Improved Linear Approximations to ARX Ciphers and Attacks Against ChaCha

Murilo Coutinho[1] and T. C. Souza Neto[1]

Research and Development Center for
Communications Security (CEPESC), Brazil
Email: murilo.coutinho@redes.unb.br, tsouzaneto@gmail.com

Abstract. In this paper, we present a new technique which can be used to find better linear approximations in ARX ciphers. Using this technique, we present the first explicitly derived linear approximations for 3 and 4 rounds of ChaCha and, as a consequence, it enables us to improve the recent attacks against ChaCha. Additionally, we present new differentials for 3 and 3.5 rounds of ChaCha that, when combined with the proposed technique, lead to further improvement in the complexity of the Differential-Linear attacks against ChaCha.

Keywords: Differential-Linear Cryptanalysis · ARX-Ciphers · ChaCha.

1 Introduction

Symmetric cryptographic primitives are heavily used in a variety of contexts. In particular, ARX-based design is a major building block of modern ciphers due to its efficiency in software. ARX stands for addition, word-wise rotation and XOR. Indeed, ciphers following this framework are composed of those operations and avoid the computation of smaller S-boxes through look-up tables. The ARX-based design approach is used to design stream ciphers (e.g., Salsa20 [7] and ChaCha [6]), efficient block ciphers (e.g., Sparx [16]), cryptographic permutations (e.g., Sparkle [3]) and hash functions (e.g., Blake [2]).

ARX-based designs are not only efficient but provide good security properties. The algebraic degree of ARX ciphers is usually high after only a very few rounds as the carry bit within one modular addition already reaches almost maximal degree. For differential and linear attacks, ARX-based designs show weaknesses for a small number of rounds. However, after some rounds the differential and linear probabilities decrease rapidly. Thus, the probabilities of differentials and the absolute correlations of linear approximations decrease very quickly as we increase the number of rounds. In fact, this property led to the long-trail strategy for designing ARX-based ciphers [16].

Ciphers and primitives based on Salsa20 and ChaCha families are heavily used in practice. In 2005, Bernstein proposed the stream cipher Salsa20 [7] as a contender to the eSTREAM [27], the ECRYPT Stream Cipher Project. As outlined by the author, Salsa20 is an ARX type family of algorithms which can be ran with several number of rounds, including the well known Salsa20/12 and Salsa20/8 versions. Latter, in 2008, Bernstein proposed some modifications to Salsa20 in order to provide better diffusion per round and higher resistance to cryptanalysis. These changes originated a new stream cipher, a variant which he called ChaCha [6]. Although Salsa20 was one of the winners of the eS-TREAM competition, ChaCha has received much more attention through the years. Nowadays, we see the usage of this cipher in several projects and applications.

ChaCha, along with Poly1305 [5], is in one of the cipher suits of the new TLS 1.3 [22], which has been used by Google on both Chrome and Android. Not only has ChaCha been used in TLS but also in many other protocols such as SSH, Noise and S/MIME 4.0. In addition, the RFC 7634 proposes the use of ChaCha in IKE and IPsec. ChaCha has been used not only for encryption, but also as a pseudo-random number generator in any operating system running Linux kernel 4.8 or newer [26, 29]. Additionally, ChaCha has been used in several applications such as WireGuard (VPN) (see [19] for a huge list of applications, protocols and libraries using ChaCha).

Related Work. Since ChaCha is so heavily used, it is very important to understand its security. Indeed, the cryptanalysis of ChaCha is well understood and several authors studied its security [1, 9, 11, 13–15, 17, 18, 20, 23–25, 28, 30] which show weaknesses in the reduced round versions of the cipher.

The cryptanalysis of Salsa20 was introduced by Crowley [11] in 2005. Crowley developed a differential attack against Salsa20/5, namely the 5-round version of Salsa20, and received the $1000 prize offered by Bernstein for the most interesting Salsa20 cryptanalysis in that year. In 2006, Fischer et al. [17] improved the attack against Salsa20/5 and presented their attack against Salsa20/6.

Probably the most important cryptanalysis in this regard was proposed by Aumasson et al. at FSE 2008 [1] with the introduction of Probabilistic Neutral Bits (PNBs), showing attacks against Salsa20/7, Salsa20/8, ChaCha20/6 and ChaCha20/7. After that, several authors proposed small enhancements on the attack of Aumasson et al. The work by Shi et al. [28] introduced the concept of Column Chaining Distinguisher (CCD) to achieve some incremental advancements over [1] for both Salsa and ChaCha.

Maitra, Paul and Meier [23] studied an interesting observation regarding round reversal of Salsa, but no significant cryptanalytic improvement could be obtained using this method. Maitra [24] used a technique of Chosen IVs to obtain certain improvements over existing results. Dey and Sarkar [14] showed how to choose values for the PNB to further improve the attack.

In a paper presented in FSE 2017, Choudhuri and Maitra [9] significantly improved the attacks by considering the mathematical structure of both Salsa and ChaCha in order to find differential characteristics with much higher correlations. Recently, Coutinho and Souza [10] proposed new multi-bit differentials using the mathematical framework of Choudhuri and Maitra. In Crypto 2020, Beierle et al. [4] proposed improvements to the framework of differential-linear cryptanalysis against ARX-based designs and further improved the attacks against ChaCha.

Our Contribution. In this work, we provide a new framework to find linear approximations for ARX ciphers. Using this framework we provide the first explicitly derived linear approximations for 3 and 4 rounds of ChaCha. Exploring these linear approximations, we can improve the attacks for 6 and 7 rounds of ChaCha. Additionally, we present new differentials for 3 and 3.5 rounds of ChaCha . We summarize our findings along with other significant attacks for comparison in Table 1. Also, we verified all theoretical results with random experiments. We provide the source code to reproduce this paper in Github `https://github.com/MurCoutinho/cryptanalysisChaCha.git`, which is, for the best of our knowledge, the first implementation of cryptanalysis against ChaCha available to the public. We should note that it is possible to find attacks with less complexity for related key attacks, but we do not consider them in this work.

Rounds	Time Complexity	Data Complexity	Reference
4	2^{6}	2^{6}	[9]
4.5	2^{12}	2^{12}	[9]
5	2^{16}	2^{16}	[9]
6	2^{139}	2^{30}	[1]
	2^{136}	2^{28}	[28]
	2^{130}	2^{35}	[9]
	$2^{127.5}$	$2^{37.5}$	[9]
	2^{116}	2^{116}	[9]
	$2^{102.2}$	2^{56}	[10]
	$2^{77.4}$	2^{58}	[4]
	2^{75}	2^{75}	[10]
	2^{51}	2^{51}	This work
7	2^{248}	2^{27}	[1]
	$2^{246.5}$	2^{27}	[28]
	$2^{238.9}$	2^{96}	[24]
	$2^{237.7}$	2^{96}	[9]
	$2^{231.9}$	2^{50}	[10]
	$2^{230.86}$	$2^{48.8}$	[4]
	2^{224}	2^{224}	This work

Table 1: The best attacks against ChaCha with 256-bit key.

Organization of the paper. In Section 2, we provide an overview of previous results, including a description of ChaCha, a summary of differential-linear cryptanalysis and a review of the techniques developed by Choudhuri and Maitra in [9]. In Section 3, we present a new technique which can be used to find better linear approximations in ARX ciphers and theoretically develop new linear relations between bits of different rounds for ChaCha. Then, in Section 4, we show that these new linear approximations lead to a better distinguishers for ChaCha reduced to 6 and 7 rounds. Finally, Section 5 presents the conclusion and future work.

2 Specifications and Preliminaries

The main notation we will use throughout the paper is defined in Table 2. Next we define the algorithm ChaCha.

Notation	Description
X	a 4×4 state matrix of ChaCha
$X^{(0)}$	initial state matrix of ChaCha
$X^{(R)}$	state matrix after application of R round functions
Z	output of ChaCha, $Z = X^{(0)} + X^{(R)}$
$x_i^{(R)}$	i^{th} word of the state matrix $X^{(R)}$ (words arranged in row major)
$x_{i,j}^{(R)}$	j^{th} bit of i^{th} word of the state matrix $X^{(R)}$
$x_i^{(R)}[j_0, j_1, ..., j_t]$	the sum $x_{i,j_0}^{(R)} \oplus x_{i,j_1}^{(R)} \oplus \cdots \oplus x_{i,j_t}^{(R)}$
$x + y$	addition of x and y modulo 2^{32}
$x - y$	subtraction of x and y modulo 2^{32}
$x \oplus y$	bitwise XOR of x and y
$x \lll n$	rotation of x by n bits to the left
$x \ggg n$	rotation of x by n bits to the right
Δx	XOR difference of x and x'. $\Delta x = x \oplus x'$
$\Delta X^{(R)}$	XOR difference of $X^{(R)}$ and $X'^{(R)}$. $\Delta X^{(R)} = X^{(R)} \oplus X'^{(R)}$
$\Delta x_i^{(R)}$	differential $\Delta x_i^{(R)} = x_i^{(R)} \oplus x'^{(R)}_i$
$\Delta x_{i,j}^{(R)}$	differential $\Delta x_{i,j}^{(R)} = x_{i,j}^{(R)} \oplus x'^{(R)}_{i,j}$
$\Pr(E)$	probability of occurrence of an event E
$\mathcal{ID}$	input difference
$\mathcal{OD}$	output difference

Table 2: Notation

2.1 ChaCha

The stream cipher Salsa20 was proposed by Bernstein [7] to the *eSTREAM* competition and later Bernstein proposed ChaCha [6] as an improvement of Salsa20. ChaCha consists of a series of ARX (addition, rotation, and XOR) operations on 32-bit words, being highly efficient in software and hardware. Each round of ChaCha has a total of 16 bitwise XOR, 16 addition modulo 2^{32} and 16 constant-distance rotations.

ChaCha operates on a state of 64 bytes, organized as a 4×4 matrix with 32-bit integers, initialized with a 256-bit key $k_0, k_1, ..., k_7$, a 64-bit nonce v_0, v_1 and a 64-bit counter t_0, t_1 (we may also refer to the nonce and counter words as IV words), and 4 constants $c_0 = 0\text{x}61707865$, $c_1 = 0\text{x}3320646e$, $c_2 = 0\text{x}79622d32$ and $c_3 = 0\text{x}6b206574$. For ChaCha, we have the following initial state matrix:

$$X^{(0)} = \begin{pmatrix} x_0^{(0)} & x_1^{(0)} & x_2^{(0)} & x_3^{(0)} \\ x_4^{(0)} & x_5^{(0)} & x_6^{(0)} & x_7^{(0)} \\ x_8^{(0)} & x_9^{(0)} & x_{10}^{(0)} & x_{11}^{(0)} \\ x_{12}^{(0)} & x_{13}^{(0)} & x_{14}^{(0)} & x_{15}^{(0)} \end{pmatrix} = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 \\ k_0 & k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 & k_7 \\ t_0 & t_1 & v_0 & v_1 \end{pmatrix}. \tag{1}$$

The state matrix is modified in each round by a *Quarter Round Function* (QRF), denoted by $QR\left(x_a^{(r-1)}, x_b^{(r-1)}, x_c^{(r-1)}, x_d^{(r-1)}\right)$, which receives and updates 4 integers in the following way:

$$\begin{array}{llll} x_{a'}^{(r-1)} & = x_a^{(r-1)} + x_b^{(r-1)}; & x_{d'}^{(r-1)} & = (x_d^{(r-1)} \oplus x_{a'}^{(r-1)}) \lll 16; \\ x_{c'}^{(r-1)} & = x_c^{(r-1)} + x_{d'}^{(r-1)}; & x_{b'}^{(r-1)} & = (x_b^{(r-1)} \oplus x_{c'}^{(r-1)}) \lll 12; \\ x_a^{(r)} & = x_{a'}^{(r-1)} + x_{b'}^{(r-1)}; & x_d^{(r)} & = (x_{d'}^{(r-1)} \oplus x_a^{(r)}) \lll 8; \\ x_c^{(r)} & = x_{c'}^{(r-1)} + x_d^{(r)}; & x_b^{(r)} & = (x_{b'}^{(r-1)} \oplus x_c^{(r)}) \lll 7; \end{array} \tag{2}$$

One round of ChaCha is defined as 4 applications of the QRF. There is, however, a difference between odd and even rounds. For odd rounds, i.e. $r \in \{1, 3, 5, 7, ...\}$, $X^{(r)}$ is obtained from $X^{(r-1)}$ by applying

$$\begin{array}{l} \left(x_0^{(r)}, x_4^{(r)}, x_8^{(r)}, x_{12}^{(r)}\right) = QR\left(x_0^{(r-1)}, x_4^{(r-1)}, x_8^{(r-1)}, x_{12}^{(r-1)}\right) \\ \left(x_1^{(r)}, x_5^{(r)}, x_9^{(r)}, x_{13}^{(r)}\right) = QR\left(x_1^{(r-1)}, x_5^{(r-1)}, x_9^{(r-1)}, x_{13}^{(r-1)}\right) \\ \left(x_2^{(r)}, x_6^{(r)}, x_{10}^{(r)}, x_{14}^{(r)}\right) = QR\left(x_2^{(r-1)}, x_6^{(r-1)}, x_{10}^{(r-1)}, x_{14}^{(r-1)}\right) \\ \left(x_3^{(r)}, x_7^{(r)}, x_{11}^{(r)}, x_{15}^{(r)}\right) = QR\left(x_3^{(r-1)}, x_7^{(r-1)}, x_{11}^{(r-1)}, x_{15}^{(r-1)}\right) \end{array}.$$

