Rotational Cryptanalysis on ChaCha Stream Cipher

symmetry

MDPI

Article

Rotational Cryptanalysis on ChaCha Stream Cipher

Stefano Barbero [1,*], Danilo Bazzanella [1] and Emanuele Bellini [2]

[1] Department of Mathematical Sciences "Giuseppe Luigi Lagrange", Politecnico di Torino, 10129 Torino, Italy; danilo.bazzanella@polito.it

[2] Cryptography Research Centre, Technology Innovation Institute, Abu Dhabi P.O. Box 9639, United Arab Emirates; emanuele.bellini@tii.ae

* Correspondence: stefano.barbero@polito.it

Abstract: In this paper we consider the ChaCha20 stream cipher in the related-key scenario and we study how to obtain rotational-XOR pairs with nonzero probability after the application of the first quarter round. The ChaCha20 input can be viewed as a 4×4 matrix of 32-bit words, where the first row of the matrix is fixed to a constant value, the second two rows represent the key, and the fourth some initialization values. Under some reasonable independence assumptions and a suitable selection of the input, we show that the aforementioned probability is about $2^{-251.7857}$, a value greater than 2^{-256}, which is the one expected from a random permutation. We also investigate the existence of constants, different from the ones used in the first row of the ChaCha20 input, for which the rotational-XOR probability increases, representing a potential weakness in variants of the ChaCha20 stream cipher. So far, to our knowledge, this is the first analysis of the ChaCha20 stream cipher from a rotational-XOR perspective.

Keywords: ChaCha20; stream cipher; rotational cryptanalysis; rotational-XOR cryptanalysis

MSC: Primary: 94A60; 11T71; Secondary: 11T06; 94D10

Citation: Barbero, S.; Bazzanella, D.; Bellini, E. Rotational Cryptanalysis on ChaCha Stream Cipher. *Symmetry* **2022**, *14*, 1087. https://doi.org/10.3390/sym14061087

Academic Editors: Chin-Ling Chen and Kuo-Hui Yeh

Received: 17 March 2022
Accepted: 20 May 2022
Published: 25 May 2022

1. Introduction

The ChaCha20 stream cypher, published in 2008 [1], was developed by Daniel J. Bernstein as a modification of Salsa20 [2], another stream cypher designed in 2005 by the same author and then later submitted to the eSTREAM project. The aim of ChaCha20 is to increase diffusion and performance on some architectures with respect to Salsa20. Google has selected ChaCha20 along with Bernstein's Poly1305 message authentication code as a replacement for RC4 in TLS, and its specifications can be found in [3]. Both ciphers are ARX (Add-Rotate-XOR) ciphers, i.e., built on a pseudorandom function based only on the following three 32-bit word operations: modular addition, circular rotation, and bitwise exclusive or (XOR). This pseudorandom function is itself built upon a 512 bit permutation. According to [4], both permutations are not designed to simulate ideal permutations: they are designed to simulate ideal permutations with certain symmetries, i.e., ideal permutations of the orbits of the state space under these symmetries. The input of the ChaCha20 function is partially fixed to specific asymmetric constants, ensuring that different inputs lie in different orbits. As a consequence of the ChaCha20 design, some cryptanalytic techniques, such as rotational cryptanalysis, seem hard to apply.

Related works. Rotational cryptanalysis studies the propagation of rotational pairs $(x, x \lll r)$ throughout the encryption steps of an ARX scheme. It is a probabilistic chosen-plaintext attack that turns to be an effective cryptanalytical tool against ciphers and hash functions that are based on the three ARX operations.

A rotational pair of keys was considered for the first time in the pioneering work on related keys by Biham [5]. This approach was extended by Kelsey et al. in several related-key attacks on block ciphers [6]. A rotational pair of inputs was later adopted in the cryptanalysis of the compression function of Shabal [7] (the term "rotational input" was not used yet), but it was traced only through bitwise operations and not through additions. Bernstein [8] explicitly prevented using rotational pairs in Salsa20 (and ChaCha20) by fixing non-symmetric constants in the input of the permutation. However, he did not provide any complexity or probability estimates for this kind of attack. The designers of the block cipher SEA [9] described the technique of rotational cryptanalysis in 2006 and defended against it with non-linear key-schedule and pseudorandom constants. In his Ph.D. thesis, Daum [10] described the links between modular addition and bit rotations. In 2010, the term "rotational cryptanalysis" appeared for the first time in the papers of Khovratovich and Nikolić [11,12], were they cryptanalyzed the reduced-round Threefish cipher, which is part of the Skein hash function, a SHA-3 competition candidate [13]. Their attack covered 53 rounds of Skein-256 and 57 rounds of Skein-512. In the aforementioned paper [11], the authors proved that for ARX primitives, under the assumption that the cipher can be modeled as a Markov chain, the probability to have a rotational pair of inputs which will produce a rotational pair at the output depends on the number of additions only. This claim was corrected in [14], where the authors showed that chained modular additions used in ARX ciphers do not always form a Markov chain with regards to rotational analysis. They provided a precise value of the probability of such chains and they gave a new algorithm for computing the rotational probability of ARX ciphers. Then, they used the algorithm to correct (from 12 to 7 rounds) the rotational attacks on BLAKE2 [15] and to provide valid rotational attacks against the simplified version of Skein. In 2013, rotational cryptanalysis was also applied to distinguish four rounds of KECCAK permutation [16] in time 2^{221}. A crucial requirement for rotational cryptanalysis to work effectively is that all constants used in the ARX primitive must preserve their values when rotated. This requirement can be relaxed to some extent and instead of assuming completely rotational constants, one can work with constants that are almost rotational, i.e., the XOR difference between the initial constants and the rotated ones gives words of small Hamming weight. A different approach was taken by Ashur and Liu in 2016 [17]. They presented a way to compute the rotational probability when constants are injected into the state and applied their approach to Speck. In particular, they exposed, in the related-key scenario, a trail suggesting that a weak key class of size 2^{39} exists, leading to a 7-round distinguisher for Speck64/32. The technique was later automated in [18], through the use of SAT solvers. In [19], a rotational cryptanalysis of the Salsa core function was presented, also finding rotational distinguishers for the Salsa and Chacha permutations: the rotational distinguisher for the ChaCha permutation performs properly only up to 8 rounds with a probability of approximately $2^{-489.6}$, while the rotational distinguisher for the Salsa permutation performs properly up to 32 rounds with a probability of approximately $2^{-506.752}$, clarifying the weakness of the Salsa core function with respect to rotational cryptanalysis. Finally, in [20] the authors applied rotational cryptanalysis to Chacha20 permutation without considering the injected constants in the state of the permutation, showing that it does not behave as a random permutation for up to 17 rounds, since the related probability is less than 2^{-505} for 17 rounds of ChaCha permutation, while, for a random permutation of the same input size, this probability is 2^{-511}.

Our contribution. In this paper we study the applicability to the ChaCha20 stream cipher quarter round function of the techniques due to Ashur and Liu [17], used for the rotational cryptanalysis of Speck in the presence of constants. In order to do this, we first consider the quarter round function $\mathcal{Q}$ of Chacha20, described in Section 2, determining the conditions which guarantee that, starting from randomly and independently chosen w-bit words, inputs x_1, x_2, x_3 and a w-bit constant c_0 giving the w–bit words y_0, y_1, y_2, y_3 as output, i.e.,

$$(y_0, y_1, y_2, y_3) = \mathcal{Q}(c_0, x_1, x_2, x_3),$$

we can find a characteristics of the type

$$((y_0 \lll \mathsf{r}) \oplus a, (y_1 \lll \mathsf{r}) \oplus b, (y_2 \lll \mathsf{r}) \oplus c, (y_3 \lll \mathsf{r}) \oplus d) = \mathcal{Q}(c_0, f_1, f_2, f_3),$$

for some *input relations* $f_i = f_i(\mathsf{r}, c_0, x_1, x_2, x_3)$, $i = 1, 2, 3$, where r is an integer less than the word size w and a, b, c, d are w-bit strings. Then, we search for a suitable selection of a, b, c, d, and key/nonce/counter relations, the above mentioned *input relations* f_i, $i = 1, 2, 3$, such that, fixing the rotational amount $\mathsf{r} = 1$, we can precisely determine which input constants c_0 yield rotational-XOR pairs with high or low probability, in a related-key scenario. We determine a formula to compute this probability as a function of c_0 and we show that in the case of the constants commonly used in ChaCha20 it is greater than 2^{-256}, revealing the non-randomness of the quarter round permutation. Moreover we investigate the value of this probability for different constants than the ones injected in the ChaCha20 initial state, determining for which of them it increases. While in [19,20] it was only considered the core of ChaCha/Salsa, not the full stream cipher construction, to our knowledge, this is the first attempt to apply rotational analysis to the Chacha20 stream cipher in order to distinguish the behavior of the Chacha20 quarter round from a random permutation and to relate this behaviour to the values of the constants used in Chacha20.

Outline of the paper. In Section 2, we introduce the notation and the essential preliminaries, describing the ChaCha20 stream cipher and the fundamentals of rotational and rotational-XOR cryptanalysis.

In Section 3 we present our analysis. In particular, in Section 3.1, we derive a set of necessary and sufficient conditions which need to be satisfied for a rotational-XOR pair to appear in the output of ChaCha20 quarter round, given that the first word of the input is fixed to a constant value. In Section 3.2, we discuss a suitable choice of the input relations f_i, $i = 1, 2, 3$, and of the parameters a, b, c, d, allowing them to meet the above conditions.In Section 3.3, under the previous choice we compute the probability that these conditions are satisfied by a randomly and independently chosen set of inputs x_i, $i = 1, 2, 3$, as a function of the constant c_0, i.e., the probability that rotational-XOR pairs (with rotational amount $\mathsf{r} = 1$) $(y, (y \lll 1) \oplus \delta)$ appear as outputs of the quarter round. In Section 4 we discuss and resume our results. In particular, in Section 4.1 we evaluate this probability for the special constants commonly used in the ChaCha20 stream cipher. We show that, in our scenario, rotational-XOR pairs can appear in ChaCha20 after one round with probability around $2^{-251.7857}$ against the value of 2^{-256} for a random permutation. In Section 4.2 we describe the form of some initial constants for which these pairs are very likely. Finally, in Section 5 we give our conclusions.

