Revisiting Higher-Order Differential-Linear Attacks from an Algebraic Perspective

Kai Hu, Thomas Peyrin, Quan Quan Tan, and Trevor Yap

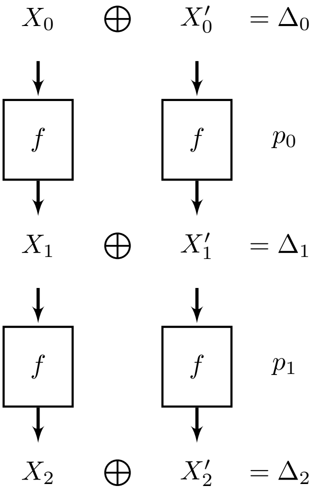
Nanyang Technological University

December 7, 2023
Guangzhou, China

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

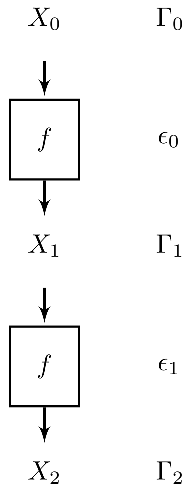
Contents

Differential Cryptanalysis



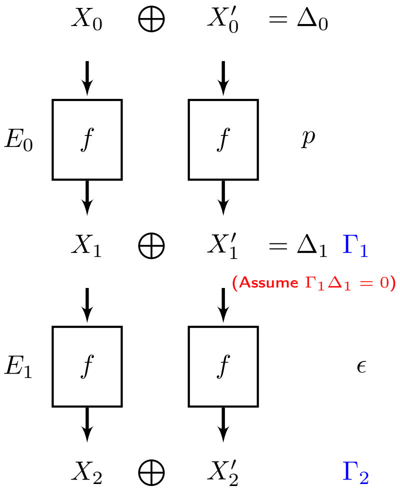
- Proposed by [BS,Crypto'91]
- Probability: $\Delta_0 \to \Delta_2$
- Traditionally studied using statistical method
  - Probability: $\Delta_0 \to \Delta_1$ with p_0
  - Probability: $\Delta_1 \to \Delta_2$ with p_1
  - Probability: $p = p_0 p_1$

Linear Cryptanalysis



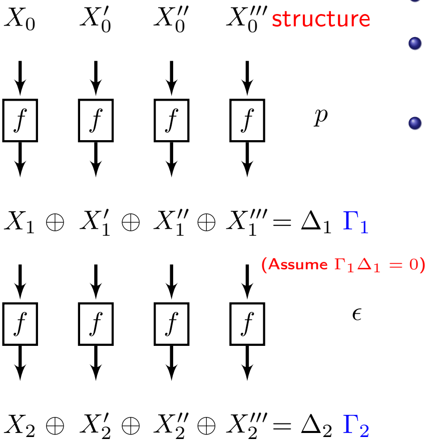
- Proposed by [Mat,Eurocypt'93]
- Correlation: $\Gamma_0 \rightarrow \Gamma_1$
- Traditionally studied using statistical method
  - Correlation: $\Gamma_0 \rightarrow \Gamma_1$ with ϵ_0
  - Correlation: $\Gamma_1 \rightarrow \Gamma_2$ with ϵ_1
  - Correlation: $\epsilon = \epsilon_0 \epsilon_1$

Differential-Linear Cryptanalysis



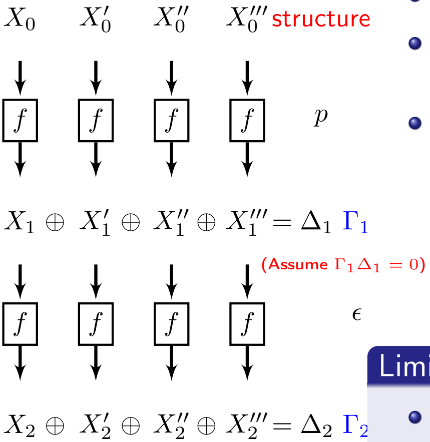
- Proposed by [LH,Crypto'94]
- Cor.: $\Gamma_2(X_2 \oplus X_2')$ w/ $X_0 \oplus X_0' = \Delta_0$
- Traditionally studied using statistical 2-phase method
  - $E = E_1 \circ E_0$
  - Probability: $\Delta_0 \to \Delta_1$ with p
  - Correlation: $\Gamma_1 \to \Gamma_2$ with ϵ
  - DL correlation: $p\epsilon^2$

