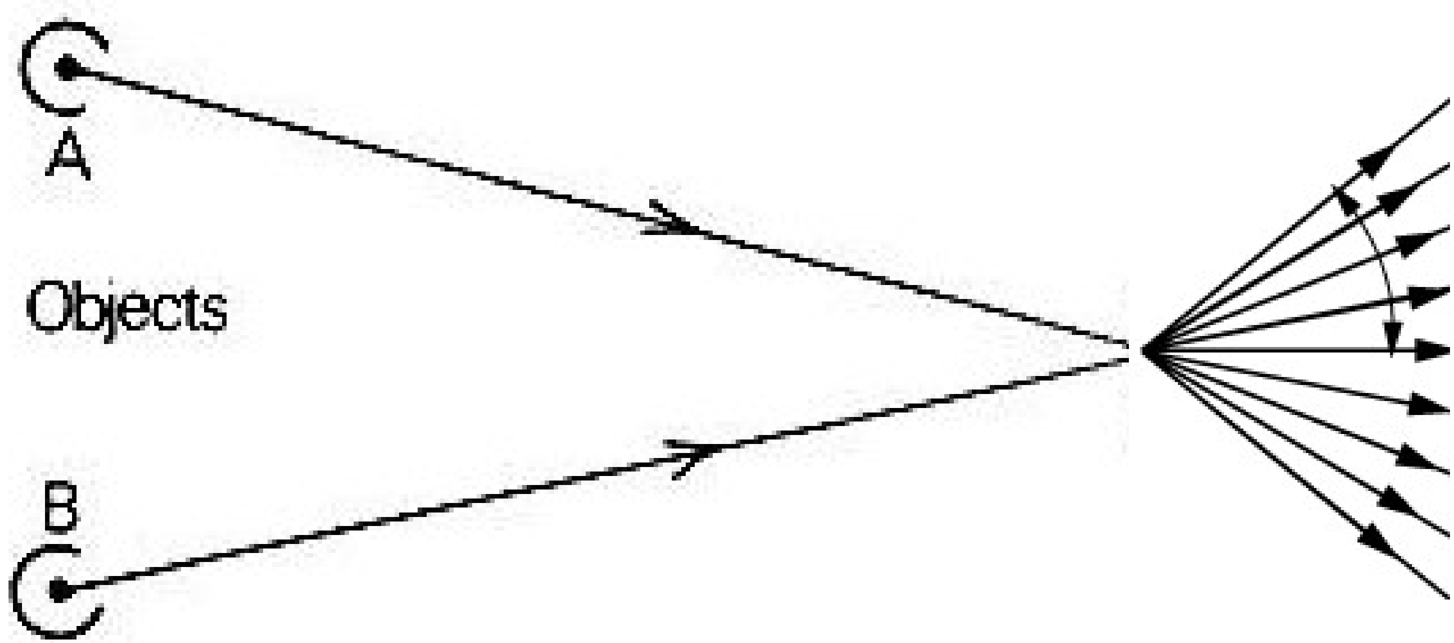
Cipher	r_e	d
ASCON [Dob+16]	3	298
GIFT [Ban+17]	3	60
KECCAK [Ber+11]	2	546
PRESENT [Bog+07]	3	43
PRIDE [Alb+14]	2	31
QARMA [Ava17]*	2	36

Grassi, Rechberger and Rønjom, 2016,
Subspace Trail Cryptanalysis



Philosophy : Aggregate Perturbations

- Can several perturbations converge somewhat? Attacker does either A or B.
- We need to improve the "channel capacity" to increase information conveyed or the likelihood of detection.

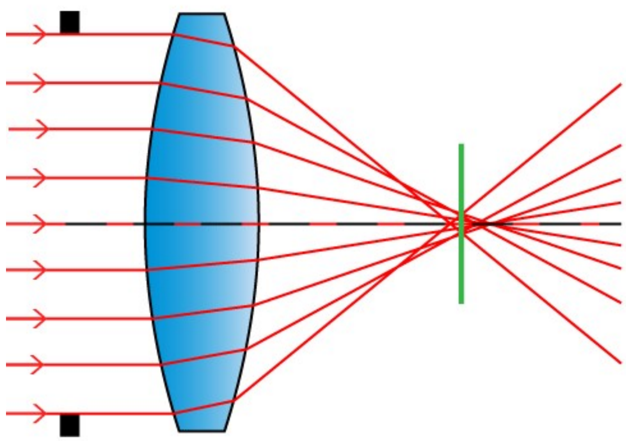


A Known Problem – Analogy with Optics

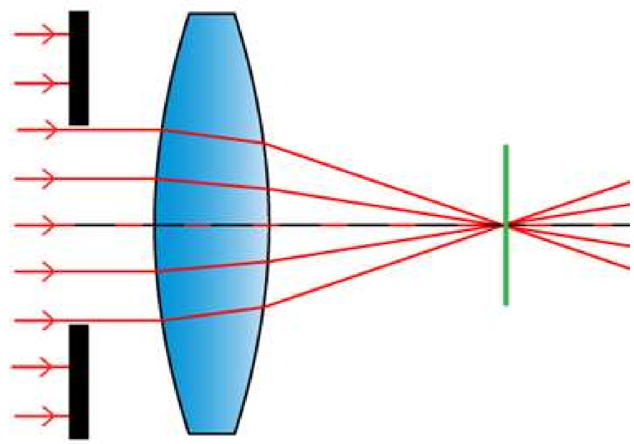
Not if we have TOO many sources!

Must restrict the input diversity.