2. Notation and ChaCha20 Stream Cipher Description

In this section, we first define our notation, then we describe the specifications of the ChaCha20 stream cipher and we provide the basics of rotational and rotational-XOR cryptanalysis.

2.1. Notation

Let $\mathbb{F}_2$ be the binary field with two elements, and $\mathcal{M}_{\mathsf{n}\times\mathsf{n}}(\mathbb{F}_2^{\mathsf{w}})$ be the set of all $\mathsf{n} \times \mathsf{n}$ matrices with elements in $\mathbb{F}_2^{\mathsf{w}}$. Depending on the context, lowercase letters stand for w-bit words or for the corresponding non-negative integers they represent, i.e.,

$$x = (x[\mathsf{w}-1], x[\mathsf{w}-2], \ldots, x[1], x[0]) \in \mathbb{F}_2^{\mathsf{w}},$$
$$x = \sum_{i=0}^{\mathsf{w}-1} x[i]2^i \in \mathbb{N}$$

where $x[i] \in \mathbb{F}_2$ for all $i = 0, \ldots, \mathsf{w}-1$. In the case of ChaCha20, we have $\mathsf{n} = 4$ and $\mathsf{w} = 32$. With uppercase letters we indicate a $\mathsf{n} \times \mathsf{n}$ matrix of n^2 words, i.e., $X \in \mathcal{M}_{\mathsf{n}\times\mathsf{n}}(\mathbb{F}_2^{\mathsf{w}})$.

We also use the following notation:

- $\oplus$ for the bitwise exclusive or (XOR), i.e., the addition in $\mathbb{F}_2^{\mathsf{w}}$;
- $\boxplus$ for the w-bit addition mod 2^{w};
- $\boxminus$ for the w-bit subtraction mod 2^{w}, i.e., the sum mod 2^{w} with the opposite of an element in $\mathbb{F}_2^{\mathsf{w}}$;

- $\boxplus_{i=1}^{k} x_i$ for the w-bit addition mod 2^{w} of k words $x_1, \dots, x_k$;
- $x|y$ for the vector bitwise OR operation between x and y;
- $x||y$ for the concatenation of x and y;
- $SHL(x)$ for a non–cyclic left shift by one bit of x;
- $(I \oplus SHL)(x) = x \oplus SHL(x)$;
- $\lll \mathsf{r}$ and $\ggg \mathsf{r}$ respectively for constant-distance left and right circular rotation of r bits of a w-bit word with $1 \leq \mathsf{r} \leq \mathsf{w} - 1$;
- $||x||$ for the Hamming weight of x;
- $|I|$ for the cardinality of a set I;
- $\mathbb{1}_Z$ for the characteristic function of a condition Z, which is equal to 1 when Z is satisfied and equal to 0 otherwise;
- $\underline{1} = (0, 0, \dots, 0, 1) \in \mathbb{F}_2^{\mathsf{w}}$;
- $x \preccurlyeq y$ if and only if we have $x[i] \leq y[i]$ for all $i = 0, \dots, \mathsf{w} - 1$;
- $x = L_h(x)||R_h(x)$, where, for $1 \leq h \leq \mathsf{w} - 1$,

  $$\begin{aligned} L_h(x) &= (x[\mathsf{w}-1], x[\mathsf{w}-2], \dots, x[\mathsf{w}-h]) \\ R_h(x) &= (x[\mathsf{w}-h-1], x[\mathsf{w}-h-2], \dots, x[1], x[0]) \end{aligned}$$

  and, considering $x \in \mathbb{N}$,

  $$x = l_h(x)2^{\mathsf{w}-h} + r_h(x), \quad \text{where} \quad l_h(x) = \textstyle\sum_{i=0}^{h-1} x[\mathsf{w}-h+i]2^i \quad \text{and} \quad r_h(x) = \textstyle\sum_{i=0}^{\mathsf{w}-h-1} x[i]2^i$$

  with $0 \leq l_h(x) \leq 2^h - 1$ and $0 \leq r_h(x) \leq 2^{\mathsf{w}-h} - 1$;
- $\lfloor \cdot \rceil_h$ for the operator which gives for any $x \in \mathbb{N}$ the integer $\lfloor x \rceil_h$ satisfying

  $$0 \leq \lfloor x \rceil_h \leq 2^{h-1} \quad \text{and} \quad \lfloor x \rceil_h \equiv x \bmod 2^h.$$

2.2. *ChaCha20 Specification*

The ChaCha20 permutation has a state of 512 bits, which can be seen as a 4×4 matrix whose elements are binary vectors of $\mathsf{w} = 32$ bits, i.e.,

$$\mathrm{X} = \{x_{i,j}\}_{\substack{i=0,\dots,3 \\ j=0,\dots,3}} = \begin{bmatrix} x_{0,0} & x_{0,1} & x_{0,2} & x_{0,3} \\ x_{1,0} & x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,0} & x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,0} & x_{3,1} & x_{3,2} & x_{3,3} \end{bmatrix} \in \mathcal{M}_{\mathsf{n}\times\mathsf{n}}(\mathbb{F}_2^{\mathsf{w}})\,.$$

The initial state of ChaCha20 is initialized by setting the first row to a 128-bit constant value, the second and third row are used to store a 256-bit key, and the fourth row contains a 64-bit nonce and a 64-bit counter.

Definition 1 (ChaCha20 quarter round)**.** *Let $x_i, y_i, i = 0, 1, 2, 3$ be* w*-bit words, and let $(y_0, y_1, y_2, y_3) = \mathcal{Q}(x_0, x_1, x_2, x_3)$, where $\mathcal{Q}$ is the ChaCha20 quarter round, defined as follows:*

$$\begin{array}{lcl} b_0 = x_0 \boxplus x_1 & (1) & y_0 = b_0 \boxplus b_1 \\ b_3 = (b_0 \oplus x_3) \lll r_1 & & y_3 = (y_0 \oplus b_3) \lll r_3 \\ b_2 = b_3 \boxplus x_2 & (2) & y_2 = y_3 \boxplus b_2 \\ b_1 = (b_2 \oplus x_1) \lll r_2 & & y_1 = (y_2 \oplus b_1) \lll r_4. \end{array}$$

We show in Figure 1 a schematic drawing of the ChaCha20 quarter round. The permutation used in the ChaCha20 stream cipher performs 20 rounds or, equivalently, 10 *double rounds*. Two consecutive rounds (or a *double round*) of the ChaCha20 permutation consist in applying the quarter round four times in parallel to the columns of the state (first round), and then four times in parallel to the diagonals of the state (second round). More formally:

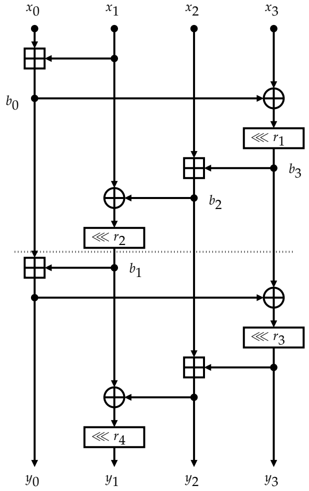

Figure 1. The ChaCha20 quarter round scheme.

Definition 2 (ChaCha20 column/diagonal round)**.** *Let* $X = \{x_{i,j}\}_{\substack{i=0,\dots,3\\j=0,\dots,3}}$, $Y = \{y_{i,j}\}_{\substack{i=0,\dots,3\\j=0,\dots,3}}$ *be two matrices in* $\mathcal{M}_{\mathsf{n}\times\mathsf{n}}(\mathbb{F}_2^{\mathsf{w}})$*.*

A column round $Y = \mathcal{R}^{\mathsf{C}}(X)$ *is defined as follows, with* $i = 0,1,2,3$*:*

$$(y_{0,i}, y_{1,i}, y_{2,i}, y_{3,i}) = \mathcal{Q}(x_{0,i}, x_{1,i}, x_{2,i}, x_{3,i})\,.$$

A diagonal round $Y = \mathcal{R}^{\mathsf{D}}(X)$ *is defined as follows, for* $i = 0,1,2,3$ *and where each subscript is computed modulo* $\mathsf{n} = 4$*:*

$$(y_{0,i}, y_{1,i+1}, y_{2,i+2}, y_{3,i+3}) = \mathcal{Q}(x_{0,i}, x_{1,i+1}, x_{2,i+2}, x_{3,i+3})\,.$$

2.3. Rotational and Rotational-XOR (RX) Cryptanalysis

Rotational cryptanalysis is essentially a distinguishing attack that exploits rotational offsets with probability higher than the one for a random permutation. If we consider a w–bit word x and a rotational offset r, we call $(x, x \lll \mathsf{r})$ a *rotational pair* and we define the *rotational property* as the property that an operation with the input of a rotational pair gives, as output, another rotational pair. Let us denote with $\mathcal{S}$ an ARX scheme with q modular additions. Rotational cryptanalysis is based on the following facts:

- The rotational property is preserved through the XOR of rotational pairs and after a rotation by a constant value r':

$$(x \oplus y) \lll \mathsf{r} = (x \lll \mathsf{r}) \oplus (y \lll \mathsf{r}), \quad (x \lll \mathsf{r}) \ggg \mathsf{r}' = (x \ggg \mathsf{r}') \lll \mathsf{r};$$

- The rotational property is preserved through a modular addition of two w–bit words with a probability given by

$$D_{\mathsf{r}} := \Pr[(x \boxplus y) \lll \mathsf{r} = (x \lll \mathsf{r}) \boxplus (y \lll \mathsf{r})] = \frac{1}{4}(1 + 2^{\mathsf{r}-\mathsf{w}} + 2^{-\mathsf{r}} + 2^{-\mathsf{w}}),$$

  and computed in Corollary 4.12 of [10], this probability is a decreasing function of r, thus it is maximized when $\mathsf{r} = 1$;
- In the case of chained modular additions of more than two w–bit strings, D_{r} must be evaluated using Lemma 2 in [14];
- $\mathcal{S}(x \lll \mathsf{r}) = \mathcal{S}(x) \lll \mathsf{r}$ with probability $(D_{\mathsf{r}})^q$
- Given a random function $\mathcal{P} : \mathbb{F}_2^{\mathsf{w}} \to \mathbb{F}_2^{\mathsf{w}}$, $\mathcal{P}(x \lll \mathsf{r}) = \mathcal{P}(x) \lll \mathsf{r}$ with probability $2^{-\mathsf{w}}$

Thus, we can detect non-randomness in the ARX scheme $\mathcal{S}$ if $(D_{\mathsf{r}})^q > 2^{-\mathsf{w}}$. For example, when $\mathsf{r} = 1$, an ARX scheme implemented with q not chained additions is vulnerable to rotational cryptanalysis if $q < \mathsf{w}/1.415$.