Higher-Order Differential-Linear Cryptanalysis



- Proposed by [BDK,FSE'05]
- Cor.: $\Gamma_2 \left(\bigoplus X_2\right)$ with $X_0, X_0', \ldots$ being a HD structure
- Traditionally studied using statistical 2-phase method
  - $E = E_1 \circ E_0$
  - Probability of HD of E_0 is p
  - Correlation: $\Gamma_1 \to \Gamma_2$ with ϵ
  - Correlation of HDL: $p\epsilon^{2^d}$

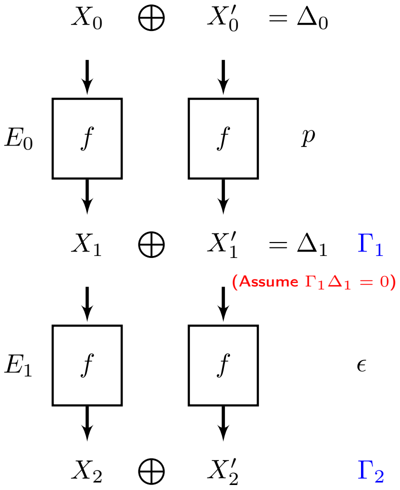
Higher-Order Differential-Linear Cryptanalysis



- Proposed by [BDK,FSE'05]
- Cor.: $\Gamma_2\left(\bigoplus X_2\right)$ with $X_0, X_0', \ldots$ being a HD structure
- Traditionally studied using statistical 2-phase method
  - $E = E_1 \circ E_0$
  - Probability of HD of E_0 is p
  - Correlation: $\Gamma_1 \to \Gamma_2$ with ϵ
  - Correlation of HDL: $p\epsilon^{2^d}$

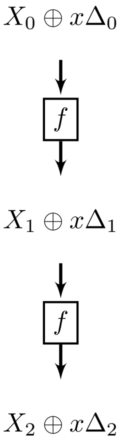
Limitations

- No method for a probabilistic HD
- If $\epsilon < 1$, HDL correlation goes to zero



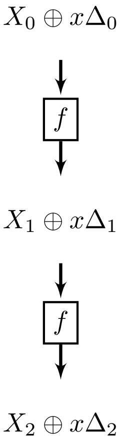
- Proposed by [LLL,Crypto'21]
- An algebraic perspective
  - DL cor. $=$ cor. of $\Gamma_2(X_2 \oplus X_2')$

Algebraic Transitional Form



- Proposed by [LLL,Crypto'21]
- An algebraic perspective
  - DL cor. $=$ cor. of $\Gamma_2(X_2 \oplus X_2')$

Algebraic Transitional Form



- Proposed by [LLL,Crypto'21]
- An algebraic perspective
  - DL cor. $=$ cor. of $\Gamma_2(X_2 \oplus X_2')$
- The form of output difference can be derived from a recursive method
  - X_1, Δ_1 are functions of X_0
  - $\Gamma_2(X_2 \oplus X_2')$ is a function of X_1, Δ_1

Contents

Contributions

- HATF: to generalize the ATF to the higher-order case
  - HATF can predict the probabilistic bias of a HDL approximation
  - New distinguishers/key-recovery attacks on Ascon and Xoodyak
- DSF: to linearize Ascon permutation
  - Improved zero-sum distinguishers for Ascon permutations

Contents

HD of a Boolean function [Lai, 1994]

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ and an ℓ^{th}-order input difference $\mathbf{\Delta} = (\Delta_0, \ldots, \Delta_{\ell-1})$ for a certain input $X \in \mathbb{F}_2^n$. The ℓ^{th} derivative of f is calculated as

$$\mathcal{D}_{\mathbf{\Delta}} f(X) = \bigoplus_{a \in X \oplus \mathsf{span}(\mathbf{\Delta})} f(a)$$

HDL Cryptanalysis from an Algebraic Perspective

HD of a Boolean function [Lai, 1994]