On the other hand, for even rounds, i.e. $r \in \{2, 4, 6, 8, , ...\}$, $X^{(r)}$ is calculated from $X^{(r-1)}$ by applying

$$\begin{array}{l} \left(x_0^{(r)}, x_5^{(r)}, x_{10}^{(r)}, x_{15}^{(r)}\right) = QR\left(x_0^{(r-1)}, x_5^{(r-1)}, x_{10}^{(r-1)}, x_{15}^{(r-1)}\right) \\ \left(x_1^{(r)}, x_6^{(r)}, x_{11}^{(r)}, x_{12}^{(r)}\right) = QR\left(x_1^{(r-1)}, x_6^{(r-1)}, x_{11}^{(r-1)}, x_{12}^{(r-1)}\right) \\ \left(x_2^{(r)}, x_7^{(r)}, x_8^{(r)}, x_{13}^{(r)}\right) = QR\left(x_2^{(r-1)}, x_7^{(r-1)}, x_8^{(r-1)}, x_{13}^{(r-1)}\right) \\ \left(x_3^{(r)}, x_4^{(r)}, x_9^{(r)}, x_{14}^{(r)}\right) = QR\left(x_3^{(r-1)}, x_4^{(r-1)}, x_9^{(r-1)}, x_{14}^{(r-1)}\right) \end{array}.$$

The output of ChaCha20/R is then defined as the sum of the initial state with the state after R rounds $Z = X^{(0)} + X^{(R)}$. One should note that it is possible to parallelize each application of the QRF on each round and also that each round is reversible. Hence, we can compute $X^{(r-1)}$ from $X^{(r)}$. For more information on ChaCha, we refer to [6].

2.2 Differential-Linear Cryptanalysis

In this section, we describe the technique of Differential-Linear cryptanalysis as used to attack ChaCha. Let E be a cipher and suppose we can write $E = E_2 \circ E_1$, where E_1 and E_2 are sub ciphers, covering m and l rounds of the main cipher, respectively. We can apply an input difference $\mathcal{ID}$ $\Delta X^{(0)}$ in the sub cipher E_1 obtaining an output difference $\mathcal{OD}$ $\Delta X^{(m)}$ (see the left side of Fig. 1). The next step is to apply Linear Cryptanalysis to the second sub cipher E_2. Using masks Γ_m and Γ_{out}, we attempt to find good linear approximations covering the remaining l rounds of the cipher E. Applying this technique we can construct a differential-linear distinguisher covering all $m + l$ rounds of the cipher E. This is the main idea in Langford and Hellman's classical approach [21].

Sometimes, however, it can be useful to divide the cipher E into three other ciphers, i.e. $E = E_3 \circ E_2 \circ E_1$. In this scenario, we can explore properties of the cipher in the first part E_1, and then apply a differential linear attack where we divide the differential part of the attack in two (see the right side of Fig. 1). Here, the $\mathcal{OD}$ from the sub cipher E_1 after r rounds, namely $\Delta X^{(r)}$, is the $\mathcal{ID}$ for the sub cipher E_2 which produces an output difference $\Delta X^{(m)}$. For more information in this regard, see [4].

It is important to understand how to compute the complexity of a differential-linear attack. We denote the differential of the state matrix as $\Delta X^{(r)} = X^{(r)} \oplus X'^{(r)}$ and the differential of individual words as $\Delta x_i^{(r)} = x_i^{(r)} \oplus x_i'^{(r)}$. Let $x_{i,j}^{(r)}$ denote the j-th bit of the i-th word of the state matrix after r rounds and let $\mathcal{J}$ be a set of bits. Also, let σ and σ' be linear combinations of bits in the set $\mathcal{J}$

$$\sigma = \left(\bigoplus_{(i,j)\in\mathcal{J}} x_{i,j}^{(r)} \right), \quad \sigma' = \left(\bigoplus_{(i,j)\in\mathcal{J}} x_{i,j}'^{(r)} \right).$$

Then

$$\Delta\sigma = \left(\bigoplus_{(i,j)\in\mathcal{J}} \Delta x_{i,j}^{(r)} \right)$$

is the linear combination of the differentials. We can write

$$\Pr\left[\Delta\sigma = 0 | \Delta X^{(0)}\right] = \frac{1}{2}(1 + \varepsilon_d), \tag{3}$$

where ε_d is the differential correlation.

Using linear cryptanalysis, it is possible to go further and find new relations between the initial state matrix and the state matrix after $R > r$ rounds. To do so, let $\mathcal{L}$ denote another set of bits and define

$$\rho = \left(\bigoplus_{(i,j)\in\mathcal{L}} x_{i,j}^{(R)} \right), \quad \rho' = \left(\bigoplus_{(i,j)\in\mathcal{L}} x_{i,j}'^{(R)} \right).$$

Then, as before,

$$\Delta\rho = \left(\bigoplus_{(i,j)\in\mathcal{L}} \Delta x_{i,j}^{(R)} \right).$$

We can define $\Pr[\sigma = \rho] = \frac{1}{2}(1+\varepsilon_L)$, where ε_L is the linear correlation. We want to find γ such that $\Pr\left[\Delta\rho = 0 | \Delta X^{(0)}\right] = \frac{1}{2}(1+\gamma)$.

To compute γ, we write (to simplify the notation we make the conditional to $\Delta X^{(0)}$ implicit):

$$\begin{aligned}\Pr[\Delta\sigma = \Delta\rho] &= \Pr[\sigma = \rho] \cdot \Pr\left[\sigma' = \rho'\right] + \Pr[\sigma = \bar{\rho}] \cdot \Pr\left[\sigma' = \overline{\rho'}\right] \\ &= \frac{1}{2}\left(1 + \varepsilon_L^2\right).\end{aligned}$$

Then,

$$\begin{aligned}\Pr[\Delta\rho = 0] &= \Pr[\Delta\sigma = 0] \cdot \Pr[\Delta\sigma = \Delta\rho] + \Pr[\Delta\sigma = 1] \cdot \Pr[\Delta\sigma = \overline{\Delta\rho}] \\ &= \frac{1}{2}\left(1 + \varepsilon_d \cdot \varepsilon_L^2\right).\end{aligned}$$

Therefore, the differential-linear correlation is given by $\gamma = \varepsilon_d \cdot \varepsilon_L^2$, which defines a distinguisher with complexity $\mathcal{O}\left(\frac{1}{\varepsilon_d^2 \varepsilon_L^4}\right)$. For further information on differential-linear cryptanalysis we refer to [8].

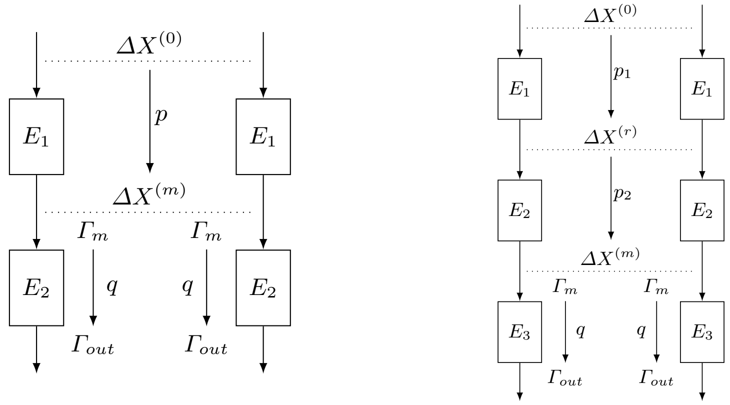

Fig. 1: A classical differential-linear distinguisher (on the left) and a differential-linear distinguisher with experimental evaluation of the correlation p_2 (on the right).

2.3 Multi-bit Differential for Reduced Round ChaCha

In this section, we review the work presented in [9] and in [10]. In these works, the authors developed the theory for selecting specific combination of bits to give high correlations for Chacha. To do that, in both papers the authors analyzed the QRF directly, representing each equation in its bit level. In the following, we change the original notation of the referred papers in order to create a notation that will be better for the purposes of this work.

Thus, let $\Theta(x, y) = x \oplus y \oplus (x + y)$ be the carry function of the sum $x + y$. Define $\Theta_i(x, y)$ as the i-th bit of $\Theta(x, y)$. By definition, we have $\Theta_0(x, y) = 0$. We can write the QRF equations of ChaCha (Eq. 2) as

$$\begin{aligned}
x'^{(m-1)}_{a,i} &= x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{b,i} \oplus \Theta_i(x^{(m-1)}_{a}, x^{(m-1)}_{b}) \\
x'^{(m-1)}_{d,i+16} &= x^{(m-1)}_{d,i} \oplus x'^{(m-1)}_{a,i} \\
x'^{(m-1)}_{c,i} &= x^{(m-1)}_{c,i} \oplus x'^{(m-1)}_{d,i} \oplus \Theta_i(x^{(m-1)}_{c}, x'^{(m-1)}_{d}) \\
x'^{(m-1)}_{b,i+12} &= x^{(m-1)}_{b,i} \oplus x'^{(m-1)}_{c,i} \\
x^{(m)}_{a,i} &= x'^{(m-1)}_{a,i} \oplus x'^{(m-1)}_{b,i} \oplus \Theta_i(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \\
x^{(m)}_{d,i+8} &= x'^{(m-1)}_{d,i} \oplus x^{(m)}_{a,i} \\
x^{(m)}_{c,i} &= x'^{(m-1)}_{c,i} \oplus x^{(m)}_{d,i} \oplus \Theta_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \\
x^{(m)}_{b,i+7} &= x'^{(m-1)}_{b,i} \oplus x^{(m)}_{c,i}
\end{aligned} \tag{4}$$

Inverting these equations, we get:

$$x'^{(m-1)}_{b,i} = x^{(m)}_{b,i+7} \oplus x^{(m)}_{c,i} \tag{5}$$

$$x'^{(m-1)}_{c,i} = x^{(m)}_{c,i} \oplus x^{(m)}_{d,i} \oplus \Theta_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \tag{6}$$

$$x'^{(m-1)}_{d,i} = x^{(m)}_{a,i} \oplus x^{(m)}_{d,i+8} \tag{7}$$

$$x'^{(m-1)}_{a,i} = x^{(m)}_{a,i} \oplus x^{(m)}_{b,i+7} \oplus x^{(m)}_{c,i} \oplus \Theta_i(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \tag{8}$$

$$x^{(m-1)}_{b,i} = \mathcal{L}^{(m)}_{b,i} \oplus \Theta_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \tag{9}$$

$$x^{(m-1)}_{c,i} = \mathcal{L}^{(m)}_{c,i} \oplus \Theta_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus \Theta_i(x^{(m-1)}_{c}, x'^{(m-1)}_{d}) \tag{10}$$

$$x^{(m-1)}_{d,i} = \mathcal{L}^{(m)}_{d,i} \oplus \Theta_i(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \tag{11}$$

$$\begin{aligned}
x^{(m-1)}_{a,i} = \mathcal{L}^{(m)}_{a,i} &\oplus \Theta_i(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \oplus \\
&\Theta_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus \Theta_i(x^{(m-1)}_{a}, x^{(m-1)}_{b})
\end{aligned} \tag{12}$$

where

$$\mathcal{L}^{(m)}_{a,i} = x^{(m)}_{a,i} \oplus x^{(m)}_{b,i+7} \oplus x^{(m)}_{b,i+19} \oplus x^{(m)}_{c,i+12} \oplus x^{(m)}_{d,i} \tag{13}$$

$$\mathcal{L}^{(m)}_{b,i} = x^{(m)}_{b,i+19} \oplus x^{(m)}_{c,i} \oplus x^{(m)}_{c,i+12} \oplus x^{(m)}_{d,i} \tag{14}$$

$$\mathcal{L}^{(m)}_{c,i} = x^{(m)}_{a,i} \oplus x^{(m)}_{c,i} \oplus x^{(m)}_{d,i} \oplus x^{(m)}_{d,i+8} \tag{15}$$

$$\mathcal{L}^{(m)}_{d,i} = x^{(m)}_{a,i} \oplus x^{(m)}_{a,i+16} \oplus x^{(m)}_{b,i+7} \oplus x^{(m)}_{c,i} \oplus x^{(m)}_{d,i+24} \tag{16}$$

Lemma 1. *It holds that* $x^{(m-1)}_{l,0} = \mathcal{L}^{(m)}_{l,0}$, *for* $l \in \{a, b, c, d\}$.