Rotational-XOR cryptanalysis is also a distinguishing attack, firstly introduced in [17] where it was applied to SPECK. This attack studies the propagation of a *rotational-XOR pair* (RX–pair), i.e., a couple $(x, (x \lll r) \oplus \alpha)$ having *RX–difference* α. Clearly RX cryptanalysis is a generalization of rotational cryptanalysis, since they coincide when $\alpha = 0$. Moreover for RX–pairs $(x, (x \lll r) \oplus \alpha)$ and $(y, (y \lll r) \oplus \beta)$ we have

- $((x \oplus y) \lll r) \oplus (\alpha \oplus \beta) = ((x \lll r) \oplus \alpha) \oplus ((y \lll r) \oplus \beta)$;
- $((x \lll \mathsf{r}) \oplus \alpha) \ggg \mathsf{r}' = ((x \ggg \mathsf{r}') \lll r) \oplus (\alpha \ggg \mathsf{r}')$ for a fixed rotational amount r'.

The propagation of RX–differences through modular addition when $\mathsf{r} = 1$ can be computed using the following theorem proved in [17], whose statement has been adapted to our notation

Theorem 1 (Ashur-Liu [17])**.** *Let $x, y \in \mathbb{F}_2^{\mathsf{w}}$ be independent random variables. Let the following $a_1, b_1, a_2, b_2, \Delta_1, \Delta_2$ be constants in $\mathbb{F}_2^{\mathsf{w}}$. Then the probability*

$$\Pr[(((x \oplus a_1) \boxplus (y \oplus b_1)) \oplus \Delta_1) \lll \mathsf{r} = (((x \lll \mathsf{r}) \oplus a_2) \boxplus ((y \lll \mathsf{r}) \oplus b_2)) \oplus \Delta_2] \tag{3}$$

when $\mathsf{r} = 1$ and w is sufficiently large is equal to

$$\mathbb{1}_{\{(a \oplus 1) \preccurlyeq b\}} 2^{-||b||} \cdot p_1 + \mathbb{1}_{\{a \preccurlyeq b\}} \cdot 2^{-||b||} \cdot p_2,$$

where

$$a = (I \oplus SHL)(\delta_1 \oplus \delta_2 \oplus \delta_3), \quad b = SHL((\delta_1 \oplus \delta_3)|(\delta_2 \oplus \delta_3)), \tag{4}$$

$$\delta_1 = R_1(a_1) \oplus L_{\mathsf{w}-1}(a_2), \quad \delta_2 = R_1(b_1) \oplus L_{\mathsf{w}-1}(b_2), \quad \delta_3 = R_1(\Delta_1) \oplus L_{\mathsf{w}-1}(\Delta_2), \tag{5}$$

and $p_1 = 2^{-3}$, $p_2 = 3 \cdot 2^{-3} \simeq 2^{-1.415}$.

3. Searching for Rotational/RX Pairs in ChaCha20

The aim of this section is to consider the ChaCha20 stream cipher, in which the first row in its initial state has constant entries, giving general conditions on the propagation of rotational/RX–pairs that we simply call *rotational propagation*. Moreover, by a suitable choice of inputs and parameters, we find a special case in which the rotational propagation can have a non-negligible probability depending on a certain family of these constants.

3.1. Conditions for the Rotational Propagation

Recalling that the first row of the ChaCha20 4×4 matrix has constant entries, the following proposition holds

Proposition 1. *Let us consider the ChaCha20 quarter round $\mathcal{Q}$, c_0 a constant w–bit string and the rotational amount r with $1 \leq \mathsf{r} \leq \mathsf{w} - 1$. If we have*

$$\mathcal{Q}(c_0, x_1, x_2, x_3) = (y_0, y_1, y_2, y_3)$$

and we consider the w–bit strings a, b, c, d and the entries $f_i = f_i(\mathsf{r}, c_0, x_1, x_2, x_3)$, $i = 1, 2, 3$, then

$$((y_0 \lll \mathsf{r}) \oplus a, (y_1 \lll \mathsf{r}) \oplus b, (y_2 \lll \mathsf{r}) \oplus c, (y_3 \lll \mathsf{r}) \oplus d) = \mathcal{Q}(c_0, f_1, f_2, f_3)$$

$$\Longleftrightarrow$$

$$\begin{gathered} ((c_0 \boxplus x_1) \lll \mathsf{r}) \oplus (x_3 \lll \mathsf{r}) \oplus (a \ggg r_1) \oplus (d \ggg (r_1 + r_3)) = (c_0 \boxplus f_1) \oplus f_3 \\ ((b_3 \boxplus x_2) \lll \mathsf{r}) \oplus (x_1 \lll \mathsf{r}) \oplus (c \ggg r_2) \oplus (b \ggg (r_2 + r_4)) = (((b_3 \lll \mathsf{r}) \oplus a \oplus (d \ggg r_3)) \boxplus f_2) \oplus f_1 \\ (c_0 \boxplus f_1) \boxplus ((b_1 \lll \mathsf{r}) \oplus c \oplus (b \ggg r_4)) = ((c_0 \boxplus x_1 \boxplus b_1) \lll \mathsf{r}) \oplus a \\ ((y_3 \lll \mathsf{r}) \oplus d) \boxplus ((b_3 \lll \mathsf{r}) \oplus a \oplus (d \ggg r_3)) \boxplus f_2 = ((y_3 \boxplus b_3 \boxplus x_2) \lll \mathsf{r}) \oplus c\,. \end{gathered}$$

We call *input relations* the functions $f_i = f_i(\mathsf{r}, c_0, x_1, x_2, x_3)$, $i = 1, 2, 3$.

Proof. If we use the input (c_0, f_1, f_2, f_3) instead of the input (c_0, x_1, x_2, x_3), we find

$$\begin{array}{lll}
\widetilde{b_0} = c_0 \boxplus f_1 & (6) & \widetilde{y_0} = \widetilde{b_0} \boxplus \widetilde{b_1} \\
\widetilde{b_3} = \left(\widetilde{b_0} \oplus f_3\right) \lll r_1 & (7) & \widetilde{y_3} = \left(\widetilde{y_0} \oplus \widetilde{b_3}\right) \lll r_3 \\
\widetilde{b_2} = \widetilde{b_3} \boxplus f_2 & (8) & \widetilde{y_2} = \widetilde{y_3} \boxplus \widetilde{b_2} \\
\widetilde{b_1} = \left(\widetilde{b_2} \oplus f_1\right) \lll r_2 & (9) & \widetilde{y_1} = \left(\widetilde{y_2} \oplus \widetilde{b_1}\right) \lll r_4.
\end{array}$$

We have to satisfy the following conditions

$$\widetilde{y_0} = (y_0 \lll \mathsf{r}) \oplus a \Longleftrightarrow \widetilde{b_0} \boxplus \widetilde{b_1} = ((b_0 \boxplus b_1) \lll \mathsf{r}) \oplus a \tag{10}$$

$$\widetilde{y_3} = (y_3 \lll \mathsf{r}) \oplus d \Longleftrightarrow \left(\widetilde{y_0} \oplus \widetilde{b_3}\right) \lll r_3 = (((y_0 \oplus b_3) \lll r_3) \lll \mathsf{r}) \oplus d \tag{11}$$

$$\widetilde{y_2} = (y_2 \lll \mathsf{r}) \oplus c \Longleftrightarrow \widetilde{y_3} \boxplus \widetilde{b_2} = ((y_3 \boxplus b_2) \lll \mathsf{r}) \oplus c \tag{12}$$

$$\widetilde{y_1} = (y_1 \lll \mathsf{r}) \oplus b \Longleftrightarrow \left(\widetilde{y_2} \oplus \widetilde{b_1}\right) \lll r_4 = (((y_2 \oplus b_1) \lll r_4) \lll \mathsf{r}) \oplus b \tag{13}$$

where (11) and (13) easily give

$$\widetilde{b_3} = (b_3 \lll \mathsf{r}) \oplus a \oplus (d \ggg r_3) \tag{14}$$

from conditions $\widetilde{y_0} = (y_0 \lll \mathsf{r}) \oplus a$ and

$$\widetilde{b_1} = (b_1 \lll \mathsf{r}) \oplus c \oplus (b \ggg r_4), \tag{15}$$

since we have the condition $\widetilde{y_2} = (y_2 \lll \mathsf{r}) \oplus c$. Now from (14) and (7) we find

$$\widetilde{b_0} \oplus f_3 = (b_0 \lll \mathsf{r}) \oplus (x_3 \lll \mathsf{r}) \oplus (a \ggg r_1) \oplus (d \ggg (r_1 + r_3)) \tag{16}$$

and from (15) and (9) we have

$$\widetilde{b_2} \oplus f_1 = (b_2 \lll \mathsf{r}) \oplus (x_1 \lll \mathsf{r}) \oplus (c \ggg r_2) \oplus (b \ggg (r_2 + r_4)) \tag{17}$$

Thus, the four conditions we need to satisfy are (16), (17), (10) and (12). If in (16) we substitute (6) and we take in account (1) in Definition 1, we find the first condition of our thesis

$$((c_0 \boxplus x_1) \lll \mathsf{r}) \oplus (x_3 \lll \mathsf{r}) \oplus (a \ggg r_1) \oplus (d \ggg (r_1 + r_3)) = (c_0 \boxplus f_1) \oplus f_3.$$

If we substitute (8), (14) and (2) in (17) we find the second condition of our thesis

$$\begin{aligned} ((b_3 \boxplus x_2) \lll \mathsf{r}) \oplus (x_1 \lll \mathsf{r}) \oplus (c \ggg r_2) \oplus (b \ggg (r_2 + r_4)) = \\ = (((b_3 \lll \mathsf{r}) \oplus a \oplus (d \ggg r_3)) \boxplus f_2) \oplus f_1. \end{aligned}$$

Moreover, if we substitute (6), (15) and (1) in (10), we obtain the third condition of our thesis

$$(c_0 \boxplus f_1) \boxplus ((b_1 \lll \mathsf{r}) \oplus c \oplus (b \ggg r_4)) = ((c_0 \boxplus x_1 \boxplus b_1) \lll \mathsf{r}) \oplus a.$$