Wide Open Aperture

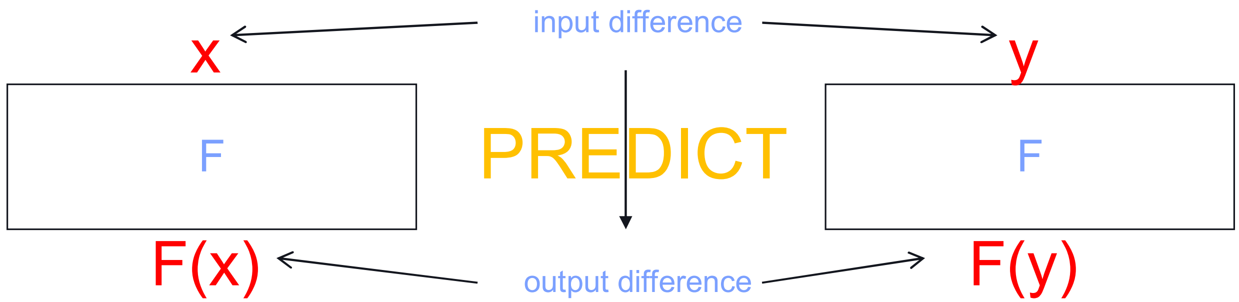


Stopped Down Aperture



a better transmission channel!

Study of Conditional Entropy and MI



$$H(F(x)-F(y) \mid x-y)$$

should be large

Prediction ability <=> Mutual Information = MI =

should be small

$$MI(F(x)-F(y);\ x-y)$$

Ascon S-Box - Proof of Concept

best =

Ent(oD) = 2.00 bits when x-y=4

Ent(oD) = 2.00 bits when x-y=12

Ent(oD) = 2.00 bits when x-y=16

Ent(oD) = 2.00 bits when x-y=17

av. Output Δ Ent = 2.00 bits

single differences
vs.
quadruple differences

we compute the entropy for the output difference

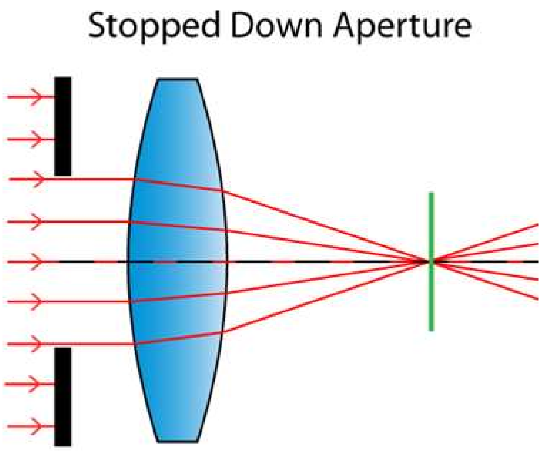
Ent(oD) = 3.69 bits when x-y \in {1,3,16,18}

Ent(oD) = 3.69 bits when x-y \in {4,8,20,24}

Ent(oD) = 3.66 bits when x-y \in {5,16,17,21}

We gain something:

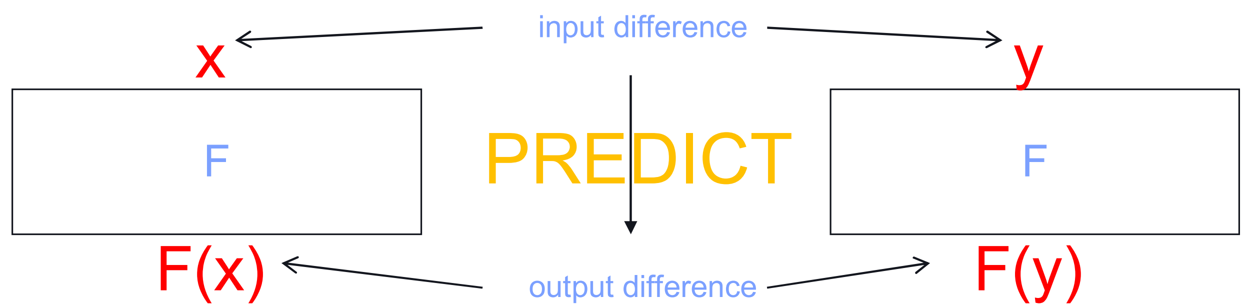
av. Output Δ Ent = only 3.69 bits instead of 4 bits



Agenda

Towards [Difference] Prediction

Prediction approach based on MI = Mutual Information



for 5 bits we get:

	Fides x^{-1} APN	RP*	Ascon/ Keccak
Entropy:	3.875	3.5	4.10
DMI = Differential Mutual Information	1.125	1.4-1.7	1.90

2x bad

*RP=Random Permutation

DDT is holographic

- Fundamental Observation:
- The fact that there are many zeros and some large integers at specific locations are actually the SAME.
- one cannot happen without the other.
- Not new - many papers on conflicting security criteria in ciphers and Boolean functions
- What is NEW? Showing they two properties coincide “EN MASSE” and in probability - they correspond to TWO large precisely measurable percentages of “basic events” which have a large intersection inside an objective information theoretic averaged measure of quality which is simply a SUM weighted by probabilities.

  Overlap between different attacks is a totally objective feature.

DDT = Difference Distribution Table

```
AS = SBox(4,11,31,20,26,21,9,2,27,5,8,18,29,3,6,28,30,19,7,14,0,13,17,24,16,12,1,25,22,10,15,23);#Ascon
```

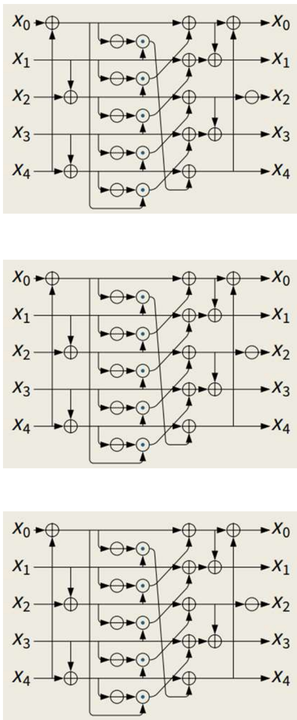
```
print(AS.difference_distribution_table())
```