Proof. This result follows directly from Eqs. (9)-(12) by using the fact that $\Theta_0(x, y) = 0$. □

From this equations, we can derive the following lemma:

Lemma 2. *(Lemma 3 of [9]) Let*

$$\begin{aligned}
\Delta A^{(m)} &= \Delta x^{(m)}_{\alpha,0} \oplus \Delta x^{(m)}_{\beta,7} \oplus \Delta x^{(m)}_{\beta,19} \oplus \Delta x^{(m)}_{\gamma,12} \oplus \Delta x^{(m)}_{\delta,0}\\
\Delta B^{(m)} &= \Delta x^{(m)}_{\beta,19} \oplus \Delta x^{(m)}_{\gamma,0} \oplus \Delta x^{(m)}_{\gamma,12} \oplus \Delta x^{(m)}_{\delta,0}\\
\Delta C^{(m)} &= \Delta x^{(m)}_{\delta,0} \oplus \Delta x^{(m)}_{\gamma,0} \oplus \Delta x^{(m)}_{\delta,8} \oplus \Delta x^{(m)}_{\alpha,0}\\
\Delta D^{(m)} &= \Delta x^{(m)}_{\delta,24} \oplus \Delta x^{(m)}_{\alpha,16} \oplus \Delta x^{(m)}_{\alpha,0} \oplus \Delta x^{(m)}_{\gamma,0} \oplus \Delta x^{(m)}_{\beta,7}
\end{aligned}$$

After m rounds of ChaCha, the following holds:

$$\left|\varepsilon_{(A(m))}\right| = \left|\varepsilon_{\left(x^{(m-1)}_{\alpha,0}\right)}\right|, \left|\varepsilon_{\left(B^{(m)}\right)}\right| = \left|\varepsilon_{\left(x^{(m-1)}_{\beta,0}\right)}\right|$$

$$\left|\varepsilon_{\left(C^{(m)}\right)}\right| = \left|\varepsilon_{\left(x^{(m-1)}_{\gamma,0}\right)}\right|, \left|\varepsilon_{\left(D^{(m)}\right)}\right| = \left|\varepsilon_{\left(x^{(m-1)}_{\delta,0}\right)}\right|$$

The tuples $(\alpha, \beta, \gamma, \delta)$ vary depending on whether m is odd or even.

- *Case I. m is odd:*

$$(\alpha, \beta, \gamma, \delta) \in \{(0, 4, 8, 12), (1, 5, 9, 13), (2, 6, 10, 14), (3, 7, 11, 15)\}.$$

- *Case II. m is even:*

$$(\alpha, \beta, \gamma, \delta) \in \{(0, 5, 10, 15), (1, 6, 11, 12), (2, 7, 8, 13), (3, 4, 9, 14)\}.$$

Proof. See [9]. □

Lemma 3. *(Lemma 9 of [9]) For one active input bit in round $m-1$ and multiple active output bits in round m, the following holds for $i > 0$.*

$$\begin{aligned}
x^{(m-1)}_{b,i} &= \mathcal{L}^{(m)}_{b,i} \oplus x^{(m)}_{d,i-1}, && w.p.\ \tfrac{1}{2}\left(1+\tfrac{1}{2}\right)\\
x^{(m-1)}_{a,i} &= \mathcal{L}^{(m)}_{a,i} \oplus x^{(m)}_{b,i+18} \oplus x^{(m)}_{c,i+11} \oplus x^{(m)}_{d,i-2} \oplus x^{(m)}_{d,i+6}, && w.p.\ \tfrac{1}{2}\left(1+\tfrac{1}{2^4}\right)\\
x^{(m-1)}_{c,i} &= \mathcal{L}^{(m)}_{c,i} \oplus x^{(m)}_{a,i-1} \oplus x^{(m)}_{d,i+7} \oplus x^{(m)}_{d,i-1}, && w.p.\ \tfrac{1}{2}\left(1+\tfrac{1}{2^2}\right)\\
x^{(m-1)}_{d,i} &= \mathcal{L}^{(m)}_{d,i} \oplus x^{(m)}_{c,i-1} \oplus x^{(m)}_{b,i+6}, && w.p.\ \tfrac{1}{2}\left(1+\tfrac{1}{2}\right)
\end{aligned}$$

Proof. See [9]. □

Finally, using Lemma 2 and Lemma 3, it is possible to find linear approximations for two rounds of ChaCha.

Lemma 4. *(Lemma 10 of [9]) The following holds with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$*

$$\begin{aligned}
x^{(3)}_{11,0} = x^{(5)}_{0}[0,8,16,24] \oplus x^{(5)}_{1,0} \oplus x^{(5)}_{3,0} \oplus x^{(5)}_{4,7} \oplus x^{(5)}_{4}[14,15] \oplus x^{(5)}_{5}[7,19] \oplus\\
x^{(5)}_{8}[0,7,8] \oplus x^{(5)}_{9,12} \oplus x^{(5)}_{11,0} \oplus x^{(5)}_{12}[0,24] \oplus x^{(5)}_{13,0} \oplus x^{(5)}_{15}[0,8].
\end{aligned}$$

Proof. See [9]. □

Recently, Coutinho and Souza [10] found linear approximations with fewer terms using the same techniques.

Lemma 5. *(Lemma 5 of [10]) When m is odd, each of the following also holds with probability $\frac{1}{2}(1+\frac{1}{2})$*

$$\begin{aligned} x_{0,0}^{(m-2)} \oplus x_{5,0}^{(m-2)} &= x_{0,0}^{(m)} \oplus x_{2,0}^{(m)} \oplus x_{4,7}^{(m)} \oplus x_{4,19}^{(m)} \oplus x_{5,26}^{(m)} \oplus x_{8,12}^{(m)} \oplus x_{9,7}^{(m)} \oplus \\ &\quad x_{9,19}^{(m)} \oplus x_{10,0}^{(m)} \oplus x_{12,0}^{(m)} \oplus x_{13,6}^{(m)} \oplus x_{13,7}^{(m)} \oplus x_{14,0}^{(m)} \oplus x_{14,8}^{(m)} \\ x_{1,0}^{(m-2)} \oplus x_{6,0}^{(m-2)} &= x_{1,0}^{(m)} \oplus x_{3,0}^{(m)} \oplus x_{5,7}^{(m)} \oplus x_{5,19}^{(m)} \oplus x_{6,26}^{(m)} \oplus x_{9,12}^{(m)} \oplus x_{10,7}^{(m)} \oplus \\ &\quad x_{10,19}^{(m)} \oplus x_{11,0}^{(m)} \oplus x_{13,0}^{(m)} \oplus x_{14,6}^{(m)} \oplus x_{14,7}^{(m)} \oplus x_{15,0}^{(m)} \oplus x_{15,8}^{(m)} \\ x_{2,0}^{(m-2)} \oplus x_{7,0}^{(m-2)} &= x_{0,0}^{(m)} \oplus x_{2,0}^{(m)} \oplus x_{6,7}^{(m)} \oplus x_{6,19}^{(m)} \oplus x_{7,26}^{(m)} \oplus x_{8,0}^{(m)} \oplus x_{10,12}^{(m)} \oplus \\ &\quad x_{11,7}^{(m)} \oplus x_{11,19}^{(m)} \oplus x_{12,0}^{(m)} \oplus x_{12,8}^{(m)} \oplus x_{14,0}^{(m)} \oplus x_{15,6}^{(m)} \oplus x_{15,7}^{(m)} \\ x_{3,0}^{(m-2)} \oplus x_{4,0}^{(m-2)} &= x_{1,0}^{(m)} \oplus x_{3,0}^{(m)} \oplus x_{4,26}^{(m)} \oplus x_{7,7}^{(m)} \oplus x_{7,19}^{(m)} \oplus x_{8,7}^{(m)} \oplus x_{8,19}^{(m)} \oplus \\ &\quad x_{9,0}^{(m)} \oplus x_{11,12}^{(m)} \oplus x_{12,6}^{(m)} \oplus x_{12,7}^{(m)} \oplus x_{13,0}^{(m)} \oplus x_{13,8}^{(m)} \oplus x_{15,0}^{(m)} \end{aligned}$$

Proof. See [10]. □

In [9], the authors showed that using as $\mathcal{ID}$ a single bit at $x_{13,13}^{(0)}$ and $\mathcal{OD}$ at $x_{11,0}^{(3)}$, it is possible to obtain $\varepsilon_d = -0.0272 \approx -\frac{1}{2^{5.2}}$, experimentally. And from Lemma 2 it is possible to extend to a 4-round differential-linear correlation with $\varepsilon_L = 1$ when the $\mathcal{OD}$ is $x_{1,0}^{(4)} \oplus x_{11,0}^{(4)} \oplus x_{12,8}^{(4)} \oplus x_{12,0}^{(4)}$. Further, it is possible to extend to a 5-round differential-linear correlation using the last equation from Lemma 4 with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. This gives a total differential-linear 5th round correlation of $\varepsilon_d \cdot \varepsilon_L^2 \approx -0.0068 = -\frac{1}{2^{7.2}}$. This leads to a 5 round distinguisher with complexity approximately 2^{16}.

Extending the linear approximation for 3 rounds comes at a cost. As discussed prior to the above lemma, for ChaCha, setting $i = 0$ in Lemma 2 allows linear approximation of probability 1 for LSB variables. The cost is thus determined by the non LSB variables. A simple count of the non LSB variables in the form (Variable Type, # non LSB occurrence) gives $(x_a, 3)$, $(x_b, 5)$, $(x_c, 3)$, and $(x_d, 2)$. Now, using the probabilities of Lemma 3 and Lemma 4, the linear correlation is $\varepsilon_L = 1/2^{1+3\cdot4+5\cdot1+3\cdot2+2\cdot1} = 2^{-26}$. This leads to a 6 round correlation of $\varepsilon_L^2 \varepsilon_d \approx \frac{1}{2^{57.2}}$. The distinguisher for this correlation has a complexity of 2^{116}.

In [10], the authors used Lemma 5 to derive a distinguisher for 6 rounds. To do that, they found a differential with correlation $\varepsilon_d = 0.00048$ for $(a, b) = (3, 4)$ when the input difference is given by $\Delta x_{14,6}^{(0)} = 1$, and 0 for all remaining bits. Therefore, expanding for 6 rounds from Lemma 5 with weights $4, 1, 2, 1$ for x_a, x_b, x_c and x_d, respectively, they got $\varepsilon_L = 1/2^{1+0\cdot4+3\cdot1+3\cdot2+3\cdot1} = 2^{-13}$. Then we have $\varepsilon_d \varepsilon_L^2 \approx 2^{-37.02}$, which leads to an attack against 6 rounds of ChaCha with complexity 2^{75}. This is the currently best known 6 round attack on ChaCha.

3 Improved Linear Approximations for ARX-Based Ciphers

The challenge of finding good linear approximations in ARX-based designs comes from the addition operation which is responsible for the non-linearity of the design. In 2003, Wallén [31] published a very important paper where a mathematical framework for finding linear approximations of addition modulo 2^n was developed. Since then, several authors used these technique to find linear approximations in ARX-based designs [9].

Therefore, as before, let $\Theta(x, y) = x \oplus y \oplus (x + y)$ be the carry function of the sum $x + y$. Define $\Theta_i(x, y)$ as the i-th bit of $\Theta(x, y)$. By definition, we have $\Theta_0(x, y) = 0$. Using Theorem 3 of [31], we can generate all possible linear approximations with a given correlation. In particular, we will use the following linear approximations:

$$\Pr(\Theta_i(x, y) = y_{i-1}) = \frac{1}{2}\left(1 + \frac{1}{2}\right), i > 0. \tag{17}$$

$$\Pr(\Theta_i(x, y) \oplus \Theta_{i-1}(x, y) = 0) = \frac{1}{2}\left(1 + \frac{1}{2}\right), i > 0. \tag{18}$$

In previous works of cryptanalysis of ARX ciphers, authors concentrated in finding approximations for particular bits in one round and then repeating the same equations to expand the linear approximation to further rounds (see [9] and [10] for some examples). However, by combining Eqs. 17 and 18 when attacking ARX ciphers we can create a strategy to improve linear approximations when considering more rounds. The main idea is that when using Eq. 17 in one round we will create consecutive terms that can be expanded together using Eq. 18.

For example, consider the sum $z = x + y$. If we want a linear approximation for the bit z_7, we can use Eq. 17 to obtain $z_7 = x_7 \oplus y_7 \oplus \Theta_7(x, y) = x_7 \oplus y_7 \oplus y_6$ with probability 0.75. Since the XOR operation will not change the indexes and the rotation will probably keep y_6 and y_7 adjacent, we can use Eq. 18 in the subsequent round to cancel out the non-linear terms rather than expanding them, leading to a linear equation with higher correlation and fewer terms to be expanded further. Next, we will use this technique to find new linear approximations for ChaCha.

3.1 Linear Approximations for the Quarter Round Function

The first step is to find linear approximations for the QRF of ChaCha. Of course, we already know some of them from previous works (Section 2.3). However, here we will consider adjacent bits and several other combinations that cancel out non-linear terms or use Eq. (18). At first glance, these results may seem innocuous, but latter they will prove themselves useful when deriving linear approximations for multiple rounds of ChaCha.

We start with a better linear approximation for $x_{a,i}^{(m-1)}$.