Finally if we substitute (11), (8), (14) and (2) in (12) we have the last condition of our thesis

$$((y_3 \lll \mathsf{r}) \oplus d) \boxplus ((b_3 \lll \mathsf{r}) \oplus a \oplus (d \ggg r_3)) \boxplus f_2 = ((y_3 \boxplus b_3 \boxplus x_2) \lll \mathsf{r}) \oplus c.$$

□

3.2. On the Choices of the Input Relations and of a, b, c, d

One can consider many different choices for the values of f_i and a, b, c, d in order to satisfy the conditions of Proposition 1. Our intention is to set these values in order to obtain simplified conditions which can be satisfied with a non-negligible probability depending on the values given to c_0. Considering Equations (14) and (15), a first simplification we apply is to choose a, b, c and d such that

$$a \oplus (d \ggg r_3) = 0, \quad c \oplus (b \ggg r_4) = 0. \tag{18}$$

With this setting the conditions of Proposition 1 become

$$\begin{aligned} (c_0 \boxplus f_1) \oplus f_3 &= ((c_0 \boxplus x_1) \lll \mathsf{r}) \oplus (x_3 \lll \mathsf{r}) & (19)\\ ((b_3 \lll \mathsf{r}) \boxplus f_2) \oplus f_1 &= ((b_3 \boxplus x_2) \lll \mathsf{r}) \oplus (x_1 \lll \mathsf{r}) & (20)\\ (c_0 \boxplus f_1) \boxplus (b_1 \lll \mathsf{r}) &= ((c_0 \boxplus x_1 \boxplus b_1) \lll \mathsf{r}) \oplus a & (21)\\ ((y_3 \lll \mathsf{r}) \oplus d) \boxplus (b_3 \lll \mathsf{r}) \boxplus f_2 &= ((y_3 \boxplus b_3 \boxplus x_2) \lll \mathsf{r}) \oplus c, & (22) \end{aligned}$$

and we have (21) and (22) satisfied when

$$\begin{aligned} a &= [(c_0 \boxplus f_1) \boxplus (b_1 \lll \mathsf{r})] \oplus [(c_0 \boxplus x_1 \boxplus b_1) \lll \mathsf{r}] \\ c &= [((y_3 \lll \mathsf{r}) \oplus d) \boxplus (b_3 \lll \mathsf{r}) \boxplus f_2] \oplus [(y_3 \boxplus b_3 \boxplus x_2) \lll \mathsf{r}]. \end{aligned} \tag{23}$$

Regarding the input relations, we consider expressions of the kind $f_i = f_i(x_i, c_0, \mathsf{r})$, which seem reasonable choices once we look to the remaining conditions (19) and (20). Indeed we may select f_3 and f_2 as

$$f_3 = x_3 \lll \mathsf{r}, \quad f_2 = x_2 \lll \mathsf{r}, \tag{24}$$

in order to simplify (19) and to give an expression for (20) very similar to the condition of rotational propagation through modular addition, obtaining

$$\begin{aligned} (c_0 \boxplus f_1) &= (c_0 \boxplus x_1) \lll \mathsf{r} & (25)\\ ((b_3 \lll \mathsf{r}) \boxplus (x_2 \lll \mathsf{r})) \oplus f_1 &= ((b_3 \boxplus x_2) \lll \mathsf{r}) \oplus (x_1 \lll \mathsf{r}). & (26) \end{aligned}$$

Now we have to think about f_1 and, consequently, about how to manage conditions (25) and (26). Clearly a possibility is to choose $f_1 = x_1 \lll \mathsf{r}$ in order to reduce (26) exactly to the condition for the rotational propagation through modular addition

$$(b_3 \lll \mathsf{r}) \boxplus (x_2 \lll \mathsf{r}) = (b_3 \boxplus x_2) \lll \mathsf{r}. \tag{27}$$

In this case (25) becomes

$$(c_0 \boxplus x_1) \lll \mathsf{r} = c_0 \boxplus (x_1 \lll \mathsf{r}), \tag{28}$$

which holds for some x_1 only under particular conditions on c_0. These conditions turn out to be too restrictive for our purposes. At the end of this subsection, we will prove them in Proposition 2, explaining why they give rise to a small number of possible choices for c_0 in order to obtain nonzero probability for (25). This part can be skipped over by the reader without problems. Therefore we may choose f_1 as something different from $x_1 \lll \mathsf{r}$ and consequently we may consider (26) as a condition of the form

$$(b_3 \lll \mathsf{r}) \boxplus (x_2 \lll \mathsf{r}) = ((b_3 \boxplus x_2) \oplus \sigma(x_1, c_0, \mathsf{r})) \lll \mathsf{r} \tag{29}$$

where σ in general is not a constant, since it depends on x_1. But we may combine the law of total probability with the Ashur-Liu Theorem 1 in [17] in order to find a probability estimate for (29) in the case $\mathsf{r} = 1$. So a good choice is to consider the case $\mathsf{r} = 1$ and use (25) in order to define f_1 as we will do in the next subsection.

Proposition 2. *If we consider*

$$\begin{aligned} c_0 &= l_{\mathsf{r}}(c_0)2^{\mathsf{w}-\mathsf{r}} + r_{\mathsf{r}}(c_0), \\ c_0 \ggg \mathsf{r} &= l_{\mathsf{r}}(c_0 \ggg \mathsf{r})2^{\mathsf{w}-\mathsf{r}} + r_{\mathsf{r}}(c_0 \ggg \mathsf{r}), \\ x_1 &= l_{\mathsf{r}}(x_1)2^{\mathsf{w}-\mathsf{r}} + r_{\mathsf{r}}(x_1), \end{aligned} \tag{30}$$

and

$$\gamma_L = \mathbb{1}_{\{l_{\mathsf{r}}(c_0 \ggg \mathsf{r}) + l_{\mathsf{r}}(x_1) \geq 2^{\mathsf{r}}\}}, \quad \gamma_R = \mathbb{1}_{\{r_{\mathsf{r}}(c_0) + r_{\mathsf{r}}(x_1) \geq 2^{\mathsf{w}-\mathsf{r}}\}}$$

then condition (28) *holds if and only if we have*

$$r_{\mathsf{r}}(c_0) = \lfloor r_{\mathsf{r}}(c_0 \ggg \mathsf{r}) + \gamma_L \rfloor_{\mathsf{w}-\mathsf{r}}, \quad l_{\mathsf{r}}(c_0 \ggg \mathsf{r}) = \lfloor l_{\mathsf{r}}(c_0) + \gamma_R \rfloor_{\mathsf{r}} \tag{31}$$

or, in other words, if and only if

$$c_0 \boxminus (c_0 \ggg \mathsf{r}) = \lfloor (-\gamma_R)2^{\mathsf{w}-\mathsf{r}} + \gamma_L \rfloor_{\mathsf{w}}.$$

Proof. The proof is quite straightforward, since from (30) we have

$$c_0 = (c_0 \ggg \mathsf{r}) \lll \mathsf{r} = r_{\mathsf{r}}(c_0 \ggg \mathsf{r})2^{\mathsf{r}} + l_{\mathsf{r}}(c_0 \ggg \mathsf{r}), \quad x_1 \lll \mathsf{r} = r_{\mathsf{r}}(x_1)2^{\mathsf{r}} + l_{\mathsf{r}}(x_1)$$

thus

$$c_0 \boxplus (x_1 \lll \mathsf{r}) = \lfloor r_{\mathsf{r}}(c_0 \ggg \mathsf{r}) + r_{\mathsf{r}}(x_1) + \gamma_L \rfloor_{\mathsf{w}-\mathsf{r}} 2^{\mathsf{r}} + \lfloor l_{\mathsf{r}}(c_0 \ggg \mathsf{r}) + l_{\mathsf{r}}(x_1) \rfloor_{\mathsf{r}} \tag{32}$$

and, since

$$c_0 \boxplus x_1 = \lfloor l_{\mathsf{r}}(c_0) + l_{\mathsf{r}}(x_1) + \gamma_R \rfloor_{\mathsf{r}} 2^{\mathsf{w}-\mathsf{r}} + \lfloor r_{\mathsf{r}}(x_1) + r_{\mathsf{r}}(c_0) \rfloor_{\mathsf{w}-\mathsf{r}},$$

we obtain

$$(c_0 \boxplus x_1) \lll \mathsf{r} = \lfloor r_{\mathsf{r}}(x_1) + r_{\mathsf{r}}(c_0) \rfloor_{\mathsf{w}-\mathsf{r}} 2^{r} + \lfloor l_{\mathsf{r}}(c_0) + l_{\mathsf{r}}(x_1) + \gamma_R \rfloor_{\mathsf{r}}. \tag{33}$$

Therefore comparing (32) and (33) we find (31). □

Following this choice, with arguments similar to the ones used by Daum in [10], we may evaluate some probability different from zero for condition (28) only when one of the following conditions is satisfied

- $c_0 \boxminus (c_0 \ggg \mathsf{r}) = 2^{\mathsf{w}} - 2^{\mathsf{w}-\mathsf{r}} + 1$, i.e., $\gamma_L = \gamma_R = 1$, which is equivalent to the equation

  $$(l_{\mathsf{r}}(c_0 \ggg \mathsf{r}) - 1)(2^{\mathsf{w}-\mathsf{r}} - 1) = r_{\mathsf{r}}(c_0 \ggg \mathsf{r})(2^{\mathsf{r}} - 1)$$

  and it gives $2^{\gcd(\mathsf{w},\mathsf{r})} - 1$ possible values for c_0;
- $c_0 \boxminus (c_0 \ggg \mathsf{r}) = 1$, i.e., $\gamma_L = 1$, $\gamma_R = 0$, which is equivalent to the equation

  $$r_{\mathsf{r}}(c_0 \ggg \mathsf{r})(2^{\mathsf{r}} - 1) - l_{\mathsf{r}}(c_0 \ggg \mathsf{r})(2^{\mathsf{w}-\mathsf{r}} - 1) = 1$$

  and it gives one value for c_0 only when $\gcd(\mathsf{w}, \mathsf{r}) = 1$ and otherwise it is impossible;
- $c_0 \boxminus (c_0 \ggg \mathsf{r}) = 0$, i.e., $\gamma_L = \gamma_R = 0$, which is equivalent to the equation

  $$l_{\mathsf{r}}(c_0 \ggg \mathsf{r})(2^{\mathsf{w}-\mathsf{r}} - 1) = r_{\mathsf{r}}(c_0 \ggg \mathsf{r})(2^{\mathsf{r}} - 1)$$

  and it gives $2^{\gcd(\mathsf{w},\mathsf{r})}$ possible values for c_0;
- $c_0 \boxminus (c_0 \ggg \mathsf{r}) = 2^{\mathsf{w}} - 2^{\mathsf{w}-\mathsf{r}}$, i.e., $\gamma_L = 0$, $\gamma_R = 1$, which is equivalent to the equation

  $$(l_{\mathsf{r}}(c_0 \ggg \mathsf{r}) - 1)(2^{\mathsf{w}-\mathsf{r}} - 1) - r_{\mathsf{r}}(c_0)(2^{\mathsf{r}} - 1) = 1$$

  and it gives one value for c_0 only when $\gcd(\mathsf{w}, \mathsf{r}) = 1$ and otherwise it is impossible.