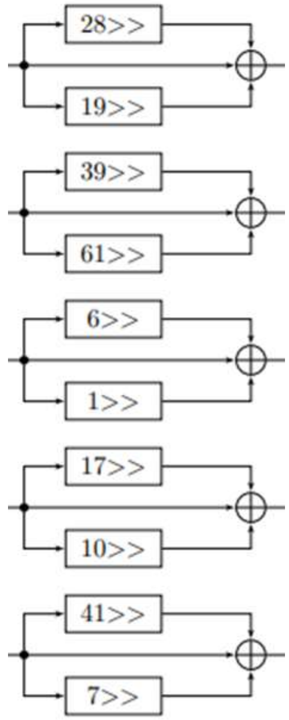
```
[32  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
[ 0  0  0  0  0  0  0  0  0  4  0  4  0  4  0  4  0  0  0  0  0  0  0  0  4  0  4  0  4  0  4  0]
[ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  4  0  4  0  4  0  4  0  4  0  4  0  4]
[ 0  4  0  0  0  4  0  0  0  4  0  0  0  4  0  0  4  0  0  0  4  0  0  0  4  0  0  0  4  0  0  0]
[ 0  0  0  0  0  0  8  0  0  0  0  0  0  0  8  0  0  0  0  0  0  0  8  0  0  0  0  0  0  0  8  0]
[ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  4  4  0  4  0  4  0  4  0  0  4  0  4]
[ 0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2]
[ 0  0  4  4  0  0  4  4  0  0  4  4  0  0  4  4  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
[ 0  0  0  0  0  0  4  4  0  0  0  0  0  0  4  4  0  0  0  0  0  0  4  4  0  0  0  0  0  0  4  4]
[ 0  2  0  2  2  0  2  0  2  0  2  0  0  2  0  2  2  0  2  0  0  2  0  2  0  2  0  2  2  0  2  0]
[ 0  2  2  0  2  0  0  2  0  2  2  0  2  0  0  2  0  2  2  0  2  0  0  2  0  2  2  0  2  0  0  2]
[ 0  0  2  2  0  0  2  2  0  0  2  2  0  0  2  2  0  0  2  2  0  0  2  2  0  0  2  2  0  0  2  2]
[ 0  8  0  0  0  0  0  0  8  0  0  0  0  0  0  0  8  0  0  0  0  0  0  0  0  8  0  0  0  0  0  0]
[ 0  2  0  2  0  2  0  2  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  0  2  0  2  0  2  0  2]
[ 0  4  4  0  4  0  0  4  0  0  0  0  0  0  0  0  0  4  4  0  4  0  0  4  0  0  0  0  0  0  0  0]
[ 0  0  0  0  0  0  0  0  4  4  0  0  4  4  0  0  0  0  0  0  0  0  0  0  4  4  0  0  4  4  0  0]
[ 0  0  0  0  0  0  0  0  0  8  0  8  0  0  0  0  0  0  0  0  0  0  0  0  8  0  8  0  0  0  0  0]
[ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  8  0  8  0  8  0  8  0  0  0  0  0  0  0  0]
[ 0  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  2  0  2  0  2  0  2  0  2  0  2  0  2  0  2  0]
[ 0  0  8  0  8  0  0  0  0  0  8  0  8  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
[ 0  0  0  0  4  4  4  4  0  0  0  0  4  4  4  4  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
[ 0  0  0  0  0  4  0  4  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0  0  0  0  0  4  0  4]
[ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2  2]
[ 0  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0]
[ 0  0  0  0  2  2  2  2  0  0  0  0  2  2  2  2  0  0  0  0  2  2  2  2  0  0  0  0  2  2  2  2]
[ 0  0  0  4  0  0  4  0  4  0  0  0  0  4  0  0  4  0  0  0  0  4  0  0  0  0  0  4  0  0  4  0]
[ 0  2  2  0  0  2  2  0  2  0  0  2  2  0  0  2  0  2  2  0  0  2  2  0  2  0  0  2  2  0  0  2]
[ 0  0  2  2  2  2  0  0  0  0  2  2  2  2  0  0  0  0  2  2  2  2  0  0  0  0  2  2  2  2  0  0]
[ 0  4  0  4  0  0  0  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0]
[ 0  0  0  4  0  4  0  0  4  0  0  0  0  0  4  0  4  0  0  0  0  0  4  0  0  0  0  4  0  4  0  0]
[ 0  0  0  0  0  0  0  0  2  2  2  2  2  2  2  2  0  0  0  0  0  0  0  0  2  2  2  2  2  2  2  2]
[ 0  0  4  4  4  4  0  0  0  0  0  0  0  0  0  0  0  0  4  4  4  4  0  0  0  0  0  0  0  0  0  0]
```

“white box security analysis” => leading to total awareness of basic facts which imply a larger family of attacks

Ascon Relies on Just ONE Tiny NL component

6,8,12 rounds like this:





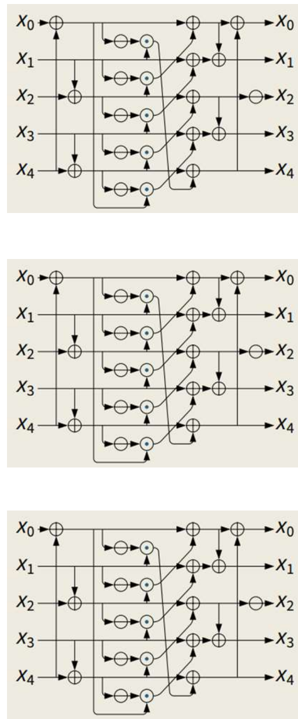
64x
columns
of 5 bits

5x
lines
of 64
bits

Key Problem

Prediction ability?

