

Analysis of the NORX Core Permutation

Alex Biryukov, Aleksei Udovenko, Vesselin Velichkov

`first-name.last-name@uni.lu`
SnT, University of Luxembourg

Abstract. NORX is one of the fifteen authenticated encryption algorithms that have reached the third round of the CAESAR competition. NORX is built using the sponge-based Monkey Duplex construction. In this note we analyze the core permutation F . We show that it has rotational symmetries on different structure levels. This yields simple distinguishing properties for the permutation, which propagate with very high probability or even probability one. We also investigate differential symmetries in NORX at the word level. A new type of truncated differentials called symmetric truncated differentials (STD) is proposed. It is shown that, under the Markov assumption, up to 2.125 rounds of the F function of NORX32 and NORX64 can be distinguished using STD. Finally, we note that our analysis covers only the permutation F and does not immediately threaten the security claims of the designers.

Keywords: NORX, CAESAR, authenticated encryption, sponge, cryptanalysis

1 Introduction

CAESAR is a competition of authenticated ciphers aiming to select a portfolio of ciphers suitable for different usage scenarios. NORX [1] is one of the fifteen candidates that have reached the third round. NORX is based on the Monkey Duplex [2,3] construction which is a sponge mode tailored for authenticated encryption schemes.

In this paper we report on some new non-random properties of the NORX permutation. More specifically, we show that it exhibits some rotational symmetries on different structure levels. The latter yields simple distinguishing properties for the permutation, which propagate with very high probability or even probability one.

We also investigate differential symmetries in NORX at the word level. A new type of truncated differentials called symmetric truncated differentials (STD) is proposed. It is shown that, under the Markov assumption, up to 2.125 rounds of the F function of NORX32 and NORX64 can be distinguished using STD.

The rest of the paper is organized as follows. We begin by briefly outlining the NORX algorithm in Sect. 2. In Section 3 we describe rotational symmetric properties in its core permutation, both at the state and at the word level. This is followed by an analysis of NORX with respect to symmetric truncated differentials in Sect. 4. Section 5 concludes the paper.

2 Description of NORX

NORX has a sponge structure and is based on the monkeyDuplex construction. It uses ARX (Addition/Rotation/XOR) primitives, with exception of modular addition. More specifically, it is inspired by the ChaCha stream cipher, where the addition operation is replaced by the 1-st order approximation: $x \oplus y \oplus (x \wedge y) \lll 1$.

The original submission proposes versions of NORX with 32- and 64-bit words called resp. NORX32 and NORX64. Subsequently two more versions were proposed with 8- and 16-bit words called resp. NORX8 and NORX16 [4]. The word size is denoted by w . The internal state of all NORX variants is composed of 16 words organized as a 4×4 matrix.

The basic building block of NORX is a permutation F on $b = r + c$ bits where b is called the *width*, r is the *rate* and c is the *capacity*. F is also called a *round*, and F^l is an l -fold iteration of F . The recommended instances of NORX use $l = 4$ or $l = 6$ rounds. The initialization phase is always followed by a data processing phase and as a result the state effectively goes through F^{2l} before any absorption. NORX allows parallelization but we consider only the sequential construction (the parameter $p = 1$). The parameter combinations of the NORX variants are given in Table 1. A description of the full scheme is shown on Fig. 1.

Table 1: Parameters of the NORX variants.

word size (w)	rounds (l)	rate (r)	capacity (c)	state size (b)	nonce	key	tag (t)
8	4 or 6	40	88	128	32	80	80
16	4 or 6	128	128	256	32	96	96
32	4 or 6	768	256	1024	128	128	128
64	4 or 6	1536	512	2048	256	256	256