The strong dependence from $\gcd(\mathsf{w},\mathsf{r})$ of the possible values of c_0 for which (25) holds for some x_1 seems too limiting for our purposes. For example, when $\gcd(\mathsf{w},\mathsf{r}) = 1$ we have only two suitable values for c_0 from the third case and only a single value from the remaining ones.

3.3. Probability of Rotational Propagation for 1 Bit Rotations

From what we have previously pointed out, we will consider from now on $\mathsf{r} = 1$ with the choices (24), (35) and (29), where

$$\sigma(x_1, c_0, 1) = (f_1 \ggg 1) \oplus x_1. \tag{34}$$

and we use condition (25) in order to define f_1

$$c_0 \boxplus f_1 = (c_0 \boxplus x_1) \lll 1 \quad \Rightarrow \quad f_1(x_1, c_0, 1) = ((c_0 \boxplus x_1) \lll 1) \boxminus c_0 \tag{35}$$

This definition of f_1 seems reasonable since we can simultaneously have (25) automatically satisfied and a better range of possible values for $\sigma(x_1, c_0, 1)$ that may give a non-negligible probability for condition (26).

We will follow this path: first we will prove the following Corollary 1 of Theorem 1, then we will use it together with the law of total probability in order to find a formula for the probability of (29) depending on c_0 and in particular on the cardinalities of some sets related to c_0. Finally, we will evaluate these cardinalities in Lemma 1 and we will give in Theorem 2 the value of this probability making explicit its dependence on c_0.

Corollary 1. *Let $x, y \in \mathbb{F}_2^{\mathsf{w}}$ be independent random variables, $\mathsf{r} = 1$ and*

$$\sigma = L_1(\sigma)||R_1(\sigma) = \sigma[\mathsf{w}-1]||(\sigma[w-2], \sigma[\mathsf{w}-3], \ldots, \sigma[1], \sigma[0])$$

a constant word in $\mathbb{F}_2^{\mathsf{w}}$. If j, $0 \leq j \leq \mathsf{w}-1$, is the index of the first bit equal to 0 in $R_1(\sigma)$ starting from the right to the left, with the convention that $j = \mathsf{w}-1$ only when $R_1(\sigma)$ has all the entries equal to 1, then

$$\Pr[((x \boxplus y) \oplus \sigma) \lll 1 = (x \lll 1) \boxplus (y \lll 1)] = P(\sigma), \tag{36}$$

where

$$P(\sigma) = \begin{cases} p_2 & \text{if } j = 0 \text{ and } \sigma[i] = 0 \text{ for all } i = 0, \ldots \mathsf{w}-2 \\ 2^{-j} p_1 & \text{if } j = \mathsf{w}-1 \text{ or, if } j \geq 1, \sigma[i] = 0 \text{ for all } i = j+1, \ldots, \mathsf{w}-2 \\ 0 & \text{otherwise.} \end{cases} \tag{37}$$

Proof. We observe that (36) is a special case of (3) when $a_1 = a_2 = b_1 = b_2 = \Delta_2 = 0$, $\sigma = \Delta_1$. Thus from (5) we find $\delta_1 = \delta_2 = 0$ and $\delta_3 = R_1(\sigma)$. Therefore

$$\delta_1 \oplus \delta_2 \oplus \delta_3 = R_1(\sigma) \quad (\delta_1 \oplus \delta_3)|(\delta_2 \oplus \delta_3) = (\delta_3|\delta_3) = \delta_3 = R_1(\sigma)$$

$$b = SHL(R_1(\sigma)) = (\sigma[\mathsf{w}-3], \sigma[\mathsf{w}-4], \ldots, \sigma[1], \sigma[0], 0) \tag{38}$$

and

$$\begin{aligned} a = (I \oplus SHL)(\delta_1 \oplus \delta_2 \oplus \delta_3) = (I \oplus SHL)(R_1(\sigma)) = R_1(\sigma) \oplus SHL(R_1(\sigma)) = \\ = (\sigma[\mathsf{w}-2] \oplus \sigma[\mathsf{w}-3], \ldots, \sigma[h+1] \oplus \sigma[h], \ldots, \sigma[1] \oplus \sigma[0], \sigma[0]) \end{aligned} \tag{39}$$

$$a \oplus \underline{1} = (\sigma[\mathsf{w}-2] \oplus \sigma[\mathsf{w}-3], \ldots, \sigma[h+1] \oplus \sigma[h], \ldots, \sigma[1] \oplus \sigma[0], \sigma[0] \oplus 1). \tag{40}$$

We obtain $\mathbb{1}_{\{(a \oplus \underline{1}) \preccurlyeq b\}} = 1$ if and only if we have

$$\sigma[0] \oplus 1 \leq 0 \Rightarrow \sigma[0] = 1$$

and, for all $0 \le h \le \mathsf{w}-3$,

$$\sigma[h+1] \oplus \sigma[h] \le \sigma[h]. \tag{41}$$

Clearly, condition (41) holds when all the bits $\sigma[h]$ are equal to 1 for any $\sigma[h+1] \in \mathbb{F}_2$, while, if $\sigma[j]$, $j \ge 1$, is the first bit equal to 0 in $R_1(\sigma)$ starting from the right to the left, condition (41) holds for $h \ge j$ if and only if $\sigma[h+1] = 0$. Thus in order to have $\mathbb{1}_{\{(a \oplus \underline{1}) \preccurlyeq b\}} = 1$, we need $R_1(\sigma)$ such that $||R_1(\sigma)|| = j \le \mathsf{w}-2$, where all the first j bits from $\sigma[0]$ to $\sigma[j-1]$ are equal to 1 and the remaining ones from $\sigma[j]$ to $\sigma[\mathsf{w}-2]$ are equal to 0, or $||R_1(\sigma)|| = \mathsf{w}-1$, i.e., all the bits of $R_1(\sigma)$ are equal to 1. In these cases we find from Theorem 1 that $P(\sigma) = 2^{-j}p_1$. On the other hand $\mathbb{1}_{\{a \preccurlyeq b\}}$ is equal to 1 if and only if we have

$$\sigma[0] \le 0 \Rightarrow \sigma[0] = 0$$

and for all $0 \le h \le \mathsf{w}-3$ condition (41) holds. Since $\sigma[h] = 0$ in (41) implies $\sigma[h+1] = 0$ and we have $\sigma[0] = 0$, an easy inductive argument shows that all the bits of $R_1(\sigma)$ must be equal to 0 in order to have $\mathbb{1}_{\{a \preccurlyeq b\}} = 1$. Thus $||R_1(\sigma)|| = 0$ and we find from Theorem 1 $P(\sigma) = p_2$. We finally observe that the two characteristic functions $\mathbb{1}_{\{(a \oplus \underline{1}) \preccurlyeq b\}}$ and $\mathbb{1}_{\{a \preccurlyeq b\}}$ can not be contemporarily equal to 1, and they are contemporarily equal to 0, with $P(\sigma) = 0$, for any other value of $R_1(\sigma)$ different from the ones we have pointed out. □

Thanks to the law of total probability, since for a fixed v we may assume $\sigma(v, c_0, 1)$ as a constant while b_3 and x_2 are independent variables, we find

$$\begin{aligned}
&\Pr[((b_3 \boxplus x_2) \oplus \sigma(x_1, c_0, 1)) \lll 1 = (b_3 \lll 1) \boxplus (x_2 \lll 1)] = \\
&= \sum_{v \in \mathbb{F}_2^{\mathsf{w}}} \Pr[x_1 = v] \cdot \Pr[((b_3 \boxplus x_2) \oplus \sigma(x_1, c_0, 1)) \lll 1 = (b_3 \lll 1) \boxplus (x_2 \lll 1) | x_1 = v] = \\
&= \frac{1}{2^{\mathsf{w}}} \sum_{v \in \mathbb{F}_2^{\mathsf{w}}} \Pr[((b_3 \boxplus x_2) \oplus \sigma(v, c_0, 1)) \lll 1 = (b_3 \lll 1) \boxplus (x_2 \lll 1)] = \\
&= \frac{1}{2^{\mathsf{w}}} \sum_{v \in \mathbb{F}_2^{\mathsf{w}}} P(\sigma(v, c_0, 1)) = P(c_0),
\end{aligned}$$

where we used the fact that x_1 is uniformly distributed, so $\Pr[x_1 = v] = \frac{1}{2^{\mathsf{w}}}$. In order to explicitly evaluate $P(c_0)$ applying the results of Corollary 1, we now define sets of w–bit words having a special form and, in the next Lemma 1, we will find their cardinalities, depending on c_0.

Definition 3. *Let us consider the vectors $u(\delta, h) \in \mathbb{F}_2^{\mathsf{w}}$ such that*

$$u(\delta, h) = (\delta, \underbrace{0,0,0,\ldots,0}_{\mathsf{w}-1-h\text{–zeros}}, \underbrace{1,1,1,\ldots,1}_{h\text{–ones}}) \quad \text{where} \quad h = 0,1,\ldots \mathsf{w}-1,\ \delta \in \mathbb{F}_2 \tag{42}$$

We define the sets I_h as

$$I_h = \{v \in \mathbb{F}_2^{\mathsf{w}} : \sigma(v, c_0, 1) = u(\delta, h),\ \delta \in \mathbb{F}_2\}. \tag{43}$$

Therefore, from Corollary 1 we have the following explicit expression for $P(c_0)$

$$\begin{aligned}
P(c_0) &= \frac{1}{2^{\mathsf{w}}} \sum_{v \in \mathbb{F}_2^{\mathsf{w}}} P(\sigma(v, c_0, 1)) = \frac{1}{2^{\mathsf{w}}} \left(|I_0| p_2 + p_1 \sum_{h=1}^{\mathsf{w}-1} |I_h| \cdot 2^{-h} \right) = \\
&= \frac{1}{2^{\mathsf{w}+3}} \left(3|I_0| + \sum_{h=1}^{\mathsf{w}-1} |I_h| \cdot 2^{-h} \right),
\end{aligned} \tag{44}$$

where, for $h = 0, \ldots, \mathsf{w}-1$, we evaluate the cardinalities $|I_h|$ in the following lemma.