64x



Δ Δ

unrelated?

Claim: Ascon is not very strong in this respect compared to other encryption algorithms… Here is why.

So What?

Imagine that the attacker is trying to break the Ascon hash function
by a sophisticated guess then determine attack
[see work of Xiaoyun Wang, Marc Stevens, Leo Perrin etc etc]

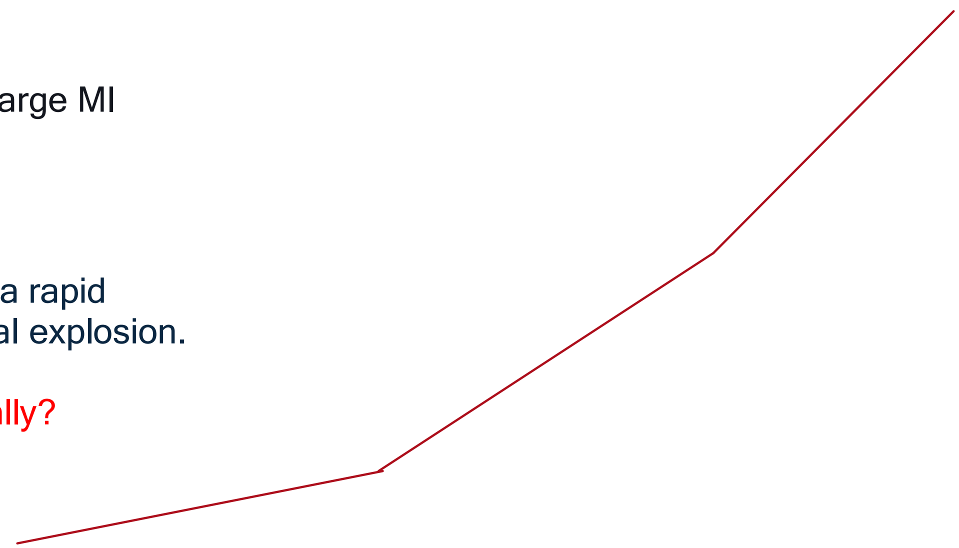
⇒ when MI is 2x bigger, the amount of information the attacker “already knows” doubles…

We intend to show [very early, just argue] and claim that a large MI has a dramatic impact enabling all of the following:

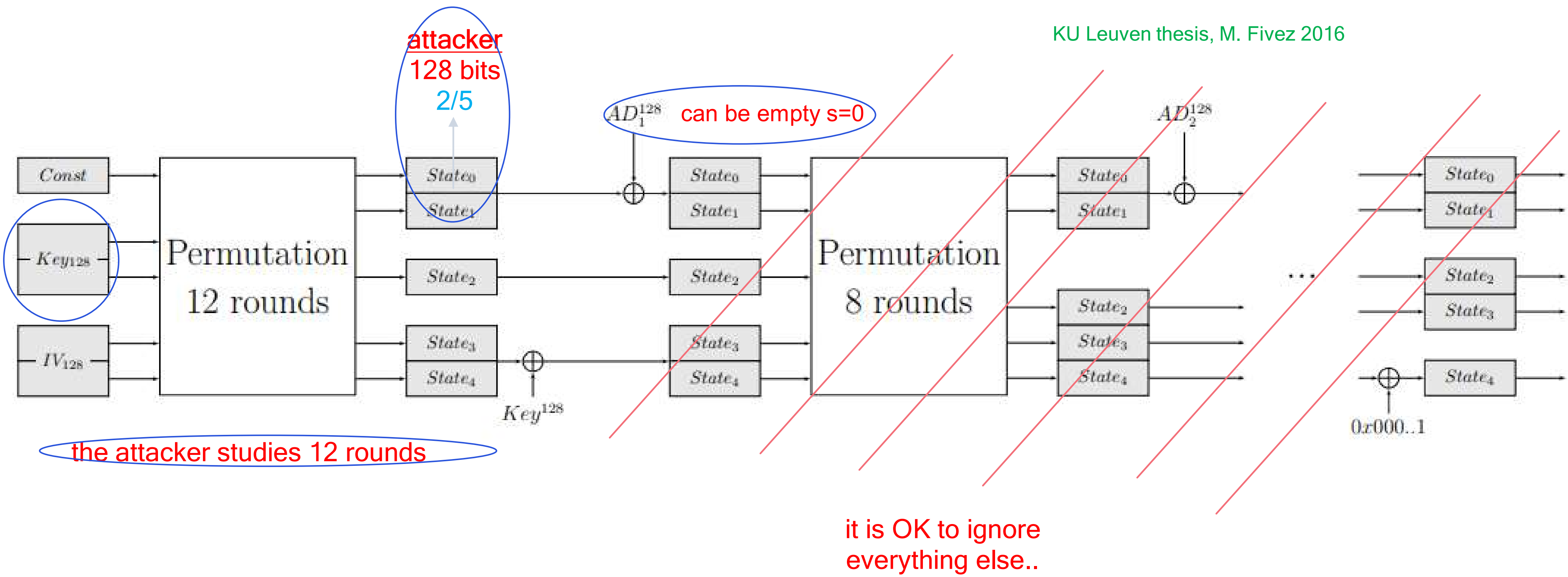
- all sorts of guess and determine attacks
- truncated differential attacks
- polynomial invariant attacks
- subspace trail attacks
- zero-sum and cube attacks

Claim: a rapid combinatorial explosion.

Really?



Ascon has very few rounds!! Example:



A common misconception in cryptanalysis

Is Ascon box good enough?

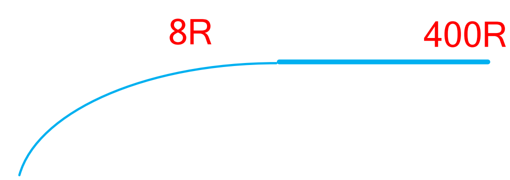
Claim that… "bad quality" NL mappings are OK if you have a large number of rounds.