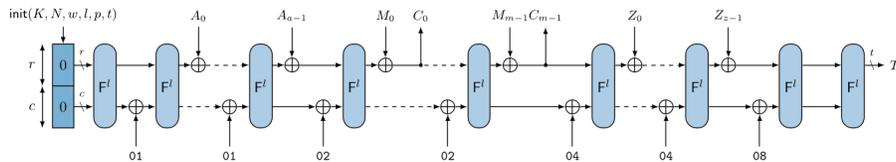


Fig. 1: The NORX AE scheme with parallelization parameter $p = 1$. K and N denote a key and a nonce resp., A and Z denote a header and a trailer resp., M_i and C_i denote plaintext and ciphertext blocks resp., T is the authentication tag. (credits: NORX specification [1])

F is composed of 8 steps: 4 column steps denoted by F_{col} , followed by 4 diagonal steps denoted by F_{diag} . 1, 2, ..., 7 and 8 steps are denoted resp. by $F^{0.125}$, $F^{0.250}$, ..., $F^{0.875}$ and $F^{1.000} = F = F_{diag} \circ F_{col}$. The first 4 steps of F represent an ARX circuit called G applied to the 4 columns of the state in parallel. The next 4 steps represent the G circuit applied to the 4 diagonals of the state in parallel. The G circuit of NORX is depicted on Fig. 2 and its application to the columns and diagonals is illustrated on Fig. 3.

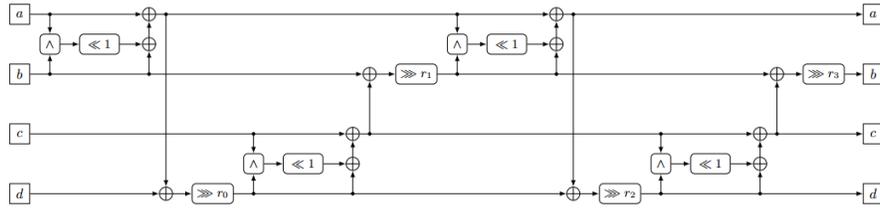


Fig. 2: The G circuit of NORX. It is applied in parallel first to each of the 4 columns of the state followed by an application to each of the 4 diagonals (see Fig. 3). This constitutes one round of F . (credits: NORX specification [1])

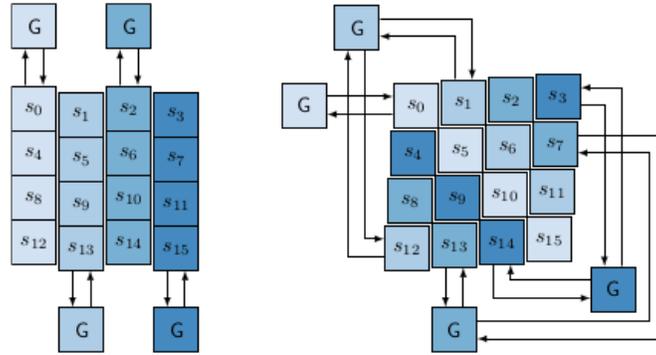


Fig. 3: The G circuit applied to the columns (left) and diagonals (right) of the state. (credits: NORX specification [1])

The security of each of the four versions of NORX is limited by the size of the key and tag. The designers require unique nonces and abort on verification failure. In addition, at most 2^e messages are allowed to be processed with a single key, where e is equal to 24, 32, 64, 128 resp. for NORX8, NORX16, NORX32, NORX64.

For a more detailed description of NORX we refer the reader to the specification [1].

3 Rotational Symmetries

In this section we describe rotational symmetries in the permutation F of NORX. They exist both on the word level and on the state level.

3.1 State Symmetries

We can see a 4x4 NORX state S as a list of 4 columns: $S = (c_0, c_1, c_2, c_3)$. The following proposition shows that the permutation F is column rotation-symmetric.