Lemma 1. *Let us consider σ and f_1 respectively defined as in (34) and (35). Then we have*

$$|I_0| = \begin{cases} 2^{\mathsf{w}-1} & \text{if } c_0 = (\underbrace{0,0,0,\ldots,0}_{\mathsf{w}-1\text{–zeros}},1) \text{ or } c_0 = (\underbrace{1,1,1,1,\ldots,1}_{\mathsf{w}\text{–ones}}) \\ 2^{\mathsf{w}} & \text{if } c_0 = (\underbrace{0,0,0,0,\ldots,0}_{\mathsf{w}\text{–zeros}}) \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

and

$$|I_h| = \begin{cases} 2 & \text{if } h = \mathsf{w}-1,\, \mathsf{w}-2 \text{ and } c_0[0] = 1 \\ 2^{\mathsf{w}-1-h} & \text{if } 1 \le h \le \mathsf{w}-3,\, c_0[0] = 1 \\ & \text{and } c_0[i] = j \in \mathbb{F}_2 \text{ for all } i = h+1,\ldots,\mathsf{w}-1 \\ 4 & \text{if } h = \mathsf{w}-1,\, \mathsf{w}-2 \text{ and } c_0[0] = 0,\, c_0[1] = 1 \\ 2^{\mathsf{w}-h} & \text{if } 1 \le h \le \mathsf{w}-3,\, c_0[0] = 0, c_0[1] = 1 \\ & \text{and } c_0[i] = j \in \mathbb{F}_2 \text{ for all } i = h+1,\ldots,\mathsf{w}-1 \\ 0 & \text{otherwise.} \end{cases}$$

Proof. First of all, we recall that $v \in I_h$ if and only if

$$\sigma(v, c_0, 1) = (f_1(v, c_0, 1) \ggg 1) \oplus v = u(\delta, h)$$

or, equivalently, if and only if

$$f_1(v, c_0, 1) = (u(\delta, h) \oplus v) \lll 1$$

and if we use (35) we finally have the condition

$$v \in I_h \iff (c_0 \boxplus v) \lll 1 = ((u(\delta, h) \oplus v) \lll 1) \boxplus c_0. \tag{46}$$

If we consider

$$\begin{aligned} c_0 &= l_1(c_0)2^{\mathsf{w}-1} + r_1(c_0) = c_0[\mathsf{w}-1]2^{\mathsf{w}-1} + \sum_{i=0}^{\mathsf{w}-2} c_0[i]2^i, \\ v &= l_1(v)2^{\mathsf{w}-1} + r_1(v) = v[\mathsf{w}-1]2^{\mathsf{w}-1} + \sum_{i=0}^{\mathsf{w}-2} v[i]2^i, \end{aligned} \tag{47}$$

we find

$$(c_0 \boxplus v) \lll 1 = \lfloor r_1(c_0) + r_1(v) \rfloor_{\mathsf{w}-1} \cdot 2 + \lfloor c_0[\mathsf{w}-1] + v[\mathsf{w}-1] + \gamma_R \rfloor_1, \tag{48}$$

where $\gamma_R = \mathbb{1}_{\{r_1(c_0)+r_1(v) \ge 2^{\mathsf{w}-1}\}}$. On the other hand, we have

$$(u(\delta, h) \oplus v) \lll 1 = (v[\mathsf{w}-2] \oplus u[\mathsf{w}-2], v[\mathsf{w}-3] \oplus u[\mathsf{w}-3], \ldots, v[0] \oplus u[0], v[\mathsf{w}-1] \oplus \delta)$$

and, since

$$c_0 = 2\left(c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \frac{1}{2}\sum_{i=1}^{\mathsf{w}-2} c_0[i]2^i\right) + c_0[0] = 2\left(c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \frac{r_1(c_0) - c_0[0]}{2}\right) + c_0[0],$$

$$u(\delta, h) \oplus v = l_1((u(\delta, h) \oplus v))2^{\mathsf{w}-1} + r_1((u(\delta, h) \oplus v),$$

where $l_1((u(\delta,h)\oplus v) = v[w-1]\oplus\delta$, and by Definition 3

$$\begin{aligned}\overline{v} = r_1((u(\delta,h)\oplus v) = \sum_{i=0}^{\mathsf{w}-2}(v[i]\oplus u[i])2^i &= \\ = \sum_{i=h}^{\mathsf{w}-2}(v[i]\oplus u[i])2^i + \sum_{i=0}^{h-1}(v[i]\oplus u[i])2^i &= \\ = \sum_{i=h}^{\mathsf{w}-2}v[i]2^i + \sum_{i=0}^{h-1}(v[i]\oplus 1)2^i&,\end{aligned} \tag{49}$$

we obtain

$$\begin{aligned}&((u(\delta,h)\oplus v)\lll 1)\boxplus c_0 = \\ &= \left\lfloor c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \frac{r_1(c_0)-c_0[0]}{2} + \overline{v} + \gamma_L\right\rfloor_{\mathsf{w}-1}\cdot 2 + \lfloor c_0[0] + (v[\mathsf{w}-1]\oplus\delta)\rfloor_1,\end{aligned} \tag{50}$$

where $\gamma_L = \mathbb{1}_{\{c_0[0]+(v[\mathsf{w}-1]\oplus\delta)\geq 2\}}$.

Therefore, comparing (48) and (50), we find that condition (46) is equivalent to the system of congruences

$$\begin{cases} r_1(c_0)+r_1(v) \equiv c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \dfrac{r_1(c_0)-c_0[0]}{2} + \overline{v} + \gamma_L \bmod 2^{\mathsf{w}-1} \\ c_0[\mathsf{w}-1] + v[\mathsf{w}-1] + \gamma_R \equiv c_0[0] + (v[\mathsf{w}-1]\oplus\delta) \bmod 2, \end{cases} \tag{51}$$

where, using (47) and (49), and observing that

$$\sum_{i=0}^{h-1}(v[i]+(v[i]\oplus 1))2^i = \sum_{i=0}^{h-1}2^i = 2^h-1, \quad \frac{r_1(c_0)+c_0[0]}{2} = \sum_{i=0}^{\mathsf{w}-3}c_0[i+1]2^i + c_0[0],$$

we can rewrite the first congruence of (51) as

$$\sum_{i=0}^{\mathsf{w}-3}c_0[i+1]2^i + c_0[0] + 2^h - 1 \equiv c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + 2\sum_{i=0}^{h-1}(v[i]\oplus 1)2^i + \gamma_L \bmod 2^{\mathsf{w}-1}. \tag{52}$$

Now, in solving (52) we have to consider the following cases.

- $c_0[0]=1,\ h\geq 1:$ in this case when $h=\mathsf{w}-1$ congruence (52) becomes

  $$\begin{aligned}&\sum_{i=1}^{\mathsf{w}-3}c_0[i+1]2^i + c_0[1] \equiv \\ \equiv &c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \sum_{i=1}^{\mathsf{w}-3}(v[i-1]\oplus 1)2^i + (v[\mathsf{w}-3]\oplus 1)2^{\mathsf{w}-2} + \gamma_L \bmod 2^{\mathsf{w}-1}.\end{aligned}$$

  Comparing the two members it clearly holds if and only if

  $$v[i-1] = c_0[i+1]\oplus 1, \quad i=1,\dots,\mathsf{w}-2, \quad \text{and} \quad c_0[1] = \gamma_L \tag{53}$$

  and, by definition of γ_L, the second congruence in (51) holds if

  $$\begin{cases} v[\mathsf{w}-1] \equiv c_0[\mathsf{w}-1] + \gamma_R \bmod 2 & \text{and, if } \gamma_L = 1 \quad \delta = v[\mathsf{w}-1]\oplus 1 \\ v[\mathsf{w}-1] \equiv c_0[\mathsf{w}-1] + \gamma_R + 1 \bmod 2 & \text{and, if } \gamma_L = 0 \quad \delta = v[\mathsf{w}-1] \end{cases} \tag{54}$$

  where γ_R is fixed by (53) and by the free choice of $v[\mathsf{w}-2]$. Therefore we always have only 2 solutions when $h=\mathsf{w}-1$. When $h=\mathsf{w}-2$ congruence (52) becomes

$$\sum_{i=1}^{\mathrm{w}-3} c_0[i+1]2^i + c_0[1] + 2^{\mathrm{w}-2} \equiv$$
$$\equiv c_0[\mathrm{w}-1]2^{\mathrm{w}-2} + \sum_{i=1}^{\mathrm{w}-3}(v[i-1]\oplus 1)2^i + (v[\mathrm{w}-3]\oplus 1)2^{\mathrm{w}-2} + \gamma_L \bmod 2^{\mathrm{w}-1}$$

giving $v[\mathrm{w}-3] = [(1+c_0[\mathrm{w}-1]) \bmod 2] \oplus 1$ and conditions similar to (53)

$$v[i-1] = c_0[i+1]\oplus 1, \quad i=1,\dots,\mathrm{w}-3, \quad \text{and} \quad c_0[1] = \gamma_L$$

and the same conditions (54) for congruence (51). Thus also when $h = \mathrm{w}-2$ we find that we do not have conditions on $v[\mathrm{w}-2]$ and we always have only 2 solutions. Finally if $1 \le h \le \mathrm{w}-3$ both members of the congruence (52) are less than $2^{\mathrm{w}-1}$ and they have the same parity if and only if $c_0[1] = \gamma_L$. Moreover, we have

$$2^{\mathrm{w}-2} - 2^{h+1} = \sum_{i=h+1}^{\mathrm{w}-3} 2^i \ge \sum_{i=h+1}^{w-3} c_0[i+1]2^i,$$

where the equality holds if and only if $c_0[i+1] = 1$ for all $i = h+1,\dots,\mathrm{w}-3$, and

$$\sum_{i=1}^{h-1} c_0[i+1]2^i \le 2^h - 2, \quad \sum_{i=1}^{h-1}(v[i-1]\oplus 1)2^i \le 2^h - 2,$$

where the equalities hold if $c_0[i+1] = (v[h-1]\oplus 1) = 1$ for all $i = 1,\dots,h-1$. Thus we find

$$\sum_{i=h+1}^{\mathrm{w}-3} c_0[i+1]2^i + (c_0[h+1]+1)2^h + \sum_{i=1}^{h-1} c_0[i+1]2^i =$$
$$= c_0[\mathrm{w}-1]2^{\mathrm{w}-2} + (v[h-1]\oplus 1)2^h + \sum_{i=1}^{h-1}(v[i-1]\oplus 1)2^i,$$

which gives $v[i-1] = c_0[i+1]\oplus 1$ for all $i = 1,\dots,h-1$ and, if $c_0[\mathrm{w}-1] = j \in \mathbb{F}_2$, we need $c_0[i+1] = j$ for all $i = h+1,\dots,\mathrm{w}-3$ and $v[h-1] = j$. Therefore we have solutions only when the constant c_0 is such that $c_0[i] = j \in \mathbb{F}_2$ for all $i = h+1,\dots,\mathrm{w}-1$ and, since also in these cases the same conditions (54) for congruence (51) hold, we have $v[i]$ free for all $i = h, h+1,\dots,\mathrm{w}-2$ and $2^{\mathrm{w}-1-h}$ possible solutions.