The problem: Ascon does NOT have a large number of rounds.

???????????

Some attacks are such that additional rounds do NOT increase security

1. polynomial invariant attacks and affine space trail attacks …



2. some differential attacks: very surprising but real…
   Composability violation: Proof of concept:
   => Attack works equally well for say 8 and 400 rounds…

Home › Information Security and Cryptology – ICISC 2020 › Conference paper Springer Link

Can a Differential Attack Work for an Arbitrarily Large Number of Rounds?

Nicolas T. Courtois & Jean-Jacques Quisquater

Conference paper | First Online: 07 February 2021

Methodology - Example

Example: consider a 5-round truncated differential property on 29 bits of Ascon out of 320.

We can MAP in a precise exact and undeniable way this property seen as an enumeration of discrete cryptographic events to an information-theoretic measure of quality of

A. the direct product of the S-box with itself (parallel application)
   which has 100% predictable properties in terms of MI and mappings.
B. the linear layer which also has well-defined entropy and mutual information properties w.r.t diffusion.

THEN one can show that our enumeration of events accounts EXACTLY for 43% and 57 % of the MI / entropy in percentage and probability mass for A and B respectively.
All attacks comes back to few basic information theoretical facts!

From here we claim that there exists a simple and robust methodology for evaluating the quality of ciphers.

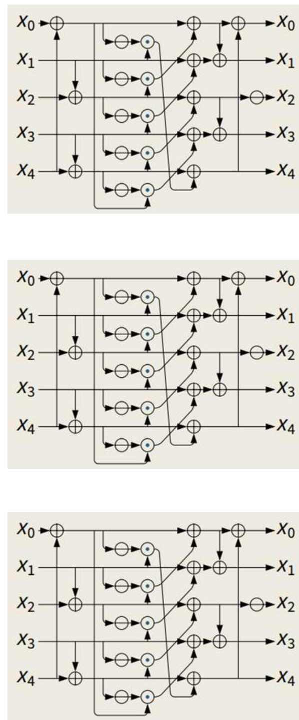
Table 10: Summary of attacks on ASCON.

Type	Rounds	Time	Method
Key Recovery	6/12	2^{66}	Cube-like
Key Recovery	5/12	2^{35}	Cube-like
Key Recovery	5/12	2^{36}	Differential-Linear
Key Recovery	5/12	2^{58} or $2^{127.99}$	Truncated/Improbable
Key Recovery	4/12	2^{18}	Differential-Linear
Key Recovery	4/12	3^{48}	Truncated/Impossible
Forgery	4/12	2^{101}	Differential
Forgery	3/12	2^{33}	Differential

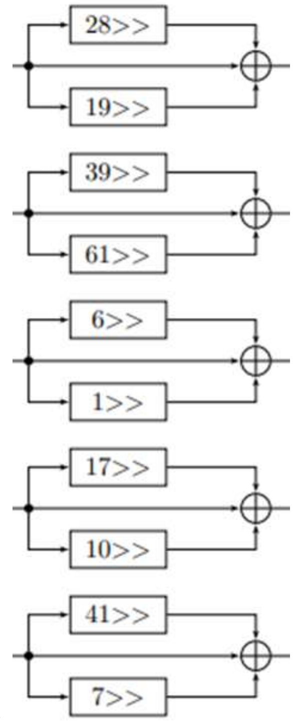
Are Many Tiny Boxes Toxic?

- 6,8,12 rounds like this:

“semi-transparent”



“totally-transparent”



Δ Δ Δ

Ascon

unrelated?

Are Many Tiny Boxes Toxic? YES, without any doubt

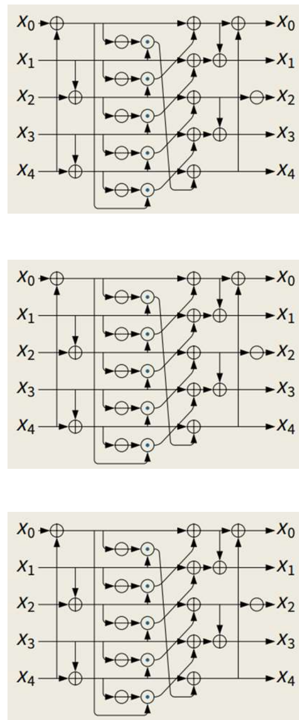
- 6,8,12 rounds like this:

This is a HUGE amount of shared information

- An UNDENIABLE information-theoretic property
- Active in essentially any of already known attacks on Ascon…

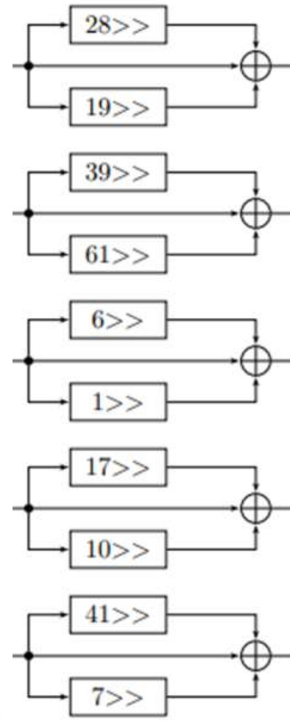
“semi-transparent”

Δ



Δ

“totally-transparent”



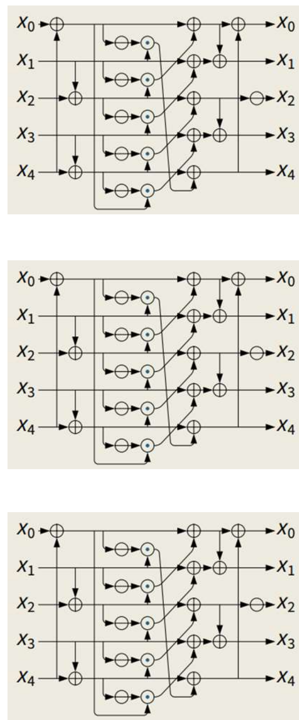
Δ

Ascon
64*1.91
122 bits

Are Many Tiny Boxes Toxic? YES, without any doubt

- 6,8,12 rounds like this:

Δ



Δ

Ascon

122 bits

would be only 42 bits total if
we used the AES S-box

Information-theoretical
undeniable
cannot be ignored

Agenda

How high MI implies

“Undesirable Properties”

or does it?