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ and an ℓ^{th}-order input difference $\boldsymbol{\Delta} = (\Delta_0, \ldots, \Delta_{\ell-1})$ for a certain input $X \in \mathbb{F}_2^n$. The ℓ^{th} derivative of f is calculated as

$$\mathcal{D}_{\boldsymbol{\Delta}} f(X) = \bigoplus_{a \in X \oplus \mathsf{span}(\boldsymbol{\Delta})} f(a)$$

Prop. (Algebraic Perspective on HD/HDL)

Let

$$\mathcal{M} : \mathbb{F}_2^\ell \to X \oplus \mathsf{span}(\boldsymbol{\Delta})$$
$$(x_0, x_1, \ldots, x_{\ell-1}) \mapsto X \oplus x_0\Delta_0 \oplus \cdots \oplus x_{\ell-1}\Delta_{\ell-1} \triangleq X \oplus \boldsymbol{x}\boldsymbol{\Delta}$$

We have

$$\mathcal{D}_{\boldsymbol{\Delta}} f(X) = \bigoplus_{\boldsymbol{x} \in \mathbb{F}_2^\ell} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta}) = D_{\boldsymbol{x}} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta})$$

HDL Cryptanalysis from an Algebraic Perspective

HD of a Boolean function [Lai, 1994]

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ and an ℓ^{th}-order input difference $\boldsymbol{\Delta} = (\Delta_0, \ldots, \Delta_{\ell-1})$ for a certain input $X \in \mathbb{F}_2^n$. The ℓ^{th} derivative of f is calculated as

$$\mathcal{D}_{\boldsymbol{\Delta}} f(X) = \bigoplus_{a \in X \oplus \mathsf{span}(\boldsymbol{\Delta})} f(a)$$

Prop. (Algebraic Perspective on HD/HDL)

Let

$$\mathcal{M} : \mathbb{F}_2^\ell \to X \oplus \mathsf{span}(\boldsymbol{\Delta})$$
$$(x_0, x_1, \ldots, x_{\ell-1}) \mapsto X \oplus x_0\Delta_0 \oplus \cdots \oplus x_{\ell-1}\Delta_{\ell-1} \triangleq X \oplus \boldsymbol{x}\boldsymbol{\Delta}$$

We have

$$\mathcal{D}_{\boldsymbol{\Delta}} f(X) = \bigoplus_{\boldsymbol{x} \in \mathbb{F}_2^\ell} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta}) = D_{\boldsymbol{x}} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta})$$

Expression of HD: Coe $(f(X \oplus \boldsymbol{x}\boldsymbol{\Delta}), \boldsymbol{x})$

Higher-Order Algebraic Transitional Form

Iterative Cipher

Ciphers are iterative composed of simple round functions

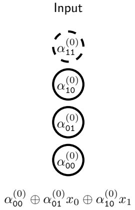
$$E = E_{R-1} \circ E_{R-2} \circ \cdots E_1 \circ E_0, \quad E_r : \mathbb{F}_2^n \to \mathbb{F}_2^n$$

We can construct the expression of HD in an iterative method

Construction of Higher-Order Algebraic Transitional Form

Write $X \oplus \boldsymbol{x}\boldsymbol{\Delta}$ as $\bigoplus_{u \in \mathbb{F}_2^\ell} \alpha_u \boldsymbol{x}^u$:

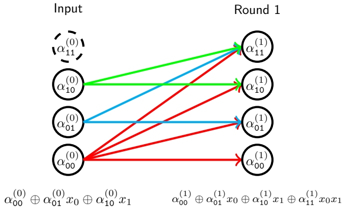
$$\alpha_u = \begin{cases} X, & u = 0 \\ \Delta_i, & u = e_i \\ \mathbf{0}, & \text{otherwise} \end{cases}$$



Construction of Higher-Order Algebraic Transitional Form

Apply E_r to $\bigoplus_{u \in \mathbb{F}_2^\ell} \alpha_u^{(r)} \boldsymbol{x}^u$

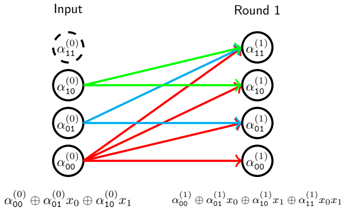
$$\bigoplus_{u \in \mathbb{F}_2^\ell} \alpha_u^{(r+1)} \boldsymbol{x}^u = E_r \left(\bigoplus_{u \in \mathbb{F}_2^\ell} \alpha_u^{(r)} \boldsymbol{x}^u \right)$$