- $c_0[0] = 0, h \ge 1$: in this case we necessarily have $\gamma_L = 0$ and $v[\mathrm{w}-1]\oplus\delta \in \mathbb{F}_2$ since $\gamma_L = 1$ only when $c_0[0] = v[\mathrm{w}-1]\oplus\delta = 1$, thus (52) becomes

$$\sum_{i=1}^{\mathrm{w}-3} c_0[i+1]2^i + c_0[1] + 2^h - 1 \equiv c_0[\mathrm{w}-1]2^{\mathrm{w}-2} + 2\sum_{i=0}^{h-1}(v[i]\oplus 1)2^i \bmod 2^{\mathrm{w}-1} \tag{55}$$

and we have solutions only when $c_0[1] = 1$, in order to preserve the same parity for both members. Under this supplementary condition, congruence (55) is equivalent to

$$\sum_{i=1}^{\mathrm{w}-3} c_0[i+1]2^i + 2^h \equiv c_0[\mathrm{w}-1]2^{\mathrm{w}-2} + 2\sum_{i=0}^{h-1}(v[i]\oplus 1)2^i \bmod 2^{\mathrm{w}-1}, \tag{56}$$

moreover in this case the second congruence in (51) gives

$$v[\mathrm{w}-1] \equiv c_0[\mathrm{w}-1] + \gamma_R \bmod 2 \quad \text{and} \quad \delta = v[\mathrm{w}-1]$$

or

$$v[\mathsf{w}-1] \equiv c_0[\mathsf{w}-1] + \gamma_R + 1 \bmod 2 \quad \text{and} \quad \delta = v[\mathsf{w}-1] \oplus 1,$$

i.e., for every solution of (56) we have also two possibilities for $v[\mathsf{w}-1]$. Thus, since (56) can be solved as in the previous case distinguishing between $h = \mathsf{w}-1$, $h = \mathsf{w}-2$, and $1 \leq h \leq \mathsf{w}-3$, the number of solutions is doubled. So, if $c_0[1] = 1$ and $h = \mathsf{w}-1, \mathsf{w}-2$, we find 4 solutions, while, if $1 \leq h \leq \mathsf{w}-3$ and $c_0[i] = j \in \mathbb{F}_2$ for all $i = h+1, \ldots, \mathsf{w}-1$, we have $2^{\mathsf{w}-h}$ solutions.

- $h = 0$: in this last case we have $\overline{v} = r_1(v)$ in the first congruence in (51) so this congruence becomes the equality

$$\sum_{i=0}^{\mathsf{w}-3} c_0[i+1]2^i + c_0[0] = c_0[\mathsf{w}-1]2^{\mathsf{w}-2} + \gamma_L.$$

Since

$$\sum_{i=0}^{\mathsf{w}-3} c_0[i+1]2^i \leq 2^{\mathsf{w}-2} - 1, \tag{57}$$

if $c_0[0] = 0$ we have solutions only when $c_0[i] = 0$ for all $i = 1, \ldots, \mathsf{w}-1$ and $\gamma_L = 0$, with the two possibilities for $v[\mathsf{w}-1]$ given by

$$v[\mathsf{w}-1] \equiv \gamma_R \bmod 2 \quad \text{and} \quad \delta = v[\mathsf{w}-1]$$

or

$$v[\mathsf{w}-1] \equiv \gamma_R + 1 \bmod 2 \quad \text{and} \quad \delta = v[\mathsf{w}-1] \oplus 1,$$

so all $v \in \mathbb{F}_2^{\mathsf{w}}$ are solutions. On the other hand, if $c_0[0] = 1$ and $c_0[\mathsf{w}-1] = 1$, from (57) we also need $c_0[i] = 1$ for all $i = 1, \ldots, \mathsf{w}-2$ and $\gamma_L = 0$, therefore $\delta = v[\mathsf{w}-1]$ and $v[\mathsf{w}-1] \equiv \gamma_R \bmod 2$. Thus only $v[\mathsf{w}-1]$ is fixed and we have $2^{\mathsf{w}-1}$ solutions. Finally, when $c_0[0] = 1$ and $c_0[\mathsf{w}-1] = 0$ we also need $c_0[i] = 0$ for all $i = 1, \ldots, \mathsf{w}-2$ and $\gamma_L = 1$. Therefore $\delta = v[\mathsf{w}-1] \oplus 1$ and $v[\mathsf{w}-1] \equiv \gamma_R \bmod 2$, thus also in this case we have only $v[\mathsf{w}-1]$ fixed and, consequently, there are $2^{\mathsf{w}-1}$ solutions.

□

Thanks to Lemma 1 and to (44), we have nonzero probabilities only if c_0 satisfies one of the following conditions

- $c_0[0] = 1$,
- $c_0[0] = 0$ and $c_0[1] = 1$,
- $c_0[i] = 0$, $i = 0, \ldots, \mathsf{w}-1$, i.e., c_0 is the zero vector in $\mathbb{F}_2^{\mathsf{w}}$,

or, in other words, there are $3 \cdot 2^{\mathsf{w}-2} + 1$ possible choices of c_0 which are about $\frac{3}{4}$ of all the possible constants.

In the following theorem we give the exact values of $P(c_0)$ excluding the only three trivial values of c_0 for which $|I_0| \neq 0$.

Theorem 2. *Let us consider c_0 such that $c_0[i] = j \in \mathbb{F}_2$ for all $i = t+1, \ldots, \mathsf{w}-2$ with $1 \leq t \leq \mathsf{w}-3$ and c_0 different from $(\underbrace{0,0,0,\ldots,0}_{\mathsf{w}-1-zeros},1)$, $(\underbrace{1,1,1,1,\ldots,1}_{\mathsf{w}-ones})$, $(\underbrace{0,0,0,0,\ldots,0}_{\mathsf{w}-zeros})$. Then we have*

$$P(c_0) = \begin{cases} \dfrac{9 + 8 \cdot \delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} \cdot \left(2^{2(\mathsf{w}-2-t)} - 1\right)}{3 \cdot 2^{2\mathsf{w}+1}} & \text{if} \quad c_0[0] = 1 \\ \dfrac{9 + 8 \cdot \delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} \cdot \left(2^{(\mathsf{w}-2-t)} - 1\right)}{3 \cdot 2^{2\mathsf{w}}} & \text{if} \quad c_0[0] = 0,\ c_0[1] = 1, \end{cases} \tag{58}$$

where

$$\delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} = \begin{cases} 1 & \textit{if} \quad c_0[\mathsf{w}-1] = c_0[\mathsf{w}-2] \\ 0 & \textit{if} \quad c_0[\mathsf{w}-1] \neq c_0[\mathsf{w}-2]. \end{cases}$$

Proof. With our choices of c_0 and from the results of Lemma 1, we have $|I_h| = 0$ for $h = 0, 1, \ldots, t-1$, $|I_h| \neq 0$ for $h = t, \ldots, \mathsf{w}-3$ if $c_0[\mathsf{w}-1] = c_0[\mathsf{w}-2]$, and $|I_h| \neq 0$ if $h = \mathsf{w}-1, \mathsf{w}-2$. Thus, from (44) we obtain

$$P(c_0) = \frac{1}{2^{\mathsf{w}+3}} \left(|I_{\mathsf{w}-1}| 2^{-(\mathsf{w}-1)} + |I_{\mathsf{w}-2}| 2^{-(\mathsf{w}-2)} + \delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} \sum_{h=t}^{\mathsf{w}-3} |I_h| \cdot 2^{-h} \right).$$

If $c_0[0] = 1$ from the results of Lemma 1 we find

$$P(c_0) = \frac{1}{2^{\mathsf{w}+3}} \left(2^{-(\mathsf{w}-2)} + 2^{-(\mathsf{w}-3)} + \delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} 2^{\mathsf{w}-1} \sum_{h=t}^{\mathsf{w}-3} 2^{-2h} \right)$$

and, since

$$2^{\mathsf{w}-1} \sum_{h=t}^{\mathsf{w}-3} 2^{-2h} = \frac{4 \cdot \left(2^{2(\mathsf{w}-2-t)} - 1 \right)}{3 \cdot 2^{\mathsf{w}-3}},$$

an easy calculation gives

$$P(c_0) = \frac{9 + 8 \cdot \delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} \cdot \left(2^{2(\mathsf{w}-2-t)} - 1 \right)}{3 \cdot 2^{2\mathsf{w}+1}}.$$

The case $c_0[0] = 0$ and $c_0[1] = 1$ is straightforward, since all the non-zero cardinalities of the sets $|I_h|$ are doubled. □

4. Discussion

We first resume our results: with the choices (18), (23), (24) and (35)

- conditions (21) and (22) hold automatically from (23);
- condition (25) holds automatically from (35);
- condition (26), for w sufficiently large, holds with probability $P(c_0)$ given by (44) and explicitly by (58), under the assumption of independence and uniform distribution for x_1, x_2 and b_3.