Academic Background:

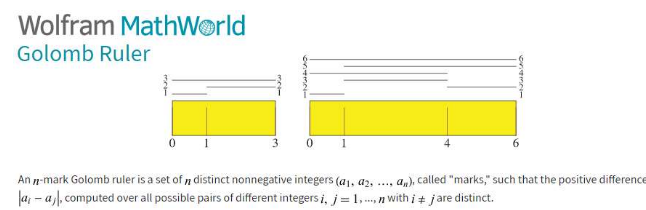
The notion of so called "Forbidden Mappings"

Def. We call Sidon-Rodier-Golomb = SRG_2 mappings all sets of 4 points which map an affine space of dim 2 to an affine space of dim 2
[⇔partial linearity on 4 points]

	APN AES-like	RP	Ascon/ Keccak
	1.125	1.57	bad
Forbidden Mappings =>	0	some exist	bad

Sidon Sequence

Number Theory, Paul Erdős, Pál Turán



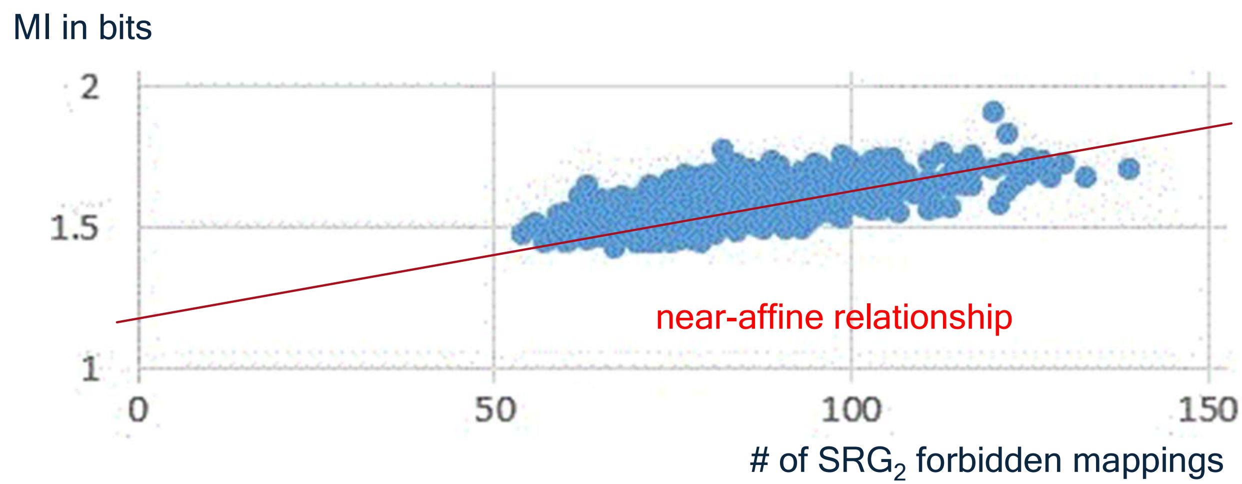
A divisibility criterion for exceptional APN functions

Florian Caullery

ABSTRACT. We are interested in the functions from $\mathbb{F}_{2^m}$ to itself which are Almost Perfectly Nonlinear over infinitely many extensions of $\mathbb{F}_2$, namely, the exceptional APN functions. In particular, we study the case of the polynomial functions of degree $4e$ with e odd and we give a necessary condition on an associated multivariate polynomial for the function to be exceptional APN. We use this condition to confirm the conjecture of Aubry, McGuire and Rodier

MI is Additive and influenced by ALL partial linearity events => Prediction Ability

The number of SRG_2 or "Forbidden Mappings" is predictable:



	APN AES-like, Fides	RP	Ascon/ Keccak
	1.125	1.57	1.90
Forbidden Mappings =>	0	>50	80

Other Ciphers?

All ciphers are the same!

Δ

MI

Δ

MI =

a PRECISE and RELIABLE

measure of quality of ciphers!

*note: the more repetition, like 8 or 10, the more we approach the concept of “space trails”

Example:
Compare 3 versions of DES on
MI and #SRG_2 mappings.

Table 9. Selected best 16 mappings of affine spaces U of dimension 2 which can be mapped to an affine space W of dimension 2, classified by input linear spaces, we report how many times $K = 0, \ldots 10$ they are re-used in distinct affine spaces.