Lemma 6. *The following holds for $i > 0$*

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus x_{b,i+6}^{(m)} \oplus x_{b,i+18}^{(m)} \oplus x_{c,i+11}^{(m)} \oplus x_{d,i-1}^{(m)}, \ w.p.\ \tfrac{1}{2}\left(1+\tfrac{1}{2^3}\right).$$

Proof. From Eq. (12) we have

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \Theta_i(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_i(x_a^{(m-1)}, x_b^{(m-1)}).$$

Using Eq. (17) and the Piling-up Lemma we can write

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus x_{b,i-1}'^{(m-1)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus x_{b,i-1}^{(m-1)},$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. Using Eq. (9) we get

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus x_{b,i-1}'^{(m-1)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathcal{L}_{b,i-1}^{(m)} \oplus \Theta_{i-1}(x_c'^{(m-1)}, x_d^{(m)}).$$

Using the approximation of Eq. (18) and the Piling-up Lemma we can write

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus x_{b,i-1}'^{(m-1)} \oplus \mathcal{L}_{b,i-1}^{(m)},$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. Finally, using Eqs. (5) and (14) we get

$$x_{a,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus x_{b,i+6}^{(m)} \oplus x_{b,i+18}^{(m)} \oplus x_{c,i+11}^{(m)} \oplus x_{d,i-1}^{(m)},$$

which completes the proof. □

Lemma 7. *For two active input bits in round $m-1$ and multiple active output bits in round m, the following holds for $i > 0$*

$$x_{\lambda,i}^{(m-1)} \oplus x_{\lambda,i-1}^{(m-1)} = \mathcal{L}_{\lambda,i}^{(m)} \oplus \mathcal{L}_{\lambda,i-1}^{(m)}, \ \ w.p.\ \frac{1}{2}\left(1+\frac{1}{2^\sigma}\right),$$

where $(\lambda, \sigma) \in \{(a,3),(b,1),(c,2),(d,1)\}$.

Proof. This proof follows directly from Eqs. (9)-(12) using the approximation of Eq. (18) and the Piling-up Lemma. □

Lemma 8. *Suppose that $(\lambda, \sigma) \in \{(i, i-2), (i-1, i-1)\}$, $i > 1$. Then for three active input bits in round $m-1$ and multiple active output bits in round m, the following holds*

$$x_{b,\lambda}^{(m-1)} \oplus x_{c,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = \mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus x_{d,\sigma}^{(m)}, \ \ w.p.\ \frac{1}{2}\left(1+\frac{1}{2^2}\right).$$

Proof. Using Eq. (9) and Eq. (10) we get

$$\begin{aligned} x_{b,\lambda}^{(m-1)} \oplus x_{c,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = {} & \mathcal{L}_{b,\lambda}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \Theta_\lambda(x_c'^{(m-1)}, x_d^{(m)}) \oplus \\ & \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}) \oplus \\ & \Theta_{i-1}(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}). \end{aligned}$$

Canceling out common factors and using the approximation of Eq. (18) we can write

$$x_{b,\lambda}^{(m-1)} \oplus x_{c,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \Theta_{\sigma+1}(x_c'^{(m-1)}, x_d^{(m)}).$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. Using Eq. (17) we get

$$x_{b,\lambda}^{(m-1)} \oplus x_{c,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus x_{d,\sigma}^{(m)},$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. □

Lemma 9. *For multiple active input bits in round $m-1$ and multiple active output bits in round m, the following linear approximations hold for ChaCha with probability $\frac{1}{2}\left(1+\frac{1}{2^k}\right)$:*

$$x_{b,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus x_{a,i-1}^{(m)} \oplus x_{d,i+7}^{(m)} \qquad k=1, i>0 \tag{19}$$

$$x_{a,i}^{(m-1)} \oplus x_{b,i}^{(m-1)} = \begin{array}{l} \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \\ x_{b,i+6}^{(m)} \oplus x_{c,i-1}^{(m)} \oplus x_{d,i-2}^{(m)} \end{array} \qquad k=3, i>1 \tag{20}$$

$$x_{a,1}^{(m-1)} \oplus x_{b,1}^{(m-1)} = \mathcal{L}_{a,1}^{(m)} \oplus \mathcal{L}_{b,0}^{(m)} \oplus \mathcal{L}_{b,1}^{(m)} \oplus x_{b,7}^{(m)} \oplus x_{c,0}^{(m)} \qquad k=2 \tag{21}$$

$$x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \begin{array}{l} \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus x_{a,i-1}^{(m)} \oplus \\ x_{b,i+6}^{(m)} \oplus x_{c,i-1}^{(m)} \oplus x_{d,i-2}^{(m)} \oplus x_{d,i+7}^{(m)} \end{array} \qquad k=4, i>1 \tag{22}$$

$$x_{a,1}^{(m-1)} \oplus x_{c,1}^{(m-1)} = \begin{array}{l} \mathcal{L}_{a,1}^{(m)} \oplus \mathcal{L}_{b,0}^{(m)} \oplus \mathcal{L}_{c,1}^{(m)} \oplus x_{a,0}^{(m)} \oplus \\ x_{b,7}^{(m)} \oplus x_{c,0}^{(m)} \oplus x_{d,8}^{(m)} \end{array} \qquad k=3 \tag{23}$$

$$x_{a,i}^{(m-1)} \oplus x_{d,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus \mathcal{L}_{b,i-1}^{(m)} \qquad k=2, i>1 \tag{24}$$

$$\begin{array}{c} x_{a,i-1}^{(m-1)} \oplus \\ x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} \end{array} = \begin{array}{l} \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \\ x_{d,i-2}^{(m)} \oplus x_{a,i-1}^{(m)} \oplus x_{d,i+7}^{(m)} \end{array} \qquad k=4, i>1 \tag{25}$$

$$\begin{array}{c} x_{a,i}^{(m-1)} \oplus \\ x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \end{array} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus x_{d,i-2}^{(m)}, \qquad k=3, i>1 \tag{26}$$

$$\begin{array}{c} x_{b,i-1}^{(m-1)} \oplus \\ x_{a,i}^{(m-1)} \oplus x_{d,i}^{(m-1)} \end{array} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus x_{d,i-1}^{(m)}, \qquad k=2, i>1 \tag{27}$$

$$\begin{array}{c} x_{b,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus \\ x_{c,i-1}^{(m-1)} \oplus x_{c,i}^{(m-1)} \end{array} = \begin{array}{l} \mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \\ \mathcal{L}_{c,i-1}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)}, \end{array} \qquad k=1, i>1 \tag{28}$$

$$\begin{array}{c} x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus \\ x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \end{array} = \begin{array}{l} \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \\ \mathcal{L}_{c,i-1}^{(m)} \oplus x_{a,i-2}^{(m)} \oplus x_{d,i+6}^{(m)}, \end{array} \qquad k=3, i>1 \tag{29}$$

$$\begin{array}{c} x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus \\ x_{c,i-1}^{(m-1)} \oplus x_{d,i}^{(m-1)} \oplus \\ x_{d,i-1}^{(m-1)} \end{array} = \begin{array}{l} \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ \mathcal{L}_{d,i-1}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus x_{d,i-1}^{(m)} \oplus \\ x_{a,i-2}^{(m)} \oplus x_{d,i+6}^{(m)}, \end{array} \qquad k=3, i>2 \tag{30}$$

Proof. The proof for each equation follows the same basic steps: (1) cancel common factors; (2) cancel adjacent non-linear terms using Eq. (18), updating the probability using the Piling-Up Lemma; (3) substitute the remaining non-linear terms using Eq. (17), updating the probability using the Piling-Up Lemma. For completeness, we list all proofs in Appendix 1. □

3.2 Linear Approximations for Multiple Rounds of ChaCha

In this section, we use the proposed technique to construct several new linear approximations for the stream cipher ChaCha which will prove useful to construct better distinguishers. We developed a program (available in `https://github.com/MurCoutinho/cryptanalysisChaCha.git`) that makes the process of finding linear approximations partly automatic. Our program is capable of expanding the equations and, after statistically verifying the correlation, it outputs the resulting linear approximation in LaTeXcode.

We start using the result of Coutinho and Souza [10]. We will only consider the equation for $x^{(3)}_{3,0} \oplus x^{(3)}_{4,0}$ of Lemma 5 but the same reasoning could be applied to any other equation in that lemma. Then, we have

$$\begin{aligned} x^{(3)}_{3,0} \oplus x^{(3)}_{4,0} = x^{(5)}_{1,0} \oplus x^{(5)}_{3,0} \oplus x^{(5)}_{4,26} \oplus x^{(5)}_{7,7} \oplus x^{(5)}_{7,19} \oplus x^{(5)}_{8,7} \oplus x^{(5)}_{8,19} \oplus \\ x^{(5)}_{9,0} \oplus x^{(5)}_{11,12} \oplus x^{(5)}_{12,6} \oplus x^{(5)}_{12,7} \oplus x^{(5)}_{13,0} \oplus x^{(5)}_{13,8} \oplus x^{(5)}_{15,0} \end{aligned} \tag{31}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$.

As presented in Section 2.3, to expand the equation to the 6-th round, we could use only Lemma 3 as proposed in [9]. In this case, we have weights $4, 1, 2, 1$ for x_a, x_b, x_c and x_d, respectively, and a count of $(x_a, 0)$, $(x_b, 3)$, $(x_c, 3)$ e $(x_d, 3)$. Thus, the linear correlation is $\varepsilon_L = 1/2^{1+0\cdot 4+3\cdot 1+3\cdot 2+3\cdot 1} = 2^{-13}$. However, we can do better with the new technique proposed in Section 3. This will lead us to the following lemma

Lemma 10. *The following linear approximation holds with probability* $\frac{1}{2}\left(1+\frac{1}{2^8}\right)$

$$\begin{aligned} x^{(3)}_{3,0} \oplus x^{(3)}_{4,0} = {} & x^{(6)}_{0}[0,16] \oplus x^{(6)}_{1}[0,6,7,11,12,22,23] \oplus x^{(6)}_{2}[0,6,7,8,16,18, \\ & 19,24] \oplus x^{(6)}_{4}[7,13,19] \oplus x^{(6)}_{5}[7] \oplus x^{(6)}_{6}[7,13,14,19] \oplus \\ & x^{(6)}_{7}[6,7,14,15,26] \oplus x^{(6)}_{8}[0,7,8,19,31] \oplus x^{(6)}_{9}[0,6,12,26] \oplus \\ & x^{(6)}_{10}[0] \oplus x^{(6)}_{11}[6,7] \oplus x^{(6)}_{12}[0,11,12,19,20,30,31] \oplus \\ & x^{(6)}_{13}[0,14,15,24,26,27] \oplus x^{(6)}_{14}[8,25,26] \oplus x^{(6)}_{15}[24]. \end{aligned}$$

Proof. First, from Eq. (31) we can use Lemma 1 to replace $x^{(5)}_{1,0}$, $x^{(5)}_{3,0}$, $x^{(5)}_{9,0}$, $x^{(5)}_{13,0}$, $x^{(5)}_{15,0}$ by $\mathcal{L}^{(6)}_{1,0}$, $\mathcal{L}^{(6)}_{3,0}$, $\mathcal{L}^{(6)}_{9,0}$, $\mathcal{L}^{(6)}_{13,0}$, $\mathcal{L}^{(6)}_{15,0}$ with probability 1. Next, note that, since we are transitioning from round 5 to 6, we have

$$(a, b, c, d) \in \{(0, 5, 10, 15), (1, 6, 11, 12), (2, 7, 8, 13), (3, 4, 9, 14)\}.$$

We have already considered the case $(a, b, c, d) = (0, 5, 10, 15)$. Then we still have 3 cases left to consider.