Therefore when $\mathsf{r} = 1$ and w is sufficiently large, if c_0 satisfies the hypotheses of Theorem 2 and following our choices for f_1, f_2, f_3, a, b, c and d, with the aforementioned assumptions of independence and uniform distribution for x_1, x_2 and b_3, the equality

$$((y_0 \lll 1) \oplus a, (y_1 \lll 1) \oplus b, (y_2 \lll 1) \oplus c, (y_3 \lll 1) \oplus d) = \mathcal{Q}(c_0, f_1, f_2, f_3)$$

holds with probability $P(c_0)$ given by (58).

4.1. *Propagation Probability of Rotational-XOR Pairs through ChaCha20 Quarter Round*

If we consider the four constants used in the ChaCha20 definition, they all satisfy $c_0[0] = 1$, thus we have

- [expa] = 01100101011110000111000001100001 and $\delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} = 0$, so $P(c_0) = \frac{3}{2^{2\mathsf{w}+1}}$, which gives $P(c_0) = 3 \cdot 2^{-65}$ when $\mathsf{w} = 32$;
- [nd 3] = 01101110011001000010000000110011 and $\delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} = 0$ with $P(c_0)$ as before;
- [2-by] = 00110010001011010110001001111001 and $\delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} = 1$, $t = \mathsf{w}-3$, so $P(c_0) = \frac{11}{2^{2\mathsf{w}+1}}$, which gives $P(c_0) = 11 \cdot 2^{-65}$ when $\mathsf{w} = 32$;

- [te k] $= 01110100011001010010000001101011$ and $\delta_{c_0[\mathsf{w}-1],c_0[\mathsf{w}-2]} = 0$ with $P(c_0)$ as for the first two constants.

A simple calculation shows that the probability of rotational propagation related to the 4 columns of the initial state of ChaCha20 is $297 \cdot 2^{-260} \simeq 2^{-251.7857}$ which is greater than $2^{-32 \cdot 8} = 2^{-256}$.

4.2. ChaCha20 Alternative Constants Giving Non-Negligible Probability

In our scenario, this probability increases for some selections of alternative constants and may represent a weakness for variants of ChaCha20 against rotational attacks. From Theorem 2 we observe that $P(c_0)$ has always a higher value when $c_0[\mathsf{w}-1] = c_0[\mathsf{w}-2]$. In this case we find

$$P(c_0) = \begin{cases} \dfrac{2^{2\mathsf{w}-2t-1}+1}{3 \cdot 2^{2\mathsf{w}+1}} & \text{if} \quad c_0[0] = 1 \\[2ex] \dfrac{2^{2\mathsf{w}-2t-1}+1}{3 \cdot 2^{2\mathsf{w}}} & \text{if} \quad c_0[0] = 0,\ c_0[1] = 1, \end{cases} \tag{59}$$

both of them are greater than $2^{-2\mathsf{w}}$ for all the possible values of $t \in \{1, 2, \ldots, \mathsf{w}-3\}$ and increase their value when t decreases. Thus, constants c_0 satisfying the hypotheses of Theorem 2 and giving a rotational propagation probability for the quarter round greater than 2^{-64} when $\mathsf{w} = 32$ have one of the two following forms

$$c_0 = \begin{cases} (\underbrace{j, j, \ldots, j,}_{\geq 2 \text{ equal bits}} c_0[t], c_0[t-1], \ldots, c_0[2], c_0[1], 1) \\[3ex] (\underbrace{j, j, \ldots, j,}_{\geq 2 \text{ equal bits}} c_0[t], c_0[t-1], \ldots, c_0[2], 1, 0). \end{cases}$$

Finally, we consider the following special values of c_0:

- $c_0 = (\underbrace{0, 0, 0, \ldots, 0,}_{\mathsf{w}-1\text{–zeros}} 1)$ and $c_0 = (\underbrace{1, 1, 1, 1, \ldots, 1}_{\mathsf{w}\text{–ones}})$ using the results from (44) and Lemma 1 we have $P(c_0) = \frac{3}{16}$ for w sufficiently large;
- $c_0 = (\underbrace{0, 0, 0, 0, \ldots, 0}_{\mathsf{w}\text{–zeros}})$ again, from (44) and Lemma 1, we obtain $P(c_0) = \frac{3}{8}$ for w sufficiently large.

5. Conclusions

We considered the ChaCha20 stream cipher and we studied the propagation of rotational-XOR pairs in the quarter round function. Under suitable choices of the inputs and assumptions of independence and uniform distribution, setting the rotational amount r to 1, we established a formula for the probability of the propagation of rotational-XOR pairs depending on the selected constants. For the standard constants in ChaCha20 we find a probability around $2^{-251.7857}$ against the probability of 2^{-256} of a random permutation. Moreover, we were able to find a family of constants which in our scenario potentially facilitate rotational propagation with non-negligible probability.

Author Contributions: Conceptualization, S.B., D.B. and E.B.; methodology, S.B. and E.B.; validation, S.B. and E.B.; formal analysis, S.B. and E.B.; investigation, S.B., D.B. and E.B.; resources, D.B. and E.B.; data curation, S.B. and E.B.; writing—original draft preparation, S.B.; writing—review and editing, S.B. and E.B.; visualization, S.B. and E.B.; supervision D.B. and E.B.; project administration, D.B. and E.B.; funding acquisition, D.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Politecnico di Torino (Funding for basic research).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous referees for their valuable suggestions and corrections which surely improved the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bernstein, D.J. ChaCha, a Variant of Salsa20. In *Workshop Record of SASC*; 2008; Volume 8, pp. 3–5. Available online: https://cr.yp.to/chacha/chacha-20080120.pdf (accessed on 23 May 2022).
2. Bernstein, D.J. Salsa20 Specification. In *Technical Report, eSTREAM Project*; 2005. Available online: http://www.ecrypt.eu.org/stream/salsa20pf.html (accessed on 23 May 2022).
3. Nir, Y.; Langley, A. Chacha20 and poly1305 for IETF protocols. *RFC* **2018**, *8439*, 1–46. [CrossRef]
4. Bernstein, D.J.; Hopwood, D.; Hülsing, A; Lange, T.; Niederhagen, R.; Papachristodoulou, L.; Schneider, M.; Schwabe, P.; Wilcox-O'Hearn, Z. SPHINCS: Practical Stateless Hash–Based Signatures. In *Advances in Cryptology—EUROCRYPT 2015*; LNCS; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9056, pp. 368–397. [CrossRef]
5. Biham, E. New Types of Cryptanalytic Attacks Using Related Keys. *J. Cryptol.* **1994**, *7*, 229–246. [CrossRef]
6. Kelsey, J.; Schneier, B.; Wagner, D. Related-key cryptanalysis of 3-way, biham-DES, CAST, DES-X, newDES, RC2, and TEA. In *Information and Communications Security. ICICS 1997*; LNCS; Springer: Berlin/Heidelberg, Germany, 1997; Volume 1334, pp. 233–246. [CrossRef]
7. Knudsen, L.R.; Matusiewicz, K.; Thomsen, S.S. Observations on the Shabal Keyed Permutation. In Official Comment; 2009. Available online: http://www2.mat.dtu.dk/people/oldusers/S.Thomsen/shabal/shabal.pdf (accessed on 23 May 2022).
8. Bernstein, D.J. Salsa20 Security. In *Technical Report, eSTREAM Project*; 2005. Available online: http://cr.yp.to/snuffle/security.pdf (accessed on 23 May 2022).
9. Standaert, F.-X.; Piret, G.; Gershenfeld, N.; Quisquater, J.-J. Sea: A Scalable Encryption Algorithm for Small Embedded Applications. In *Smart Card Research and Advanced Applications. CARDIS 2006*; LNCS; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3928, pp. 222–236. [CrossRef]
10. Daum, M. Cryptanalysis of Hash Functions of the MD4-Family. Ph.D. Thesis, Ruhr University Bochum, Bochum, Germany, 2005. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.88.7847&rep=rep1&type=pdf (accessed on 23 May 2022).
11. Khovratovich, D.; Nikolić, I. Rotational cryptanalysis of ARX. In *Fast Software Encryption. FSE 2010*; LNCS; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6147, pp. 333–346. [CrossRef]
12. Khovratovich, D.; Nikolić, I.; Rechberger, C. Rotational Rebound Attacks on Reduced Skein. In *Advances in Cryptology—ASIACRYPT 2010*; LNCS; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6477, pp. 1–19. [CrossRef]
13. Ferguson, N.; Lucks, S.; Schneier, B.; Whiting, D.; Bellare, M.; Kohno, T.; Callas, J.; Walker, J. The Skein Hash Function Family. *Submiss. NIST (Round 3)* **2010**, *7*, 3.
14. Khovratovich, D.; Nikolić, I.; Pieprzyk, J.; Sokolowski, P.; Steinfeld, R. Rotational Cryptanalysis of ARX Revisited. In *Fast Software Encryption. FSE 2015*; LNCS; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9054, pp. 519–536. [CrossRef]
15. Guo, J.; Karpman, P.; Nikolić, I.; Wang, L.; Wu, S. Analysis of BLAKE2. In *Topics in Cryptology—CT-RSA 2014*; LNCS; Springer: Cham, Switzerland, 2014; Volume 8366, pp. 402–423. [CrossRef]
16. Morawiecki, P.; Pieprzyk, J.; Srebrny, M. Rotational Cryptanalysis of Round-reduced Keccak. In *Fast Software Encryption: 20th International Workshop*; LNCS; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8424, pp. 241–262. [CrossRef]
17. Ashur, T.; Liu, Y. Rotational Cryptanalysis in the Presence of Constants. *IACR Trans. Symmetric Cryptol.* **2016**, *2016*, 57–70. [CrossRef]
18. Ashur, T.; De Witte, G.; Liu, Y. An Automated Tool for Rotational-XOR Cryptanalysis of ARX-based Primitives. In Proceedings of the 2017 Symposium on Information Theory and Signal Processing in the Benelux (SITB 2017), Delft, The Netherlands, 11–12 May 2017.
19. Ito, R. Rotational Cryptanalysis of Salsa Core Function. In *Information Security. ISC 2020*; LNCS; Springer: Cham, Switzerland, 2020; Volume 12472, pp. 129–145. [CrossRef]
20. Barbero, S.; Bellini, E.; Makarim, R. Rotational Analysis of ChaCha Permutation. *Adv. Math. Commun.* **2021**. [CrossRef]