1.71

$U1$	$U2$	DES S-box 1	2	3	4	5	6	7	8
1	4	1	3	–	–	1	–	7	–
1	8	1	2	5	4	3	5	5	2
1	C	1	1	–	–	2	–	5	3
1	E	5	1	1	4	1	1	–	–
2	4	4	7	3	4	–	4	4	6
2	8	–	2	3	4	–	1	–	–
3	4	1	3	2	4	1	1	2	–
3	5	–	–	4	–	1	2	2	1
3	8	1	1	1	4	1	1	1	1
3	C	3	–	2	4	1	2	2	–
3	D	–	2	5	4	1	3	1	3
4	A	4	3	1	–	2	1	1	1
5	A	3	2	–	–	2	–	2	1
5	B	2	4	3	4	2	4	2	4
6	B	3	1	1	4	3	–	1	2
7	B	3	2	–	4	2	2	3	2
total		255							

2.11

$U1$	$U2$	s^5DES S-box 1	2	3	4	5	6	7	8
1	2	2	4	**8**	6	2	4	2	2
1	4	4	2	4	4	6	6	2	4
1	8	5	4	5	3	4	3	6	2
2	4	6	4	6	6	**8**	4	4	4
2	5	4	**8**	**8**	**10**	2	**10**	4	**8**
2	8	5	2	2	1	5	3	2	3
3	4	4	4	6	2	4	4	6	2
3	5	2	2	6	2	6	**8**	2	6
3	8	4	3	6	3	3	1	4	4
4	8	4	4	1	3	5	3	2	3
5	8	5	3	2	4	3	4	4	6
5	A	2	3	4	5	1	2	3	6
7	8	4	7	4	3	4	7	**8**	2
7	9	1	4	5	–	3	4	3	3
7	A	–	7	6	3	1	6	2	4
7	B	3	4	5	–	2	5	5	1
total		505							

1.67 bits

$U1$	$U2$	S*DES S-box 1	2	3	4	5	6	7	8
1	2	–	–	–	–	3	2	1	2
1	4	1	3	–	2	–	–	1	–
1	6	1	1	2	3	2	1	–	2
1	8	3	3	3	3	4	4	–	3
2	4	6	5	5	3	3	4	**8**	6
2	5	–	1	1	–	1	2	1	3
2	C	2	–	2	3	1	2	1	2
3	5	1	2	2	1	2	1	–	3
3	8	1	3	3	1	–	–	2	–
3	D	1	2	3	4	2	2	3	1
4	8	3	1	–	–	3	1	–	2
4	A	3	2	–	1	1	3	–	–
4	B	1	–	3	4	2	–	–	2
5	A	1	1	1	3	4	3	2	3
5	B	1	2	4	4	4	2	3	2
7	B	3	2	4	–	–	2	–	1
total		236							

a precise near-linear relationship!

\# of SRG_2 forbidden mappings

Related Work

Tezcan 2014:

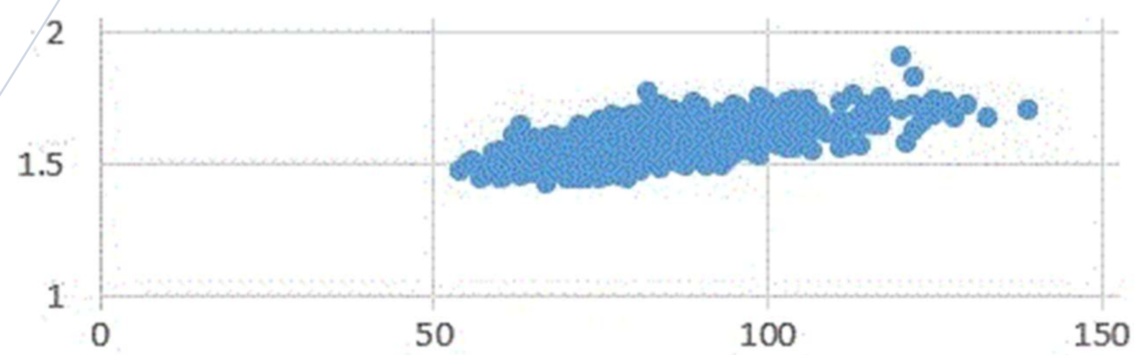
For a specific input difference of an S-box, some bits of the output difference remain invariant…

single input difference focus in prior work

2x input difference in one property is what we study here

Table 2: Undisturbed Bits of ASCON's S-box.

Input Difference	Output Difference	Input Difference	Output Difference
00001	?1???	10000	?10??
00010	1???1	10001	10??1
00011	???0?	10011	0???0
00100	??110	10100	0?1??
00101	1????	10101	????1
00110	????1	10110	1????
00111	0??1?	10111	????0
01000	??11?	11000	??1??
01011	???1?	11100	??0??
01100	??00?	11110	?1???
01110	?0???	11111	?0???
01111	?1?0?		