- **Case 1:** When $(a, b, c, d) = (1, 6, 11, 12)$, we have the factors $x^{(5)}_{11,12}$, $x^{(5)}_{12,6}$, $x^{(5)}_{12,7}$. Then we can use Lemma 3 and Lemma 7 in order to get

  $$\Pr\left(x^{(5)}_{11,12} = \mathcal{L}^{(6)}_{11,12} \oplus x^{(6)}_{1,11} \oplus x^{(6)}_{12,19} \oplus x^{(6)}_{12,11}\right) = \frac{1}{2}\left(1 + \frac{1}{2^2}\right)$$

  and

  $$\Pr\left(x^{(5)}_{12,7} \oplus x^{(5)}_{12,6} = \mathcal{L}^{(6)}_{12,7} \oplus \mathcal{L}^{(6)}_{12,6}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

- **Case 2:** If $(a, b, c, d) = (2, 7, 8, 13)$, we have the factors $x^{(5)}_{7,7}$, $x^{(5)}_{7,19}$, $x^{(5)}_{8,7}$, $x^{(5)}_{8,19}$, $x^{(5)}_{13,8}$ and we can use Lemma 3 and Eq. (19) of Lemma 9 to get

  $$\Pr\left(x^{(5)}_{13,8} = \mathcal{L}^{(6)}_{13,8} \oplus x^{(6)}_{8,7} \oplus x^{(6)}_{7,14}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right),$$

  $$\Pr\left(x^{(5)}_{7,7} \oplus x^{(5)}_{8,7} = \mathcal{L}^{(6)}_{7,7} \oplus \mathcal{L}^{(6)}_{8,7} \oplus x^{(6)}_{2,6} \oplus x^{(6)}_{13,14}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right),$$

  $$\Pr\left(x^{(5)}_{7,19} \oplus x^{(5)}_{8,19} = \mathcal{L}^{(6)}_{7,19} \oplus \mathcal{L}^{(6)}_{8,19} \oplus x^{(6)}_{2,18} \oplus x^{(6)}_{13,26}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

- **Case 3:** When considering $(a, b, c, d) = (3, 4, 9, 14)$, we have $x^{(5)}_{4,26}$ and we can use Lemma 3 to obtain

  $$\Pr\left(x^{(5)}_{4,26} = \mathcal{L}^{(6)}_{4,26} \oplus x^{(6)}_{14,26}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

  By the Piling-up Lemma, we have that all these changes result in a probability of $\frac{1}{2}\left(1 + \frac{1}{2^8}\right)$. Expanding the linear terms using Eqs. (13)-(16) and canceling out common factors completes the proof. □

Computational Result 1 *The linear approximation of Lemma 10 holds computationally with* $\varepsilon_{L_0} = 0.006942 \approx 2^{-7.17}$. *This correlation was verified using* 2^{38} *random samples.*

In [9], the authors remarked that an expansion of this method to 7 rounds would be unlikely to be useful. Indeed, if we only apply Lemma 3 (which are the linear approximations proposed by Choudhuri and Maitra) we would have $(x_a, 14)$, $(x_b, 13)$, $(x_c, 9)$, $(x_d, 15)$. Therefore, the aggregated correlation would be $\varepsilon_L = 1/2^{7+14\cdot4+13\cdot1+9\cdot2+15\cdot1} = 2^{-109}$. Thus, using this linear expansion in a differential-linear attack would lead to a distinguisher with complexity no less then 2^{436}. However, using our new linear approximations we can get a much better result.

Lemma 11. *The following linear approximation holds with probability* $\frac{1}{2}\left(1+\frac{1}{2^{55}}\right)$

$$\begin{aligned} x_{3,0}^{(3)} \oplus x_{4,0}^{(3)} = {} & x_0^{(7)}[0,3,4,7,8,11,12,14,15,18,20,27,28] \oplus x_1^{(7)}[0,5,7,8,10,\\ & 11,14,15,16,22,23,24,25,27,30,31] \oplus x_2^{(7)}[6,7,9,10,16,18,19,\\ & 25,26] \oplus x_3^{(7)}[6,7,8,24] \oplus x_4^{(7)}[0,2,3,5,18,22,23,27] \oplus x_5^{(7)}[1,2,\\ & 9,10,13,14,18,21,22,25,29,30] \oplus x_6^{(7)}[2,3,5,7,10,11,13,14,19,\\ & 22,23,27,30,31] \oplus x_7^{(7)}[1,2,13,25,26,30,31] \oplus x_8^{(7)}[8,11,13,20,\\ & 25,27,28,30,31] \oplus x_9^{(7)}[2,3,6,7,14,15,18,27] \oplus x_{10}^{(7)}[0,3,4,8,12,\\ & 13,14,18,20,27,28,30] \oplus x_{11}^{(7)}[6,14,15,18,19,23,24,27] \oplus \\ & x_{12}^{(7)}[3,4,6,11,13,22,23,24,26,27,30,31] \oplus x_{13}^{(7)}[1,2,6,7,8,10,\\ & 11,13,14,16,18,20,22,23,24,25,26] \oplus x_{14}^{(7)}[0,6,13,14,15,16,\\ & 23,24] \oplus x_{15}^{(7)}[16,25,26]. \end{aligned}$$

Proof. If we start from Lemma 10 then we want to expand the equation one more round. To do so, first note that since we are transitioning from round 6 to 7, we have $(a,b,c,d) \in \{(0,4,8,12),(1,5,9,13),(2,6,10,14),(3,7,11,15)\}$. Therefore, we can divide the factors of the equation in 4 distinct groups:

- Group I - $x_0^{(6)}[0,16], x_4^{(6)}[7,13,19], x_8^{(6)}[0,7,8,19,31]$, $x_{12}^{(6)}[0,11,12,19,20,30,31]$.
- Group II - $x_1^{(6)}[0,6,7,11,12,22,23]$, $x_5^{(6)}[7]$, $x_9^{(6)}[0,6,12,26]$, $x_{13}^{(6)}[0,14,15,24,26,27]$.
- Group III - $x_2^{(6)}[0,6,7,8,16,18,19,24], x_6^{(6)}[7,13,14,19]$, $x_{10}^{(6)}[0]$, $x_{14}^{(6)}[8,25,26]$.
- Group IV - $x_7^{(6)}[6,7,14,15,26]$, $x_{11}^{(6)}[6,7], x_{15}^{(6)}[24]$.

The procedure to expand and compute the correlation is similar to that in the proof of Lemma 10. To simplify the notation we will compute the probability given by the Piling-up Lemma by summing values k where the probability of a particular linear equation will be given by $\frac{1}{2}\left(1+\frac{1}{2^k}\right)$.

In Group I, the factors $x_{0,0}^{(6)}, x_{8,0}^{(6)}, x_{12,0}^{(6)}$ can be expanded using Lemma 1 with probability 1. Next, we can combine the following factors: $x_{4,7}^{(6)}, x_{8,7}^{(6)}, x_{8,8}^{(6)}$ using Lemma 8 ($k=2$); $x_{4,19}^{(6)}, x_{8,19}^{(6)}$ using Eq. (19) of Lemma 9 ($k=1$); $x_{12,11}^{(6)}, x_{12,12}^{(6)}$ using Lemma 7 with ($k=1$); $x_{12,19}^{(6)}, x_{12,20}^{(6)}$ using Lemma 7 with ($k=1$); $x_{12,30}^{(6)}, x_{12,31}^{(6)}$ using Lemma 7 with ($k=1$). Finally, it remains some single terms to be expanded: $x_{0,16}^{(6)}$ using Lemma 6 ($k=3$); $x_{4,13}^{(6)}$ using Lemma 3 ($k=1$); $x_{8,31}^{(6)}$ using Lemma 3 ($k=2$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$\begin{aligned} & x_0^{(6)}[0,16] \oplus x_4^{(6)}[7,13,19] \oplus x_8^{(6)}[0,7,8,19,31] \oplus x_{12}^{(6)}[0,11,12,19,20,30,31] = \\ & x_0^{(7)}[0,3,4,7,8,11,12,14,15,18,20,27,28] \oplus x_4^{(7)}[0,2,3,5,18,22,23,27] \oplus \\ & x_8^{(7)}[8,11,13,20,25,27,28,30,31] \oplus x_{12}^{(7)}[3,4,6,11,13,22,23,24,26,27,30,31] \end{aligned} \tag{32}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^{12}}\right)$.

In Group II, the factors $x_{1,0}^{(6)}, x_{9,0}^{(6)}, x_{13,0}^{(6)}$ can be expanded using Lemma 1 with probability 1. Next, we can combine the following factors: $x_{1,6}^{(6)}, x_{1,7}^{(6)}, x_{5,7}^{(6)}, x_{9,6}^{(6)}$ using Eq. (29) of Lemma 9 ($k=3$); $x_{1,11}^{(6)}, x_{1,12}^{(6)}, x_{9,12}^{(6)}$ using Eq. (25) of Lemma 9 ($k=4$); $x_{1,22}^{(6)}, x_{1,23}^{(6)}$ using Lemma 7 ($k=3$); $x_{13,14}^{(6)}, x_{13,15}^{(6)}$ using Lemma 7 ($k=1$); $x_{13,26}^{(6)}, x_{13,27}^{(6)}$ using Lemma 7 ($k=1$). Finally, it remains some single terms to be expanded: $x_{9,26}^{(6)}$ using Lemma 3 ($k=2$); $x_{13,24}^{(6)}$ using Lemma 3 ($k=1$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$x_1^{(6)}[0,6,7,11,12,22,23] \oplus x_5^{(6)}[7] \oplus x_9^{(6)}[0,6,12,26] \oplus x_{13}^{(6)}[0,14,15,24,26, 27] = x_1^{(7)}[0,5,7,8,10,11,14,15,16,22,23,24,25,27,30,31] \oplus x_5^{(7)}[1,2,9,10, 13,14,18,21,22,25,29,30] \oplus x_9^{(7)}[2,3,6,7,14,15,18,27] \oplus x_{13}^{(7)}[1,2,6,7,8,10, 11,13,14,16,18,20,22,23,24,25,26] \tag{33}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^{15}}\right)$.

In Group III, the factors $x_{2,0}^{(6)}$ and $x_{10,0}^{(6)}$ can be expanded using Lemma 1 with probability 1. Next, we can combine the following factors: $x_{2,6}^{(6)}, x_{2,7}^{(6)}$ using Lemma 7 ($k=3$); $x_{6,13}^{(6)}, x_{6,14}^{(6)}$ using Lemma 7 ($k=1$); $x_{14,25}^{(6)}, x_{14,26}^{(6)}$ using Lemma 7 ($k=1$); $x_{2,18}^{(6)}, x_{2,19}^{(6)}, x_{6,19}^{(6)}$ using Eq. (26) of Lemma 9 ($k=3$); $x_{2,8}^{(6)}, x_{6,7}^{(6)}, x_{14,8}^{(6)}$ using Eq. (27) of Lemma 9 ($k=2$). Finally, it remains some single terms to be expanded: $x_{2,16}^{(6)}$ using Lemma 6 ($k=3$); $x_{2,24}^{(6)}$ using Lemma 6 ($k=3$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$x_2^{(6)}[0,6,7,8,16,18,19,24] \oplus x_6^{(6)}[7,13,14,19] \oplus x_{10}^{(6)}[0] \oplus x_{14}^{(6)}[8,25,26] = x_2^{(7)}[6,7,9,10,16,18,19,25,26] \oplus x_6^{(7)}[2,3,5,7,10,11,13,14,19,22,23,27,30, 31] \oplus x_{10}^{(7)}[0,3,4,8,12,13,14,18,20,27,28,30] \oplus x_{14}^{(7)}[0,6,13,14,15,16,23,24] \tag{34}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^{16}}\right)$.

In Group IV, we can combine the following factors: $x_{7,14}^{(6)}, x_{7,15}^{(6)}$ using Lemma 7 ($k=1$); $x_{7,6}^{(6)}, x_{7,7}^{(6)}, x_{11,6}^{(6)}, x_{11,7}^{(6)}$ using Eq. (28) of Lemma 9 ($k=1$). It remains some single terms to be expanded: $x_{7,26}^{(6)}$ using Lemma 3 ($k=1$); $x_{15,24}^{(6)}$ using Lemma 3 ($k=1$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$x_7^{(6)}[6,7,14,15,26] \oplus x_{11}^{(6)}[6,7] \oplus x_{15}^{(6)}[24] = x_3^{(7)}[6,7,8,24] \oplus x_7^{(7)}[1,2, 13,25,26,30,31] \oplus x_{11}^{(7)}[6,14,15,18,19,23,24,27] \oplus x_{15}^{(7)}[16,25,26] \tag{35}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^{4}}\right)$.

Finally, using the Piling-up Lemma we can combine the results from Lemma 10 and Eqs. (32)-(35), which leads to a correlation of $\varepsilon_L = 1/2^{8+12+15+16+4} = 2^{-55}$. $\square$

Computational Result 2 *The linear approximation of Eq.* (32) *holds computationally with* $\varepsilon_{L_1} = 0.000301 \approx 2^{-11.70}$*. This correlation was verified using* 2^{42} *random samples.*

Computational Result 3 *The linear approximation of Eq.* (33) *holds computationally with* $\varepsilon_{L_2} = 0.000100 \approx 2^{-13.29}$*. This correlation was verified using* 2^{42} *random samples.*

Computational Result 4 *The linear approximation of Eq.* (34) *holds computationally with* $\varepsilon_{L_3} = 0.000051 \approx 2^{-14.26}$*. This correlation was verified using* 2^{42} *random samples.*

Computational Result 5 *The linear approximation of Eq.* (35) *holds computationally with* $\varepsilon_{L_4} = 0.0625 \approx 2^{-4}$*. This correlation was verified using* 2^{38} *random samples.*

Next, we will only work with a linear approximation for the bit $x_{5,0}^{(3.5)}$ (as introduced in [9], half a round of ChaCha consists in applying half the operations of the QRF. Thus, from Eq. (2) we can write $x_a^{(r-1/2)} = x_{a\prime}^{(r-1)}, \ldots\ x_d^{(r-1/2)} = x_{d\prime}^{(r-1)}$). Using Eq. (5) it is easy to see that we have $x_{5,0}^{(3.5)} = x_{5,7}^{(4)} \oplus x_{10,0}^{(4)}$. Additionally, using Lemma 3 we can expand one more round and we get

$$x_{5,0}^{(3.5)} = x_{2,0}^{(5)} \oplus x_{5,26}^{(5)} \oplus x_{9,7}^{(5)} \oplus x_{9,19}^{(5)} \oplus x_{10,0}^{(5)} \oplus x_{13,6}^{(5)} \oplus x_{13,7}^{(5)} \oplus x_{14,0}^{(5)} \oplus x_{14,8}^{(5)}, \quad (36)$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2}\right)$.