Proposition 1. *Let c_0, c_1, c_2, c_3 be four arbitrary 4-word columns and denote by R_n the function rotating of the columns left by n positions, for example $R_1(c_0, c_1, c_2, c_3) = (c_1, c_2, c_3, c_0)$. Then R_n can be moved freely through any number of rounds of the permutation F , that is, $F^l \circ R_n = R_n \circ F^l$ for any positive integer l .*

Proof. Clearly, rotation of columns do not affect the column step F_{col} , since it transforms each column separately: $F_{col} \circ R_n = R_n \circ F_{col}$. Such rotations do not break the diagonals as well, the diagonals are simply reordered by rotation with the same amount. Therefore, $F_{diag} \circ R_n = R_n \circ F_{diag}$. By applying the two equations l times, we obtain the required result.

Consider any state S that is column rotation-invariant with respect to R_n , i.e. $R_n(S) = S$ for a fixed integer $n, 1 \leq n \leq 3$. By the proposition, the state $F(S)$ is also column rotation-invariant with respect to R_n : $R_n(F(S)) = F(R_n(S)) = F(S)$. There are two cases:

1. $n = 1$ or $n = 3$. It means that $(c_0, c_1, c_2, c_3) = (c_1, c_2, c_3, c_0)$ and it follows that all columns are equal: $c_0 = c_1 = c_2 = c_3$. There are 2^{4w} out of 2^{16w} such states. The designers of NORX noted these states in [5].
2. $n = 2$. It means that $(c_0, c_1, c_2, c_3) = (c_2, c_3, c_0, c_1)$ and it follows that the two pairs of columns are equal: $c_0 = c_2$ and $c_1 = c_3$. There are 2^{8w} out of 2^{16w} such states.

The first case is quite obvious: if all columns are equal, then they will stay equal after any number of rounds of F . The second case shows that if the left and the right halves of the state are equal, then they will be equal after any number of rounds of F as well.

Hitting such a special state is not easy under the NORX's security claims. However, 2^{8w} is a more serious fraction of states than the 2^{4w} weak states which were known to the designers. To illustrate possible dangers of such properties, we describe a hypothetical weak-key set in NORX8 [4], an 8-bit NORX version for low-end devices. We remark though that NORX8 is not a part of the CAESAR submission.

A hypothetical attack on NORX8. The initial state of NORX8 is shown in Equation 1. n_i and k_i denote bytes of nonce and key respectively, u_i are constants and w, l, p, t are constants encoding parameters of NORX. It is possible to construct valid initial states with two equal halves, i.e. a special state from case $n = 2$ described before. Indeed, let us fix the four key bytes (k_2, k_3, k_6, k_7) and let us choose the two nonce bytes (n_2, n_3) arbitrarily. Then we can set the left half of the state equal to the right half, i.e. $(n_0, n_1) = (n_2, n_3)$, $(k_0, k_1) = (k_2, k_3)$ and so on. There are 2^{32} keys for which this construction works. The column rotation-invariant of such state is preserved through arbitrary number of rounds of F . After the first F^l rounds the domain separation constant will be XORed to the last word of the state. This constant is not symmetric and therefore it will break the property. Therefore, we consider a slightly modified version of NORX8 where the domain separation constant is symmetric (for example, if original constant is XORed not only to the last word, but to all words of the state or to all words in the last row). In such case the invariant is be preserved through the next F^l rounds and the rate part of the state is then observed by an adversary. This leads to an obvious distinguisher: the adversary simply compares the left and right halves of the exposed part of the state. In NORX8 the rate part consists of only 5 bytes. It allows to check only the topmost 4 words with error probability 2^{-16} . By using a few more encryptions (with another symmetric nonces) the error probability can be decreased to negligible.

$$\begin{pmatrix} n_0 & n_1 & n_2 & n_3 \\ k_0 & k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 & k_7 \\ k_8 \oplus w & k_9 \oplus l & u_{14} \oplus p & u_{15} \oplus t \end{pmatrix} \quad (1)$$

We remark that this attack is quite weak and requires symmetric domain separation constants. On the other hand, it applies independently of the number of rounds.