Construction of Higher-Order Algebraic Transitional Form

$\alpha_u^{(r+1)}$ is a function of $\alpha_u^{(r)}$

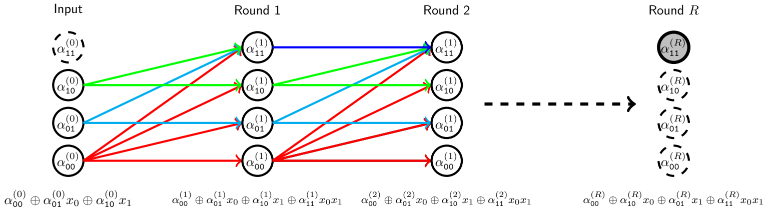
$$\alpha_u^{(r+1)} = \mathsf{Coe}\left(E_r\left(\bigoplus_{u\in\mathbb{F}_2^\ell}\alpha_u^{(r)}\boldsymbol{x}^u\right), \boldsymbol{x}^u\right)$$



Construction of Higher-Order Algebraic Transitional Form

Connecting all round functions, we obtain HATF of E,

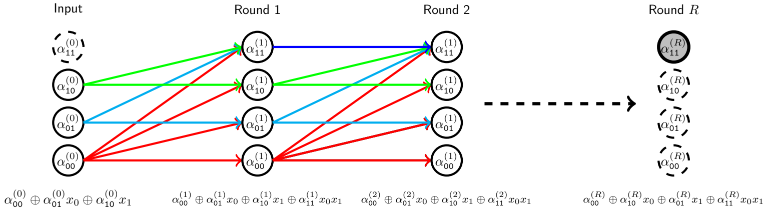
$$\mathcal{E} = \mathcal{E}_{R-1} \circ \mathcal{E}_{R-2} \circ \cdots \circ \mathcal{E}_0, \quad \mathcal{E}_r : (\mathbb{F}_2^n)^{2^\ell} \to (\mathbb{F}_2^n)^{2^\ell}$$



Construction of Higher-Order Algebraic Transitional Form

Time complexity of constructing the HATF:

- Dominated by the calculations of ANFs round by round
- Most time-consuming step is to calculate the d-degree monomials for $\bigoplus_{u\in\mathbb{F}_2^\ell} \alpha_u^{(r)} \boldsymbol{x}^u$
- $2^{d\ell}$ multiplications/additions
- Final time complexity: $\mathcal{O}(2^{d\ell})$ (detailed analysis can be found in the paper)



Computing the Bias of HDL

$\alpha_{\mathbf{1}}^{(R)}$ is a composite form:

$$\left(\alpha_u^{(0)}, u \in \mathbb{F}_2^n\right) \xrightarrow{\mathcal{E}_0} \cdots \xrightarrow{\mathcal{E}_{R-2}} \left(\alpha_u^{(R-1)}, u \in \mathbb{F}_2^n\right) \xrightarrow{\mathcal{E}_{R-1}} \alpha_{\mathbf{1}}^{(R)}$$

Lemma (LLL, Crypto'21)

Assume the bias of $x_0, x_1, \ldots, x_{n-1}$ are $\epsilon_0, \epsilon_1, \ldots, \epsilon_{n-1}$, respectively.

$$\text{Bias}(f) = \sum_{\substack{x_0, x_1, \ldots, x_{n-1} \\ s.t. f(x_0, \ldots, x_{n-1})=0}} \prod_{i=0}^{n-1} \left(\frac{1}{2} + (-1)^{x_i} \varepsilon_i\right) - \frac{1}{2}$$

- Time complexity is exponential in the number of variables in the ANF
- The number of variables is at most $d \times 2^\ell$
- Final time complexity: $\mathcal{O}(2^{\ell + d \times 2^\ell})$ (detailed analysis can be found in the paper)

Reduce the Complexity for Primitives with Quadratic Round Functions

- Primitives with quadratic round functions are more and more popular
- Higher-order differential related attacks are one of the main threats