Lemma 12. *The following linear approximation holds with probability* $\frac{1}{2}\left(1 + \frac{1}{2^8}\right)$

$$\begin{aligned} x_{5,0}^{(3.5)} = {} & x_0^{(6)}[0] \oplus x_2^{(6)}[0,6,7,22,23] \oplus x_3^{(6)}[0,6,7,8,16,18,19,24] \oplus \\ & x_4^{(6)}[7,14,15] \oplus x_5^{(6)}[13] \oplus x_7^{(6)}[7,13,14,19] \oplus x_8^{(6)}[6,7,12] \oplus \\ & x_9^{(6)}[0,8,19] \oplus x_{10}^{(6)}[0,6,26] \oplus x_{13}^{(6)}[0,30,31] \oplus \\ & x_{14}^{(6)}[0,6,7,14,15,18,19,24,26,27] \oplus x_{15}^{(6)}[0,8,25,26]. \end{aligned}$$

Proof. We start from Eq. (36) and we can use Lemma 1 $x_{2,0}^{(5)}, x_{10,0}^{(5)}, x_{14,0}^{(5)}$ by $\mathcal{L}_{2,0}^{(6)}, \mathcal{L}_{10,0}^{(6)}, \mathcal{L}_{14,0}^{(6)}$ with probability 1. Next, note that, since we are transitioning from round 5 to 6, we have $(a,b,c,d) \in \{(0,5,10,15), (1,6,11,12), (2,7,8,13), (3,4,9,14)\}$. Considering $(a,b,c,d) = (0,5,10,15)$, we have the factor $x_{5,26}^{(5)}$ and we can apply Lemma 3 to get

$$\Pr\left(x_{5,26}^{(5)} = \mathcal{L}_{5,26}^{(6)} \oplus x_{15,25}^{(6)}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

Considering $(a,b,c,d) = (2,7,8,13)$, we have the factors $x_{13,6}^{(5)}$ and $x_{13,7}^{(5)}$. Then we can use Lemma 7 to get

$$\Pr\left(x_{13,6}^{(5)} \oplus x_{13,7}^{(5)} = \mathcal{L}_{13,6}^{(6)} \oplus \mathcal{L}_{13,7}^{(6)}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

Considering $(a,b,c,d) = (3,4,9,14)$ we have $x^{(5)}_{9,7}, x^{(5)}_{9,19}$ and $x^{(5)}_{14,8}$, and then we can apply Lemma 3 to obtain

$$\Pr\left(x^{(5)}_{9,7} = \mathcal{L}^{(6)}_{9,7} \oplus x^{(6)}_{3,6} \oplus x^{(6)}_{14,14} \oplus x^{(6)}_{14,6}\right) = \frac{1}{2}\left(1 + \frac{1}{2^2}\right),$$

$$\Pr\left(x^{(5)}_{9,19} = \mathcal{L}^{(6)}_{9,19} \oplus x^{(6)}_{3,18} \oplus x^{(6)}_{14,26} \oplus x^{(6)}_{14,18}\right) = \frac{1}{2}\left(1 + \frac{1}{2^2}\right),$$

$$\Pr\left(x^{(5)}_{14,8} = \mathcal{L}^{(6)}_{14,8} \oplus x^{(6)}_{9,7} \oplus x^{(6)}_{4,14}\right) = \frac{1}{2}\left(1 + \frac{1}{2}\right).$$

By the Piling-up Lemma, we have that all these changes result in a probability of $\frac{1}{2}\left(1 + \frac{1}{2^8}\right)$. Expanding the linear terms using Eqs. (13)-(16) and canceling out common factors completes the proof. □

Computational Result 6 *The linear approximation of Lemma 12 holds computationally with* $\varepsilon_{L_0} = 0.00867 \approx 2^{-6.85}$.

Lemma 13. *The following linear approximation holds with probability* $\frac{1}{2}\left(1 + \frac{1}{2^{47}}\right)$

$$\begin{aligned} x^{(3.5)}_{5,0} = {} & x^{(7)}_{0}[0,6,7,11,12] \oplus x^{(7)}_{1}[7,8,14,15,16,18,19,30,31] \oplus \\ & x^{(7)}_{2}[0,2,3,5,6,8,10,11,14,15,16,18,19,24,25,27,30,31] \oplus \\ & x^{(7)}_{3}[6,7,9,10,18,19,25,26] \oplus x^{(7)}_{4}[1,2,7,19,26] \oplus x^{(7)}_{5}[0,5,6,7] \oplus \\ & x^{(7)}_{6}[1,2,9,10,19,21,22,29,31] \oplus x^{(7)}_{7}[2,3,5,10,11,13,14,19,22,23, \\ & 27,30,31] \oplus x^{(7)}_{8}[6,14,15,19,26,27] \oplus x^{(7)}_{9}[8,13,19,25,30,31] \oplus \\ & x^{(7)}_{10}[2,3,7,12,14,15,23,24,27] \oplus x^{(7)}_{11}[0,3,4,8,12,13,14,18,20,27, \\ & 28,30] \oplus x^{(7)}_{12}[0,5,6,11,12,19,20] \oplus x^{(7)}_{13}[0,7,12,13,15,16,18,19,22, \\ & 23,24,26,27] \oplus x^{(7)}_{14}[1,2,8,10,11,13,14,16,18,19,22,23,24,25,26, \\ & 30,31] \oplus x^{(7)}_{15}[5,6,7,8,13,14,15,16,23]. \end{aligned}$$

Proof. If we start from Lemma 12 we want to expand the equation one more round. To do so, first note that since we are transitioning from round 6 to 7, we have $(a,b,c,d) \in \{(0,4,8,12),(1,5,9,13),(2,6,10,14),(3,7,11,15)\}$. Therefore, we can divide the factors of the equation in 4 distinct groups:

- Group I - $x^{(6)}_{0}[0], x^{(6)}_{4}[7,14,15], x^{(6)}_{8}[6,7,12]$.
- Group II - $x^{(6)}_{5}[13], x^{(6)}_{9}[0,8,19], x^{(6)}_{13}[0,30,31]$.
- Group III - $x^{(6)}_{2}[0,6,7,22,23], x^{(6)}_{10}[0,6,26], x^{(6)}_{14}[0,6,7,14,15,18,19,24,26,27]$.
- Group IV - $x^{(6)}_{3}[0,6,7,8,16,18,19,24], x^{(6)}_{7}[7,13,14,19], x^{(6)}_{15}[0,8,25,26]$.

Here, we follow the same strategy as in the proof of Lemma 11. In Group I, the factor $x^{(6)}_{0,0}$ can be expanded using Lemma 1 with probability 1. Next, we can combine the following factors: $x^{(6)}_{4,7}, x^{(6)}_{8,6}, x^{(6)}_{8,7}$ using Lemma 8 ($k = 2$); $x^{(6)}_{4,14}, x^{(6)}_{4,15}$

using Lemma 7 with ($k = 1$). Finally, we expand $x^{(6)}_{8,12}$ using Lemma 3 ($k = 2$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$\begin{aligned}&x^{(6)}_0[0] \oplus x^{(6)}_4[7, 14, 15] \oplus x^{(6)}_8[6, 7, 12] = x^{(7)}_0[0, 6, 7, 11, 12]\oplus\\&x^{(7)}_4[1, 2, 7, 19, 26] \oplus x^{(7)}_8[6, 14, 15, 19, 26, 27] \oplus x^{(7)}_{12}[0, 5, 6, 11, 12, 19, 20]\end{aligned} \tag{37}$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^5}\right)$.

In Group II, the factors $x^{(6)}_{9,0}, x^{(6)}_{13,0}$ can be expanded using Lemma 1 with probability 1. Next, we can combine $x^{(6)}_{13,30}, x^{(6)}_{13,31}$ using Lemma 7 ($k = 1$). The remaining terms can be expanded with Lemma 3: $x^{(6)}_{9,8}$ ($k = 2$); $x^{(6)}_{9,19}$ ($k = 2$); $x^{(6)}_{5,13}$ ($k = 1$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$\begin{aligned}&x^{(6)}_5[13] \oplus x^{(6)}_9[0, 8, 19] \oplus x^{(6)}_{13}[0, 30, 31] = x^{(7)}_1[7, 8, 14, 15, 16, 18, 19, 30, 31]\oplus\\&x^{(7)}_5[0, 5, 6, 7] \oplus x^{(7)}_9[8, 13, 19, 25, 30, 31] \oplus x^{(7)}_{13}[0, 7, 12, 13, 15, 16, 18, 19, 22,\\&23, 24, 26, 27]\end{aligned} \tag{38}$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^6}\right)$.

In Group III, the factors $x^{(6)}_{2,0}, x^{(6)}_{10,0}$ and $x^{(6)}_{14,0}$ can be expanded using Lemma 1 with probability 1. Next, we can combine the following factors: $x^{(6)}_{2,6}, x^{(6)}_{2,7}, x^{(6)}_{10,6}$, $x^{(6)}_{14,6}, x^{(6)}_{14,7}$ using Eq. (30) of Lemma 9 ($k = 3$); $x^{(6)}_{2,22}, x^{(6)}_{2,23}$ using Lemma 7 ($k = 3$); $x^{(6)}_{14,14}, x^{(6)}_{14,15}$ using Lemma 7 ($k = 1$); $x^{(6)}_{14,18}, x^{(6)}_{14,19}$ using Lemma 7 ($k = 1$); $x^{(6)}_{14,26}, x^{(6)}_{14,27}$ using Lemma 7 ($k = 1$). Finally, it remains some single terms to be expanded: $x^{(6)}_{10,26}$ using Lemma 3 ($k = 2$); $x^{(6)}_{14,24}$ using Lemma 6 ($k = 1$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$\begin{aligned}&x^{(6)}_2[0, 6, 7, 22, 23], x^{(6)}_{10}[0, 6, 26], x^{(6)}_{14}[0, 6, 7, 14, 15, 18, 19, 24, 26, 27] =\\&x^{(7)}_2[0, 2, 3, 5, 6, 8, 10, 11, 14, 15, 16, 18, 19, 24, 25, 27, 30, 31]\oplus\\&x^{(7)}_6[1, 2, 9, 10, 19, 21, 22, 29, 31] \oplus x^{(7)}_{10}[2, 3, 7, 12, 14, 15, 23, 24, 27]\oplus\\&x^{(7)}_{14}[1, 2, 8, 10, 11, 13, 14, 16, 18, 19, 22, 23, 24, 25, 26, 30, 31]\end{aligned} \tag{39}$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^{12}}\right)$.

In Group IV, the factors $x^{(6)}_{3,0}$ and $x^{(6)}_{15,0}$ can be expanded using Lemma 1 with probability 1. Then we can combine the following factors: $x^{(6)}_{3,6}, x^{(6)}_{3,7}, x^{(6)}_{7,7}$ using Eq. (26) of Lemma 9 ($k = 3$); $x^{(6)}_{3,18}, x^{(6)}_{3,19}, x^{(6)}_{7,19}$ using Eq. (26) of Lemma 9 ($k = 3$); $x^{(6)}_{3,8}, x^{(6)}_{15,8}$ using Eq. (24) of Lemma 9 ($k = 2$); $x^{(6)}_{15,25}, x^{(6)}_{15,26}$ using Lemma 7 ($k = 1$); $x^{(6)}_{7,13}, x^{(6)}_{7,14}$ using Lemma 7 ($k = 1$). It remains some single terms to be expanded: $x^{(6)}_{3,16}$ using Lemma 6 ($k = 3$); $x^{(6)}_{3,24}$ using Lemma 6 ($k = 3$). By the Piling-up Lemma, we can combine these linear relations to obtain

$$\begin{aligned}&x_3^{(6)}[0,6,7,8,16,18,19,24], x_7^{(6)}[7,13,14,19], x_{15}^{(6)}[0,8,25,26] =\\ &x_3^{(7)}[6,7,9,10,18,19,25,26] \oplus x_7^{(7)}[2,3,5,10,11,13,14,19,22,23,27,30,31]\oplus\\ &x_{11}^{(7)}[0,3,4,8,12,13,14,18,20,27,28,30] \oplus x_{15}^{(7)}[5,6,7,8,13,14,15,16,23]\end{aligned} \tag{40}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^{16}}\right)$.