Cycles of F and Nonlinear Invariants. States consisting of four equal columns under application of the F_{col} or F_{diag} functions form cycles that correspond directly to cycles of the G function. Indeed, such states always consist of four copies of a single column and application of F_{col} or F_{diag} to a state is equivalent to applying G to the corresponding column. For instance, it is possible to enumerate all cycles of the G function for NORX8, where G permutes 2^{32} elements. All these cycles of G can be transformed into cycles of F_{col} or F_{diag} by simply making 4 copies of the column. Since in this case $F_{col} = F_{diag} = F^{0.5}$, these cycles will work for F as well, except that all even cycles will split into two cycles each. We provide the cycle decomposition of G from NORX8 in Appendix A.1.

Recently, a nonlinear invariant attack was introduced by Todo *et al.* in [6]. They show that, for any SPN-based cipher, if there exists a quadratic invariant for the S-Box and the binary matrix of the linear layer is orthogonal, then it is possible to construct a nonlinear invariant for the full round of the cipher.

Moreover, they show that all nonlinear invariants of the S-Box can be obtained from its cycle structure.

We investigate this possibility for NORX8 using the cycle decomposition of G . We treat the G function as a 32-bit S-Box. There are 2^{22} combinations of 22 cycles and they correspond to the same amount of invariants of G . Computing Algebraic Normal Form of these functions in order to find their algebraic degree would take 32×2^{32} operations per function and it is infeasible to check all of them. However, we can use the following property: XOR-sum of any boolean function of degree at most d on any affine subspace of dimension d is equal to zero. Consider multiple random 2-dimensional affine subspaces such that there are values from all cycles involved. A necessary condition for any invariant to have degree at most 2 then is that any such subspace contains even number of values for which the invariant is equal to one. It is possible to check if these constraints are consistent using simple linear algebra. Our computations yielded that there are no linear or quadratic invariants for the 32-bit function G . Moreover, the algebraic degree of any such invariant is at least equal to 20 (except for the trivial constant invariants). Therefore, it is not possible to find a nonlinear invariant for the full F round of NORX8 using the method from [6].

3.2 Word Symmetries

A similar symmetry exists on the word level too. Denote by r_n the mapping which rotates its input word left by n bits. We call a word v rotation-invariant with respect to r_n if $r_n(v) = v$. Rotation invariant is preserved through all operations in G except the left shift by 1 bit inside the H function (for binary operations we require both operands to satisfy same rotation-invariant). Nevertheless, the left shift by 1 bit is very similar to the left rotation by 1, which also preserves the invariant. In fact, the left shift by one is equivalent to the left rotation by one when the most significant bit of the input word is equal to zero. The words that are shifted inside the H function are computed as binary ANDs of some two other words. Therefore, any bit of that words is biased to be zero in a $3/4$ fraction of all inputs and all four shifts inside the G function can be approximated with rotations with probability (over all inputs of G) equal to $(3/4)^4 \approx 2^{-1.66}$. However, the bits that must be equal to zero in the approximation are not independent and experimentally we observed a slightly smaller probability of $2^{-1.84}$. The function F contains 8 applications of the G function (4 times on columns and 4 on diagonals), therefore we can approximate all shifts in F with rotations with probability around $2^{-14.72}$.

The largest fraction of states consisting of rotation-invariant words is obtained when n is equal to half of the word size. Moreover, this set contains words that are rotation-invariant for all other n . There are 2^{8w} out of 2^{16w} such states. The fraction is the same as for the state symmetry described in the previous section, but here the invariant is probabilistic.

4 Differential Symmetries

In this section we investigate non-random properties of NORX with respect to symmetric differences. We begin with introducing some terminology.