Quadratic Boolean function can be transformed into a disjoint form [JA, 1977]

$$f = x_0x_1 + x_2x_3 \quad (\checkmark)$$
$$f = x_0x_1 + x_0x_2 \quad (\times)$$

Reduce the Complexity for Primitives with Quadratic Round Functions

- Primitives with quadratic round functions are more and more popular
- Higher-order differential related attacks are one of the main threats

Quadratic Boolean function can be transformed into a disjoint form [JA, 1977]

$$f = x_0x_1 + x_0x_2(\times) \rightarrow f = x_0(x_1 + x_2) \xrightarrow{\text{Sub}} f = t_0t_1(\checkmark)$$

Reduce the Complexity for Primitives with Quadratic Round Functions

A quicker method

- Apply a linear substitution to all the variables to make f be disjoint

$$f = g \circ M(x_0, x_1, \ldots, x_{n-1})$$

- Compute the correlation of new variables by Piling-up lemma

$$y = x_0 \oplus x_1 \oplus x_2 \oplus \cdots$$

- Compute the correlation of each individual part

$$g = x_0 x_1 + x_0 + x_1 + 1$$

- Compute the correlation of f

$$f = g_0 \oplus g_1 \oplus g_2 \cdots$$

Reduce the Complexity for Primitives with Quadratic Round Functions

A quicker method

- The variable substitution is the most time-consuming: $\mathcal{O}(n^{3.8})$
  (n is the number of variables)
- The number of variables in an ANF is $2 \times 2^{\ell}$
- Final time complexity: $\mathcal{O}(2^{3.8\ell})$

Assumption Made for the Method

Assumption

- The construction of HATF does not require assumptions
- The calculation of bias of variables requires the variables to be independent

Trouble and Solution

- If a variable is linear, then it is more risky not to be independent

$$\alpha_{\boldsymbol{u}}^{(r+1)}[i] = \alpha_{\boldsymbol{u}}^{(r)}[i_0] \oplus \alpha_{\boldsymbol{u}}^{(r)}[i_1] \oplus \cdots$$

Assumption Made for the Method

Assumption

- The construction of HATF does not require assumptions
- The calculation of bias of variables requires the variables to be independent

Trouble and Solution

- If a variable is linear, then it is more risky not to be independent

Not introduce new variables

Assumption Made for the Method

Assumption

- The construction of HATF does not require assumptions
- The calculation of bias of variables requires the variables to be independent

Trouble and Solution

- Different bits of $\alpha_{\boldsymbol{u}}^{(r)}$ can be highly related

$$\alpha_{\boldsymbol{u}}^{(r)}[i] = \alpha_{\boldsymbol{u}}^{(r)}[j] \text{ or } \alpha_{\boldsymbol{u}}^{(r)}[i] = \alpha_{\boldsymbol{u}}^{(r)}[j] + 1$$

Assumption Made for the Method

Assumption

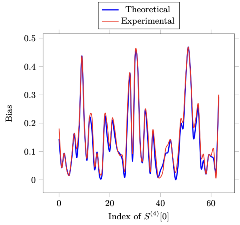
- The construction of HATF does not require assumptions
- The calculation of bias of variables requires the variables to be independent

Trouble and Solution

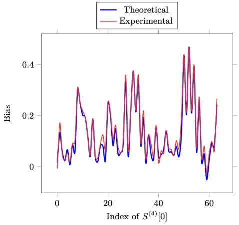
- Different bits of $\alpha_{\boldsymbol{u}}^{(r)}$ can be highly related

$$\alpha_{\boldsymbol{u}}^{(r)}[i] \text{ can be represented by } \alpha_{\boldsymbol{u}}^{(r)}[j]$$

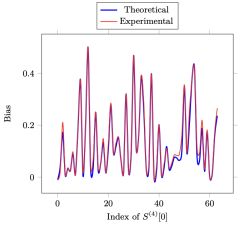
Some curves for 2nd order HDL of 4-round Ascon initialization