Finally, using the Piling-up Lemma we can combine the results from Lemma 12 and Eqs. (37)-(40), which leads to a correlation of $\varepsilon_L = 1/2^{8+5+6+12+16} = 2^{-47}$. □

Computational Result 7 *The linear approximation of Eq.* (37) *holds computationally with* $\varepsilon_{L_1} = 0.0416 \approx 2^{-4.59}$. *This correlation was verified using* 2^{38} *random samples.*

Computational Result 8 *The linear approximation of Eq.* (38) *holds computationally with* $\varepsilon_{L_2} = 0.0278 \approx 2^{-5.19}$. *This correlation was verified using* 2^{38} *random samples.*

Computational Result 9 *The linear approximation of Eq.* (39) *holds computationally with* $\varepsilon_{L_3} = 0.000398 \approx 2^{-11.29}$. *This correlation was verified using* 2^{42} *random samples.*

Computational Result 10 *The linear approximation of Eq.* (40) *holds computationally with* $\varepsilon_{L_4} = 0.000047 \approx 2^{-14.38}$. *This correlation was verified using* 2^{42} *random samples.*

It is interesting to note that the experimental correlation is higher than expected in several cases. Of course, since the hypothesis of independence for the Piling-Up Lemma does not hold, it is expected to see deviations between what is predicted theoretically and what we see in practice. The fact that the correlation is usually higher indicates a positive correlation between some equations. In future works, it may be interesting to try to understand why ChaCha has this behavior.

4 Improved Differential-Linear Attacks Against ChaCha

4.1 New Differentials

In this section, we present new differentials for 3.5 rounds of ChaCha. As in previous works, these differential correlations were found experimentally. To find these correlations we used the technique proposed by Beierle et al. at Crypto 2020 [4], and described in Section 2.2. Here, the cipher is divided into the sub ciphers E_1 covering 1 round and E_2 covering 2.5 rounds to find a differential path for 3.5 rounds. Thus we want a particular differential characteristic of the form

$$\varDelta X^{(0)} \xrightarrow{\text{1 round}} \varDelta X^{(1)} \xrightarrow{\text{2.5 rounds}} \varDelta X^{(3.5)}.$$

The idea is to generate consistent $\Delta X^{(1)}$ whose Hamming weight is minimized. In [4], the authors showed that the following differential characteristic occurs with probability 2^{-5} on average for the QRF of ChaCha

$$\begin{aligned}\Delta X^{(0)} = (([]),([]),([]),([i])) \rightarrow \Delta X^{(1)} = &(([i+28]),([i+31,i+23,i+11,\\ &i+3]),([i+24,i+16,i+4]),\\ &([i+24,i+4])).\end{aligned} \tag{41}$$

From there we computed $\Delta X^{(3.5)}$ by generating random states $X^{(1)}$ and $X'^{(1)}$ and statistically testing for correlations in particular bits of $\Delta X^{(3.5)}$. We note that this procedure is computationally intensive as some of the correlations are very small. For some bits, we executed this procedure up to 2^{50} pairs of random states in the first round. To achieve this amount of computation we used 8 NVIDIA GPUs (RTX 2080ti). As in the referred paper, we used $i = 6$. Also, we fixed the differential of Eq. (41) in the third column of the state matrix. Table 3 shows the results[1].

$\mathcal{OD}$	$\lvert\varepsilon_d\rvert$
$\Delta x_{0,0}^{(3.5)}$	0.00002797
$\Delta x_{13,0}^{(3.5)}$	0.000003032

Table 3: New differentials after 3.5 rounds, starting from $\Delta X^{(1)}$ in the third column of the state matrix with $i = 6$ in Eq. (41).

4.2 Distinguishers

Using the linear approximations of Lemma 10 and Lemma 11, the differential correlation $\varepsilon_d = 0.00048$ for $(a,b) = (3,4)$ described in [10], and the estimated correlations from the Computational Results 1-5, we get $\varepsilon_d \varepsilon_{L_0}^2 \approx 2^{-25.37}$ which gives us a distinguisher for 6 rounds of ChaCha with complexity less than 2^{51}. Also, we get $\varepsilon_d(\varepsilon_{L_0}\varepsilon_{L_1}\varepsilon_{L_2}\varepsilon_{L_3}\varepsilon_{L_4})^2 \approx 2^{-111.86}$ which gives us a distinguisher for 7 rounds of ChaCha with complexity less than 2^{224}.

[1] Since the first version of this paper was published, several independent researches reviewed our results and code. We would like to thank Juan C. G. Vásquez (juan.grados@tii.ae) for identifying an error in the code we made publicly available. That error affected the results of this table in the first version of the paper. Dey et al. [12] independently noticed that the results reported were not accurate and computed an alternative version of this table. However, we were only able to reproduce the results reported for $\Delta x_{0,0}^{(3.5)}$ and $\Delta x_{13,0}^{(3.5)}$. More precisely, it seems that Dey et al. had inaccuracies of their own, caused by a small number of samples (2^{37}) which is not enough to compute the true correlation for these bits. After correcting the code, we could not find significant results for $\Delta x_{1,0}^{(3.5)}$, $\Delta x_{12,0}^{(3.5)}$ and $\Delta x_{5,0}^{(3.5)}$ as previously reported, even considering 2^{52} samples.

4.3 New Attack using Probabilistic Neutral Bits (PNBs)

One of the most important attacks against ChaCha is the proposal of Aumasson [1]. The attack first identifies good choices of truncated differentials, then it uses probabilistic backwards computation with the notion of PNBs, estimating the complexity of the attack. This attack is described in several previous works [1, 24, 23], thus, in our description, we skip several details.

The PNB-based key recovery is a fully experimental approach. We summarize the technique as follows:

- Let the correlation in the forward direction (a.k.a, differential-linear distinguisher) after r rounds be ε_d.
- Let n be the number of PNBs given by a correlation γ. Namely, even if we flip one bit in PNBs, we still observe correlation γ.
- Let the correlation in the backward direction, where all PNB bits are fixed to 0 and non-PNB bits are fixed to the correct ones, is ε_a

Then, the time complexity of the attack is estimated as $2^{256-n}N + 2^{256-\alpha}$, where the data complexity N is given as

$$N = \left(\frac{\sqrt{\alpha \log(4)} + 3\sqrt{1 - \varepsilon_a^2 \varepsilon_d^2}}{\varepsilon_a \varepsilon_d} \right)^2 ,$$

where α is a parameter that the attacker can choose.

We can implement new attacks for 7 rounds of ChaCha using this technique by considering the new differential correlation for $\Delta_{13,0}^{(3.5)}$ presented in Table 3. Using Eq. (7) it is easy to see that we have $x_{13,0}^{(3.5)} = x_{2,0}^{(4)} \oplus x_{13,8}^{(4)}$. Therefore, we consider $\mathcal{ID}$ given by Eq. (41) with $i = 6$ and $\mathcal{OD}$ $x_{2,0}^{(4)} \oplus x_{13,8}^{(4)}$. Using $\gamma = 0.35$ we found 83 PNBs, and we obtained $\varepsilon_a = 0.000509$. From that, we get an attack with data complexity of $2^{64.59}$ and time complexity $2^{237.59}$. As in [4], we have to repeat this attack 2^5 times on average. Thus, the final attack has data complexity of $2^{69.58}$ and time complexity $2^{242.59}$, which does not improve previous results. Bellow we list all PNBs:

$$\begin{aligned}\text{PNB} = (&0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 19, 20, 21, 31, 32, 33, 34, 35, 36, 39, 43, 47, 48,\\ &49, 53, 54, 55, 59, 63, 67, 68, 69, 70, 71, 72, 73, 89, 90, 95, 99, 100, 103,\\ &104, 105, 123, 124, 125, 126, 127, 128, 129, 130, 140, 141, 142, 152, 153, 154,\\ &155, 156, 157, 158, 159, 168, 169, 170, 174, 175, 176, 184, 185, 186, 187, 188,\\ &189, 190, 191, 192, 193, 210, 223, 248, 255).\end{aligned}$$

5 Conclusion

In this paper, we presented a new technique to find linear approximations for ARX ciphers. Applying this technique we presented new linear approximations to the stream cipher ChaCha which gave us new and improved distinguishers. In addition, we presented new differential characteristics for 3.5 rounds of ChaCha and use them to create new attacks based on Probabilistic Neutral Bits. For future works, we expect that the proposed technique can be used to improve attacks against similar ARX-based designs, as the stream cipher Salsa and the hash function Blake. Additionally, Lemma 13 shows that it is possible to improve further the linear correlation used in our distinguisher for 7 rounds. Thus, it may be possible to further improve attacks to ChaCha given a better differential correlation.

Acknowledgements. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions which helped us to improve our work.

References

1. Aumasson, J.P., Fischer, S., Khazaei, S., Meier, W., Rechberger, C.: New features of latin dances: analysis of Salsa, ChaCha, and Rumba. In: International Workshop on Fast Software Encryption. pp. 470–488. Springer (2008)
2. Aumasson, J.P., Henzen, L., Meier, W., Phan, R.C.W.: SHA-3 proposal blake. Submission to NIST **92** (2008)
3. Beierle, C., Biryukov, A., Cardoso Dos Santos, L., Groszschädl, J., Perrin, L.P., Udovenko, A., Velichkov, V., Wang, Q.: Schwaemm and Esch: lightweight authenticated encryption and hashing using the Sparkle permutation family (2019)
4. Beierle, C., Leander, G., Todo, Y.: Improved differential-linear attacks with applications to ARX ciphers. In: Annual International Cryptology Conference. pp. 329–358. Springer (2020)
5. Bernstein, D.J.: The Poly1305-AES message-authentication code. In: International Workshop on Fast Software Encryption. pp. 32–49. Springer (2005)
6. Bernstein, D.J.: ChaCha, a variant of Salsa20. In: Workshop Record of SASC. vol. 8, pp. 3–5 (2008)
7. Bernstein, D.J.: The Salsa20 family of stream ciphers. In: New stream cipher designs, pp. 84–97. Springer (2008)
8. Blondeau, C., Leander, G., Nyberg, K.: Differential-linear cryptanalysis revisited. Journal of Cryptology **30**(3), 859–888 (2017)
9. Choudhuri, A.R., Maitra, S.: Significantly improved multi-bit differentials for reduced round Salsa and Chacha. IACR Transactions on Symmetric Cryptology pp. 261–287 (2016)
10. Coutinho, M., Neto, T.S.: New multi-bit differentials to improve attacks against ChaCha. IACR Cryptol. ePrint Arch. **2020**, 350 (2020)
11. Crowley, P.: Truncated differential cryptanalysis of five rounds of Salsa20. The State of the Art of Stream Ciphers SASC **2006**, 198–202 (2006)
12. Dey, S., Dey, C., Sarkar, S., Meier, W.: Revisiting cryptanalysis on chacha from crypto 2020 and eurocrypt 2021. Cryptology ePrint Archive, Report 2021/1059 (2021), `https://ia.cr/2021/1059`
13. Dey, S., Roy, T., Sarkar, S.: Revisiting design principles of Salsa and ChaCha. Advances in Mathematics of Communications **13**(4) (2019)
14. Dey, S., Sarkar, S.: Improved analysis for reduced round Salsa and Chacha. Discrete Applied Mathematics **227**, 58–69 (2017)
15. Ding, L.: Improved related-cipher attack on Salsa20 stream cipher. IEEE Access **7**, 30197–30202 (2019)

16. Dinu, D., Perrin, L., Udovenko, A., Velichkov, V., Großschädl, J., Biryukov, A.: Design strategies for ARX with provable bounds: Sparx and LAX. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 484–513. Springer (2016)
17. Fischer, S., Meier, W., Berbain, C., Biasse, J.F., Robshaw, M.J.: Non-randomness in eSTREAM candidates Salsa20 and TSC-4. In: International Conference on Cryptology in India. pp. 2–16. Springer (2006)
18. Hernandez-Castro, J.C., Tapiador, J.M., Quisquater, J.J.: On the Salsa20 core function. In: International Workshop on Fast Software Encryption. pp. 462–469. Springer (2008)
19. IANIX: ChaCha usage & deployment. `https://ianix.com/pub/chacha-deployment.html` (2020), accessed: 2020-01-13
20. Ishiguro, T., Kiyomoto, S., Miyake, Y.: Latin dances revisited: new analytic results of Salsa20 and ChaCha. In: International Conference on Information and Communications Security. pp. 255–266. Springer (2011)
21. Langford, S.K., Hellman, M.E.: Differential-linear cryptanalysis. In: Annual International Cryptology Conference. pp. 17–25. Springer (1994)
22. Langley, A., Chang, W., Mavrogiannopoulos, N., Strombergson, J., Josefsson, S.: ChaCha20-Poly1305 cipher suites for transport layer security (TLS). RFC 7905 (10) (2016)
23. Maitra, S., Paul, G., Meier, W.: Salsa20 cryptanalysis: New moves and revisiting old styles. In: The Ninth International Workshop on Coding and Cryptography (2015)
24. Maitra, S.: Chosen IV cryptanalysis on reduced round ChaCha and Salsa. Discrete Applied Mathematics **208**, 88–97 (2016)
25. Mouha, N., Preneel, B.: A proof that the ARX cipher Salsa20 is secure against differential cryptanalysis. IACR Cryptology ePrint Archive **2013**, 328 (2013)
26. Muller, S.: Documentation and analysis of the Linux random number generator - federal office for information security (germany's) (2019), `https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Studies/LinuxRNG/LinuxRNG_EN.pdf;jsessionid=6B0F8D7795B80F5EADA3DB3DB3E4043B.1_cid360?__blob=publicationFile&v=19`
27. Robshaw, M., Billet, O.: New stream cipher designs: the eSTREAM finalists, vol. 4986. Springer (2008)
28. Shi, Z., Zhang, B., Feng, D., Wu, W.: Improved key recovery attacks on reduced-round Salsa20 and ChaCha. In: International Conference on Information Security and Cryptology. pp. 337–351. Springer (2012)
29. Torvalds, L.: Linux kernel source tree (2016), `https://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git/commit/?id=818e607b57c94ade9824dad63a96c2ea6b21baf3`
30. Tsunoo, Y., Saito, T., Kubo, H., Suzaki, T., Nakashima, H.: Differential cryptanalysis of Salsa20/8. In: Workshop Record of SASC. vol. 28 (2007)
31. Wallén, J.: Linear approximations of addition modulo 2^n. In: International Workshop on Fast Software Encryption. pp. 261–273. Springer (2003)

A Proofs

In this appendix, we expand the proof of Lemma 9 for each individual linear approximation.