By definition, a *symmetric difference* (SD) $\Delta = (\Delta_L || \Delta_R)$ is a difference in which the left half is equal to the right half: $\Delta_L = \Delta_R$. A *symmetric differential state* (of NORX) is a state composed of symmetric differences. By analogy, a *symmetric differential characteristic* (SDC) is a characteristic composed of SD. Finally, a *symmetric truncated differential* (STD) over l rounds of NORX is a differential composed of many symmetric characteristics that share the same input SD and can have any output SD after l rounds. Note that symmetric differences propagate through XOR and bit rotation with probability 1 i.e. if the inputs to XOR or rotation are SD, then the output is also SD.

4.1 Symmetric Truncated Differentials in NORX

We are interested in estimating the probability of STDs over multiple rounds of NORX. The motivation is the huge number of SDs per word (bounded by $2^{n/2}$ for every n -bit word), which causes the number of SDC conforming to the same STD to increase exponentially in the number of rounds. As a result the probability of the corresponding STD is also expected to increase significantly.

We apply a branch-and-bound strategy to find high probability SDC in NORX. Our algorithm is similar to Matsui’s algorithm for finding the best differential characteristics in DES [7]. From a given SDC found in this way, we construct an STD containing a huge number of SDCs and we estimate its probability. The search is performed for NORX32 and NORX64 in the scenario `initN` [5] in which the attacker is allowed to modify only the nonce. This is also the most realistic scenario. The results are shown in Table 2.

Table 2 shows that at most 2.125 rounds of F can be distinguished from random using symmetric truncated differentials. Note that since one half of each word of the state is equal to the other half, the probability to randomly obtain a symmetric output state for NORX32 and NORX64 is respectively 2^{-256} and 2^{-512} .

4.2 Estimating the Probability of STD

The STD probabilities are estimated under the Markov assumption i.e. it is assumed that the column/diagonal rounds of NORX are independent. This allows to multiply the probabilities of several non-linear components.

In more detail, the probabilities given in Table 2 are computed in the following manner. Let α, β and γ resp. be the input and output XOR differences to an H operation and let xdp^H be the probability of the corresponding differential $(\alpha, \beta \rightarrow \gamma)$. Let x_i denote the i -th bit of the w -bit word x : $0 \leq i < w$.

In [5, Sect.3.1, Lemma 3], the designers of NORX state the following sufficient and necessary condition for the differential $(\alpha, \beta \rightarrow \gamma)$ to have non-zero

Table 2: Symmetric truncated differentials in NORX32 and NORX64: $F^l - l$ applications of the F function of NORX; $\#H$ – total number of H operations for the given value of l ; HW_{avg} – Hamming weight of the quantity $(\alpha \vee \beta) \ll 1$, averaged over the number of H operations, where α and β are input differences to an H operation; P_{std} – estimation of the probability of the best found symmetric truncated differential (STD) for the given F^l .

F^l	$\#H$	NORX32		NORX64	
		$HW_{\text{avg}}/2$	P_{std}	$HW_{\text{avg}}/2$	P_{std}
$F^{0.125}$	4	0.00	-0.0	0.00	-0.0
$F^{0.250}$	8	0.19	-1.5	0.19	-1.5
$F^{0.375}$	12	0.25	-3.0	0.25	-3.0
$F^{0.500}$	16	0.34	-5.5	0.38	-6.0
$F^{0.625}$	20	0.40	-8.0	0.53	-10.5
$F^{0.750}$	24	0.63	-15.0	0.85	-20.5
$F^{0.875}$	28	0.91	-25.5	1.21	-34.0
$F^{1.000}$	32	1.28	-41.0	1.68	-53.5
$F^{1.125}$	36	1.46	-52.5	2.33	-84.0
$F^{1.250}$	40	1.92	-77.0	3.05	-122.0
$F^{1.375}$	44	2.15	-94.5	3.65	-160.5
$F^{1.500}$	48	2.46	-118.0	4.22	-202.5
$F^{1.625}$	52	2.76	-143.5	4.87	-253.0
$F^{1.750}$	56	3.00	-168.0	5.52	-309.0
$F^{1.875}$	60	3.32	-199.0	6.17	-370.0
$F^{2.000}$	64	3.49	-223.5	6.84	-437.5
$F^{2.125}$	68	3.75	-255.0	7.39	-502.5

probability:

$$(\alpha \oplus \beta \oplus \gamma) \wedge (\neg((\alpha \vee \beta) \ll 1)) = 0 . \quad (2)$$