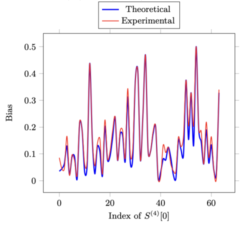
(a) $\Delta(0,1)$



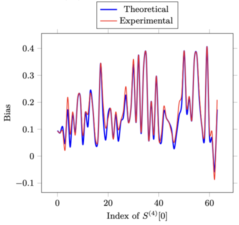
(b) $\Delta(0,2)$



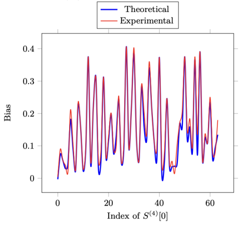
(c) $\Delta(0,3)$



(a) $\Delta(0,4)$



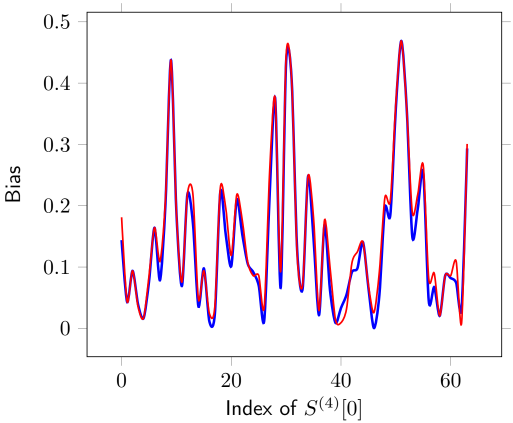
(b) $\Delta(0,5)$



(c) $\Delta(0,6)$

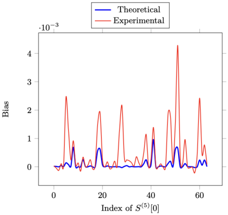
Precision of HATF

Curve of one 2^{th}-order HDL for 4-round Ascon initialization

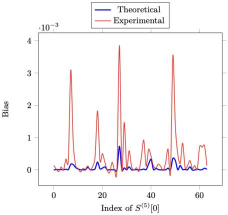


Precision of HATF

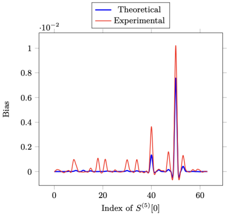
Some curves for 2nd order HDL of 5-round Ascon initialization:



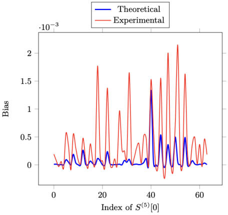
(a) $\Delta(0,1)$



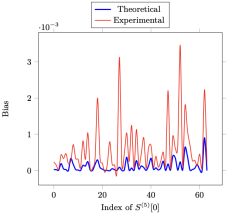
(b) $\Delta(0,2)$



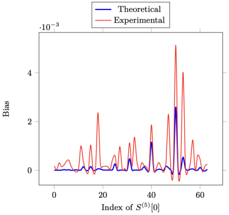
(c) $\Delta(0,3)$



(a) $\Delta(0,4)$



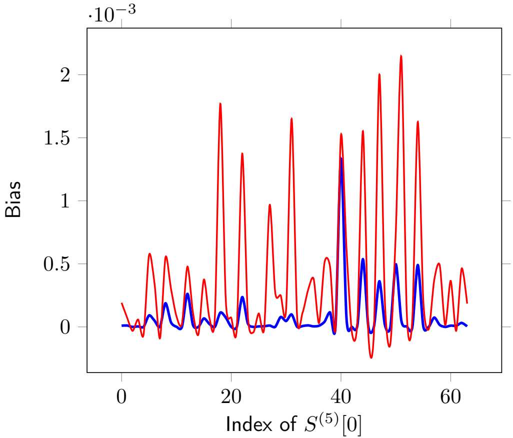
(b) $\Delta(0,5)$



(c) $\Delta(0,6)$

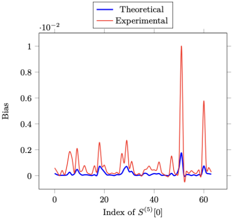
Precision of HATF

Curve of one 2^{nd}-order HDL for 5-round Ascon initialization

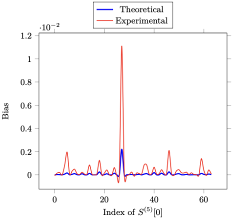


Precision of HATF

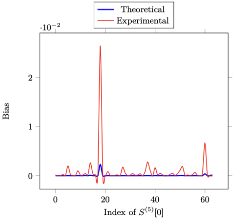
Some curves for 3rd to 8th order HDL of 5-round Ascon initialization:



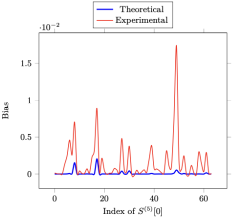
(a) $\Delta(0, 24, 33)$



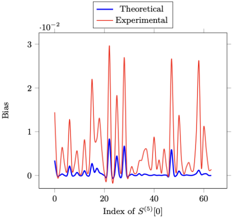
(b) $\Delta(0, 9, 15, 41)$



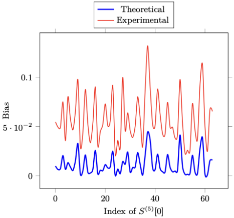
(c) $\Delta(0, 9, 24, 51, 55)$



(a) $\Delta(1, 12, 18, 22, 21, 52)$



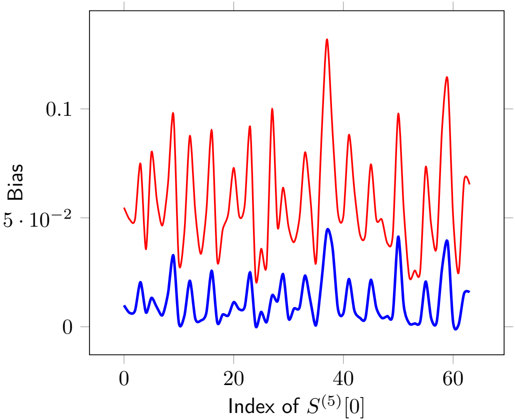
(b) $\Delta(10, 13, 21, 31, 49, 55, 61)$



(c) $\Delta(0, 8, 9, 13, 14, 26, 43, 60)$

Precision of HATF

Curve of one 8^{th}-order HDL for 5-round Ascon initialization



Precision of HATF

Discussion on Precision

- HATF CANNOT provide any upper/lower bound for HDL biases
- Quite precise to predict biased bits
- When the reported bias is high, the real bias is also high
- We have not observed any counterexamples during our experiments

Results

Results for Ascon initialization

Primitive	Round	Order	Bias Expr.	Bias Theory	Method	Reference
Ascon Init.	4	1^{st}	2^{-2}	2^{-20}	Classical	[DEMS, CT-RSA'15]
				2^{-5}	DLCT	[BDKW, Eurocrypt'19]
				$2^{-2.365}$	ATF	[LLL, Crypto'21]
				$\mathbf{2^{-2.09}}$	HATF	Here
		2^{nd}	2^{-1}	$\mathbf{2^{-1}}$	HATF	Here
	5	1^{st}	2^{-9}	–	Experimental	[DEMS, CT-RSA'15]
				$\mathbf{2^{-10}}$	HATF	Here
		2^{nd}	$2^{-6.60}$	$\mathbf{2^{-7.05}}$	HATF	Here
		8^{th}	$2^{-3.35}$	$\mathbf{2^{-4.73}}$	HATF	Here
	6	3^{rd}	$2^{-22}\dagger$	$\mathbf{2^{-25.97}}\dagger$	HATF	Here

† This bias holds when 24 conditions are satisfied

Results

Primitive	Round	Order	Bias		Method	Reference
			Expr.	Theory		
Xoodyak Init.	4	1^{st}	$2^{-9.7}$	–	Experimental	[DW, SAC'22]
				$\mathbf{2^{-9.67}}$	HATF	Here
			$-2^{-5.36}$	–	Experimental	[DW, SAC'22]
				$\mathbf{-2^{-6.0}}$	HATF	Here
		2^{nd}	$2^{-5.72}$	$\mathbf{2^{-5.72}}$	HATF	Here
		4^{th}	2^{-1}	$\mathbf{2^{-1}}$	HATF	Here
	5	2^{nd}	–	$\mathbf{2^{-45}}$	HATF	Here
Xoodoo	4	-	2^{-1}	2^{-1}	Rot. DL	[LSL, Eurocrypt'21]
		4^{th}	2^{-1}	$\mathbf{2^{-1}}$	HATF	Here
	5	3^{rd}	$2^{-8.79}$	$\mathbf{2^{-8.96}}$	HATF	Here