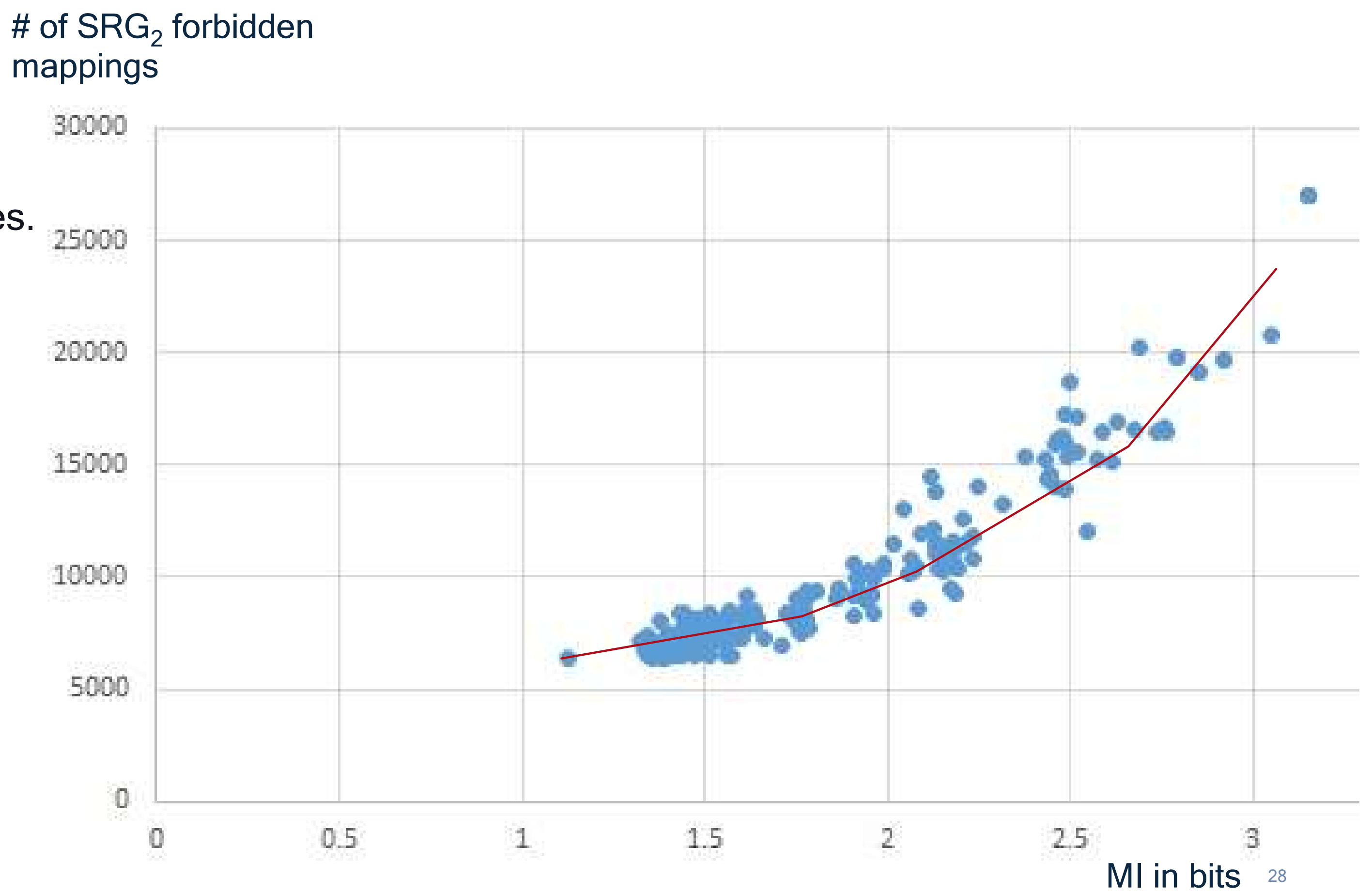
New! Combinatorial Explosion of Undesirable Properties – 2 S-boxes

Same SRG_2 mappings. 2 active boxes.

Output limited to

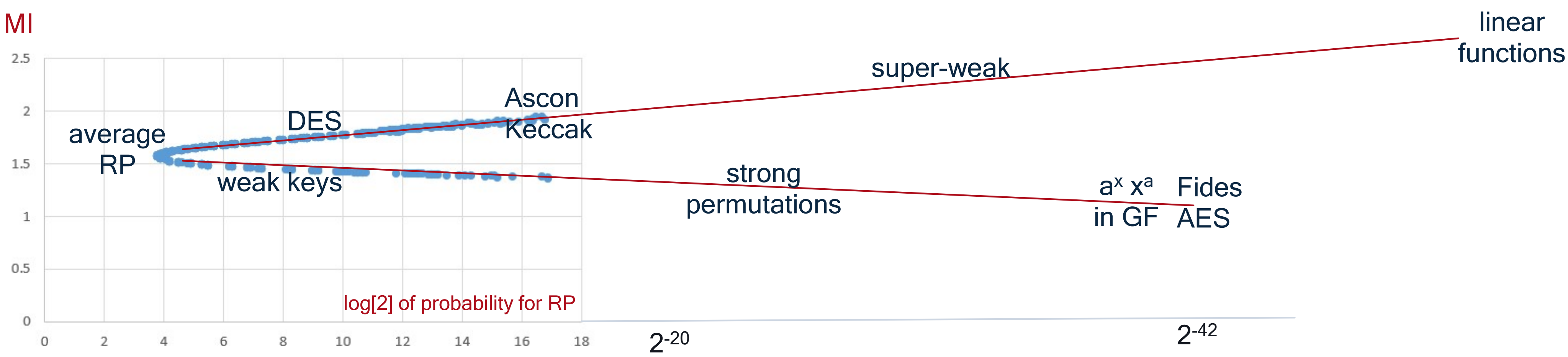
HW=6 active bits / 30 max.

=>the same type of prediction law is expected to hold for 3,4,5,… S-boxes and for most of the attacks we study: they are all based on the same basic events.



Random Permutations, APN and Ascon S-box

A new way of classifying S-boxes from strong to weak on a 2D scale.



CLAIM: we need to contemplate the large distance which separates the Ascon S-box (MI=1.91) and an ideal S-box not in terms of differential and linear properties (the distance seems small) but in terms of:

- How hard it is for a RP to move to this area – close to impossible!
- The combinatorial explosion of undesirable properties [previous slide].

Executive Summary:

Many cryptanalytics attacks can be MAPPED to combinations of discrete combinatorial events which are pure information theoretic events: they correspond to a certain probability mass of events inside a small finite number of small scale undesirable local linearity properties.

It is possible to see that Ascon is unnecessarily weak:

1. Many tiny S-boxes are somewhat inherently weak: like 42 => 122 bits of Mutual Information
2. Large MI => prediction capability=>combinatorial explosion in # undesirable properties (faster than linear).
3. We claim that the Ascon S-box is an unfortunate choice, and it never was optimized in the full light of how it leads to undesirable properties. Needs some more work.

Open problem: is there a NL layer for Ascon which simultaneously:

- has lower HW cost and low depth (possibly avoiding any XORs which are slow).
- is easy to protect against side channels and has a reversible Toffoli implementation
- has a much lower information theoretic security measure of MI for I/O differences.
- has zero or near zero undesirable mappings.

We are willing to work on any new Ascon update/proposal and to evaluate it against enclosed concerns.

Thank you

Qualcomm

Follow us on:

For more information, visit us at:

qualcomm.com & qualcomm.com/blog