If condition (2) holds, the probability xdp^H is computed according to the formula:

$$\text{xdp}^H(\alpha, \beta \rightarrow \gamma) = 2^{-HW((\alpha \vee \beta) \ll 1)} , \quad (3)$$

otherwise it is zero. Eq. (2) translates into the following system of bitwise conditions:

$$\begin{cases} (\alpha_i \oplus \beta_i \oplus \gamma_i) = 0 & \text{if } i = 0 , \\ (\alpha_i \oplus \beta_i \oplus \gamma_i) \wedge (\neg(\alpha_{i-1} \vee \beta_{i-1})) = 0 & \text{if } i > 0 . \end{cases} \quad (4)$$

In (4), note that whenever $(\neg(\alpha_{i-1} \vee \beta_{i-1})) = 0$ any value of γ_i is possible. This equation is fulfilled if either $\alpha_{i-1} = 1$ or $\beta_{i-1} = 1$ or both. Therefore, for fixed α, β the quantity $2^{HW((\alpha \vee \beta) \ll 1)}$ gives a number of possible output differences γ . If we fix half of each γ to be equal to the other half, then the number of *symmetric differences* in this set is $A = 2^{HW((\alpha \vee \beta) \ll 1)/2}$.

Assume that for F^l we have found a symmetric differential characteristic with probability P using the mentioned Matsui-like algorithm. Denote with B the number of H operations in this characteristic and let A_{avg} be the Hamming weight of the quantity $(\alpha \vee \beta) \ll 1$ divided by 2 and averaged over all H operations. Then $(A_{\text{avg}})^B$ is an estimation of the number of possible symmetric characteristics derived from the original one and forming a symmetric truncated differential. Under the assumption that all these characteristics have probability equal or close to P , the probability of the STD can be estimated as $P_{\text{std}} = P(A_{\text{avg}})^B$ which is the quantity shown in Table 2.

Example 1. The probability of the best SDC for $F^{2.125}$ of NORX32 found with our algorithm is $P = 2^{-510}$. The number of H operations in this characteristic is $B = 68$ and the average Hamming weight of $((\alpha \vee \beta) \ll 1)$ across all H operations is $510/68 = 7.5$. Therefore the average number of symmetric output differences γ per H operation is $A_{\text{avg}} = 2^{7.5/2} = 2^{3.75}$. Thus the estimated number of possible symmetric characteristics that can be derived from the original one is $(A_{\text{avg}})^B = (2^{3.75})^{68} = 2^{255}$. The probability of the cluster composed of these characteristics is estimated as $P(A_{\text{avg}})^B = 2^{-510} 2^{255} = 2^{-255}$, which is the value shown in the last row of Table 2 for NORX32.

4.3 Closed Sets of Differences

As a second contribution within this section, we performed a search for a set of four differences that is closed under the atomic component of NORX8 composed of one application of the H function, one rotation (under all 4 rotation constants) and one XOR. By *closed*, we mean that if an input difference to the component belongs to a given set, then there will always be an output difference from the same set that has non-zero probability. Our search showed that the only such set is composed of the following 4 symmetric differences (in binary) $(0000 \dots 00_2, 0101 \dots 01_2, 1010 \dots 10_2, 1111 \dots 11_2)$, where the dots denote repetition of the preceding pattern.

In this section we presented analysis of NORX w.r.t. symmetric differences. The latter are a useful tool for clustering multiple symmetric characteristics into single truncated differentials (STD). Although the number of such characteristics in NORX is huge, it is still not enough to compensate for the relatively low probability of every individual characteristic. We conclude that STDs do not pose a threat to the security of NORX.