Contents

Differential Supporting Function

We know:

$$\mathcal{D}_{\mathbf{\Delta}} f(X) = \bigoplus_{\boldsymbol{x} \in \mathbb{F}_2^{\ell}} f(X \oplus \boldsymbol{x}\mathbf{\Delta}) = D_{\boldsymbol{x}} f(X \oplus \boldsymbol{x}\mathbf{\Delta})$$

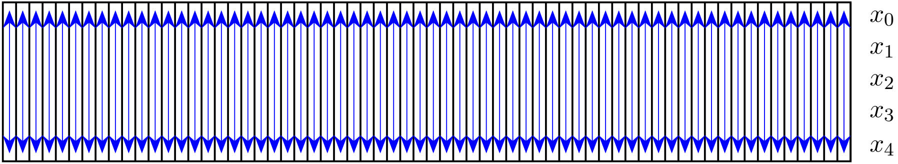
Differential Supporting Function

We know:

$$\mathcal{D}_{\boldsymbol{\Delta}} f(X) = \bigoplus_{\boldsymbol{x} \in \mathbb{F}_2^{\ell}} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta}) = D_{\boldsymbol{x}} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta})$$

- X and $\boldsymbol{\Delta}$ are parameters
- With X and $\boldsymbol{\Delta}$ being properly chosen, $D_{\boldsymbol{x}} f(X \oplus \boldsymbol{x}\boldsymbol{\Delta})$ can be made simpler

DSF on Ascon Permutation



- Intuition: Let all Sboxes have the same $\bar{X} + x\bar{\Delta}$
- $32 \times 31 = 992$ choices
- Evaluate the algebraic degree of r-round Ascon with $X = \bar{X}^{64}$, $\mathbf{\Delta} = \bar{\Delta}^{64}$

DSF on Ascon Permutation

$$(\bar{X}, \bar{\Delta}) \in \left\{ \begin{array}{l} (\texttt{0x6}, \texttt{0x13}), (\texttt{0xa}, \texttt{0x13}), (\texttt{0xc}, \texttt{0x17}), (\texttt{0xf}, \texttt{0x18}), \\ (\texttt{0x15}, \texttt{0x13}), (\texttt{0x17}, \texttt{0x18}), (\texttt{0x19}, \texttt{0x13}), (\texttt{0x1b}, \texttt{0x17}) \end{array} \right\}$$

Round r	**Upper bounds on the algebraic degree**				
	$S^{(r)}[0]$	$S^{(r)}[1]$	$S^{(r)}[2]$	$S^{(r)}[3]$	$S^{(r)}[4]$
4	3	3	2	2	3
5	6	5	5	6	6
6	11	11	12	12	11
7	23	24	23	23	22
8	47	47	45	46	47

Improved Zero-Sum Results for Ascon Permutation

New zero-sum distinguishers on Ascon permutation:

Type	Rnd	Data(log)	Time (log)	Method	Reference
From Start	8	130	130	Integral	[Todo, Eurocrypt'15]
		48	**48**	HD	Here
Best	11	315	315	Integral	[Todo, Eurocrypt'15]
Inside-outside	12	130	130	Zero-Sum	[Todo, Eurocrypt'15]
		55	**55**	Zero-Sum	Here

Discussion on the new zero-sum distinguishers

- The inputs (outputs) are fixed, so they are different from/weaker than the previous zero-sum distinguishers (derived from division property)
- More information is captured

Contents

Conclusion

- A generalization of the algebraic perspective on DL to HDL cases
- The first theoretical method for a probabilistic HDL distinguisher: HATF
- Improved distinguishers/key-recovery attacks for some round-reduced Ascon and Xoodyak
- A systematic method for linearization and finding zero-sum distinguishers for Ascon: DSF

Conclusion

- A generalization of the algebraic perspective on DL to HDL cases
- The first theoretical method for a probabilistic HDL distinguisher: HATF
- Improved distinguishers/key-recovery attacks for some round-reduced Ascon and Xoodyak
- A systematic method for linearization and finding zero-sum distinguishers for Ascon: DSF

Thank You!