A.1 Eq. (19)

Proof. Using Eqs. (9) and (10) we can write

$$\begin{aligned} x_{b,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = {} & \mathcal{L}_{b,i}^{(m)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \\ & \mathcal{L}_{c,i}^{(m)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}). \end{aligned}$$

Using the approximation of Eq. (17) we can write $\Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}) = x_{d,i-1}'^{(m-1)}$ with probability $\frac{1}{2}\left(1 + \frac{1}{2}\right)$. Thus, using Eq. (7) and canceling out common factors we get

$$x_{b,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus x_{a,i-1}^{(m)} \oplus x_{d,i+7}^{(m)},$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2}\right)$, which concludes the proof. □

A.2 Eqs. (20) and (21)

Proof. Using Eqs. (9) and (12) we can write

$$\begin{aligned} x_{a,i}^{(m-1)} \oplus x_{b,i}^{(m-1)} = {} & \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \\ & \Theta_i(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \Theta_i(x_a^{(m-1)}, x_b^{(m-1)}). \end{aligned}$$

Cancelling out common factors, using the approximation of Eq. (17) and the Piling-up Lemma we can write

$$x_{a,i}^{(m-1)} \oplus x_{b,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus x_{b,i-1}'^{(m-1)} \oplus x_{b,i-1}^{(m-1)}$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^2}\right)$. Now we can replace $x_{b,i-1}'^{(m-1)}$ using Eq. (5) and $x_{b,i-1}^{(m-1)}$ using Lemma 3, which leads to

$$x_{a,i}^{(m-1)} \oplus x_{b,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus x_{b,i+6}^{(m)} \oplus x_{c,i-1}^{(m)} \oplus \mathcal{L}_{b,i-1}^{(m)} \oplus x_{d,i-2}^{(m)},$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^3}\right)$ by the Piling-up Lemma. We can also use Lemma 1 in order to obtain

$$x_{a,1}^{(m-1)} \oplus x_{b,1}^{(m-1)} = \mathcal{L}_{a,1}^{(m)} \oplus \mathcal{L}_{b,1}^{(m)} \oplus x_{b,7}^{(m)} \oplus x_{c,0}^{(m)} \oplus \mathcal{L}_{b,0}^{(m)},$$

with probability $\frac{1}{2}\left(1 + \frac{1}{2^2}\right)$. □

A.3 Eqs. (22) and (23)

Proof. Combining Eq. (10) and Eq. (12) we have

$$\begin{aligned} x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = {} & \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}) \oplus \\ & \Theta_i(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \Theta_i(x_a^{(m-1)}, x_b^{(m-1)}). \end{aligned}$$

Using the approximation of Eq. (17) and the Piling-up Lemma we can write

$$x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus x_{d,i-1}'^{(m-1)} \oplus x_{b,i-1}'^{(m-1)} \oplus x_{b,i-1}^{(m-1)}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. Now we can replace $x'^{(m-1)}_{d,i-1}$ using Eq. (7), $x'^{(m-1)}_{b,i-1}$ using Eq. (5) and $x^{(m-1)}_{b,i-1}$ using Lemma 3 if $i > 1$ or 1 if $i = 1$, which leads to

$$\begin{aligned} x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{c,i} = {} & \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{c,i} \oplus x^{(m)}_{a,i-1} \oplus x^{(m)}_{d,i+7} \oplus x^{(m)}_{b,i+6} \\ & \oplus x^{(m)}_{c,i-1} \oplus \mathcal{L}^{(m)}_{b,i-1} \oplus x^{(m)}_{d,i-2}, \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^4}\right)$ by the Piling-up Lemma or

$$\begin{aligned} x^{(m-1)}_{a,1} \oplus x^{(m-1)}_{c,1} = {} & \mathcal{L}^{(m)}_{a,1} \oplus \mathcal{L}^{(m)}_{c,1} \oplus x^{(m)}_{a,0} \oplus x^{(m)}_{d,8} \oplus x^{(m)}_{b,7} \\ & \oplus x^{(m)}_{c,0} \oplus \mathcal{L}^{(m)}_{b,0}, \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. □

A.4 Eq. (24)

Proof. Using Eq. (11) and Eq. (12) we can write

$$x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{d,i} = \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{d,i} \oplus \mathit{\Theta}_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus \mathit{\Theta}_i(x^{(m-1)}_{a}, x^{(m-1)}_{b}).$$

Using Eq. (17) we get

$$x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{d,i} = \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{d,i} \oplus \mathit{\Theta}_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus x^{(m-1)}_{b,i-1},$$

and from Eq. (9)

$$\begin{aligned} x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{d,i} = {} & \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{d,i} \oplus \mathit{\Theta}_i(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus \\ & \mathcal{L}^{(m)}_{b,i-1} \oplus \mathit{\Theta}_{i-1}(x'^{(m-1)}_{c}, x^{(m)}_{d}), \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. Thus, using the approximation of Eq. (18) and the Piling-up Lemma we can write

$$x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{d,i} = \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{d,i} \oplus \mathcal{L}^{(m)}_{b,i-1},$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. □

A.5 Eq. (25)

Proof. Using Eq. (12) and Eq. (10) and canceling out common factors we get

$$\begin{aligned} x^{(m-1)}_{a,i-1} \oplus x^{(m-1)}_{a,i} \oplus x^{(m-1)}_{c,i} = {} & \mathcal{L}^{(m)}_{a,i-1} \oplus \mathcal{L}^{(m)}_{a,i} \oplus \mathcal{L}^{(m)}_{c,i} \oplus \\ & \mathit{\Theta}_{i-1}(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \oplus \mathit{\Theta}_{i-1}(x'^{(m-1)}_{c}, x^{(m)}_{d}) \oplus \\ & \mathit{\Theta}_{i-1}(x^{(m-1)}_{a}, x^{(m-1)}_{b}) \oplus \mathit{\Theta}_{i}(x'^{(m-1)}_{a}, x'^{(m-1)}_{b}) \oplus \\ & \mathit{\Theta}_{i}(x^{(m-1)}_{a}, x^{(m-1)}_{b}) \oplus \mathit{\Theta}_{i}(x^{(m-1)}_{c}, x'^{(m-1)}_{d}) \end{aligned}$$

Using the approximation of Eq. (18) and the Piling-up Lemma we obtain

$$\begin{aligned} x_{a,i-1}^{(m-1)} \oplus x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = {} & \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \\ & \mathit{\Theta}_{i-1}(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathit{\Theta}_i(x_c^{(m-1)}, x_d'^{(m-1)}) \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. Using Eq. (17) and Eq. (7) we get

$$\begin{aligned} x_{a,i-1}^{(m-1)} \oplus x_{a,i}^{(m-1)} \oplus x_{c,i}^{(m-1)} = {} & \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \\ & x_{d,i-2}^{(m)} \oplus x_{a,i-1}^{(m)} \oplus x_{d,i+7}^{(m)} \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^4}\right)$. □

A.6 Eq. (26)

Proof. Using Eq. (9) and Eq. (12) and canceling out common factors we can write

$$\begin{aligned} & x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \\ & \mathit{\Theta}_{i-1}(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \mathit{\Theta}_{i-1}(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathit{\Theta}_{i-1}(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \\ & \mathit{\Theta}_i(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \mathit{\Theta}_i(x_a^{(m-1)}, x_b^{(m-1)}). \end{aligned}$$

Using the approximation of Eq. (18) and the Piling-up Lemma we can write

$$\begin{aligned} x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} = {} & \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \\ & \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathit{\Theta}_{i-1}(x_c'^{(m-1)}, x_d^{(m)}). \end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. Using the approximation of Eq. (17) we get

$$x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus x_{d,i-2}^{(m)}.$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. □

A.7 Eq. (27)

Proof. Using Eq. (11) and Eq. (12), and canceling out common factors we have

$$\begin{aligned} x_{b,i-1}^{(m-1)} \oplus x_{a,i}^{(m-1)} \oplus x_{d,i}^{(m-1)} = {} & x_{b,i-1}^{(m-1)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathit{\Theta}_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \\ & \mathit{\Theta}_i(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \mathcal{L}_{d,i}^{(m)}. \end{aligned}$$

Using the approximation of Eq. (17) we have $\mathit{\Theta}_i(x_a^{(m-1)}, x_b^{(m-1)}) = x_{b,i-1}^{(m-1)}$ occurring with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. Then

$$x_{b,i-1}^{(m-1)} \oplus x_{a,i}^{(m-1)} \oplus x_{d,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus \mathit{\Theta}_i(x_c'^{(m-1)}, x_d^{(m)}).$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. Finally, using the approximation of Eq. (17) and the Piling-up Lemma we get

$$x_{b,i-1}^{(m-1)} \oplus x_{a,i}^{(m-1)} \oplus x_{d,i}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus x_{d,i-1}^{(m)}.$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. □

A.8 Eq. (28)

Proof. Using Eq. (9) and Eq. (10), we can write

$$\begin{aligned}&x_{b,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \mathcal{L}_{b,i-1}^{(m)} \oplus \Theta_{i-1}(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathcal{L}_{b,i}^{(m)} \oplus \\ &\Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \Theta_{i-1}(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}) \oplus \\ &\mathcal{L}_{c,i}^{(m)} \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}).\end{aligned}$$

Canceling out common factors we get

$$\begin{aligned}x_{b,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{c,i}^{(m-1)} = {}&\mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)} \oplus \\ &\Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}) \oplus \\ &\Theta_i(x_c^{(m-1)}, x_d'^{(m-1)}).\end{aligned}$$

Thus, using the approximation of Eq. (18) we get

$$x_{b,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{c,i}^{(m-1)} = \mathcal{L}_{b,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \mathcal{L}_{c,i}^{(m)}.$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. □

A.9 Eq. (29)

Proof. Using equations (9), (10) and (12)

$$\begin{aligned}&x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ &\Theta_i(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \Theta_i(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \Theta_{i-1}(x_a'^{(m-1)}, x_b'^{(m-1)}) \oplus \\ &\Theta_{i-1}(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}).\end{aligned}$$

Using the approximation of Eq. (18) and the Piling-up Lemma we can write

$$\begin{aligned}x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = {}&\mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ &\Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}).\end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^2}\right)$. Therefore, equations (17) and (7) give us

$$\begin{aligned}x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{b,i}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} = {}&\mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{b,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ &x_{a,i-2}^{(m)} \oplus x_{d,i+6}^{(m)}.\end{aligned}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. □

A.10 Eq. (30)

Proof. Using equations (10), (11) and (12), we can write

$$\begin{aligned}&x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{d,i}^{(m-1)} \oplus x_{d,i-1}^{(m-1)} = \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ &\mathcal{L}_{d,i-1}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus \Theta_{i-1}(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \Theta_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \\ &\Theta_i(x_a^{(m-1)}, x_b^{(m-1)}) \oplus \Theta_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}).\end{aligned}$$

Using the approximation of Eq. (18) we have

$$\begin{gathered} x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{d,i}^{(m-1)} \oplus x_{d,i-1}^{(m-1)} = \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ \mathcal{L}_{d,i-1}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus \mathit{\Theta}_i(x_c'^{(m-1)}, x_d^{(m)}) \oplus \mathit{\Theta}_{i-1}(x_c^{(m-1)}, x_d'^{(m-1)}) \end{gathered}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2}\right)$. Finally, by the Piling-up Lemma and using the approximation of Eq. (17) and Eq. (7), we get

$$\begin{gathered} x_{a,i}^{(m-1)} \oplus x_{a,i-1}^{(m-1)} \oplus x_{c,i-1}^{(m-1)} \oplus x_{d,i}^{(m-1)} \oplus x_{d,i-1}^{(m-1)} = \mathcal{L}_{a,i-1}^{(m)} \oplus \mathcal{L}_{a,i}^{(m)} \oplus \mathcal{L}_{c,i-1}^{(m)} \oplus \\ \mathcal{L}_{d,i-1}^{(m)} \oplus \mathcal{L}_{d,i}^{(m)} \oplus x_{d,i-1}^{(m)} \oplus x_{a,i-2}^{(m)} \oplus x_{d,i+6}^{(m)} \end{gathered}$$

with probability $\frac{1}{2}\left(1+\frac{1}{2^3}\right)$. □