5 Conclusion

In this paper we presented new non-random properties of the NORX permutation. We showed that it possesses rotational symmetries on different structure levels, which allow to construct efficient distinguishers with very high probability. The security of NORX was also evaluated with respect to a new type of truncated differentials called symmetric truncated differentials (STD). It was

shown that up to 2.125 rounds of the F function of NORX32 and NORX64 can be distinguished using STD. Our results do not pose an immediate threat to the security of the full scheme.

6 Acknowledgements

The work of Aleksei Udovenko is supported by the Fonds National de la Recherche, Luxembourg (project reference 9037104). Vesselin Velichkov is supported by the Internal Research Project CAESAREA of the University of Luxembourg (reference I2R-DIR-PUL-15CAES).

References

1. Aumasson, J.P., Jovanovic, P., Neves, S.: NORX v3.0. CAESAR candidate. <https://competitions.cr.yp.to/round3/norxv30.pdf> (2016)
2. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Duplexing the sponge: Single-pass authenticated encryption and other applications. In Miri, A., Vaudenay, S., eds.: Selected Areas in Cryptography: 18th International Workshop, SAC 2011, Toronto, ON, Canada, August 11-12, 2011, Revised Selected Papers, Berlin, Heidelberg, Springer Berlin Heidelberg (2012) 320–337
3. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Permutation-based encryption, authentication and authenticated encryption. Presented at DIAC 2012, 05–06 July 2012, Stockholm, Sweden. (2012)
4. Aumasson, J.P., Jovanovic, P., Neves, S.: NORX8 and NORX16: Authenticated Encryption for Low-End Systems. Trustworthy Manufacturing and Utilization of Secure Devices—TRUDEVICE (2015)
5. Aumasson, J., Jovanovic, P., Neves, S.: Analysis of NORX: investigating differential and rotational properties. In Aranha, D.F., Menezes, A., eds.: Progress in Cryptology - LATINCRYPT 2014 - Third International Conference on Cryptology and Information Security in Latin America, Florianópolis, Brazil, September 17-19, 2014, Revised Selected Papers. Volume 8895 of Lecture Notes in Computer Science., Springer (2014) 306–324
6. Todo, Y., Leander, G., Sasaki, Y.: Nonlinear Invariant Attack - Practical Attack on Full SCREAM, iSCREAM, and Midori64. In Cheon, J.H., Takagi, T., eds.: Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part II. Volume 10032 of Lecture Notes in Computer Science. (2016) 3–33
7. Matsui, M.: On Correlation Between the Order of S-boxes and the Strength of DES. In Santis, A.D., ed.: EUROCRYPT. Volume 950 of Lecture Notes in Computer Science., Springer (1994) 366–375

A Appendix

A.1 Cycle Decomposition of G from NORX8

In Table 3 we provide the starting points and lengths of cycles of the function G from NORX8.

Table 3: Cycles of G from NORX8. Starting points are of the form (a, b, c, d) (see Figure 2).

Starting point	Cycle length
(00,00,00,02)	3294443807
(00,00,00,11)	621984749
(00,00,00,01)	212798071
(00,00,00,05)	56236016
(00,00,00,0c)	55712043
(00,00,00,bc)	21461014
(00,00,00,ca)	9062510
(00,00,00,f3)	7374122
(00,00,03,7d)	7328319
(00,00,01,6b)	5608893
(00,00,02,45)	2463170
(00,00,1a,c2)	399843
(00,02,57,2c)	52972
(00,01,bd,15)	23344
(00,0f,3a,7a)	8301
(00,07,f5,35)	6339
(00,1d,8b,54)	2124
(00,92,69,ea)	848
(00,05,31,d5)	595
(02,46,c2,5e)	137
(01,c0,8e,d5)	78
(00,00,00,00)	1