Proving TLS-attack related open biases of RC4

Santanu Sarkar[1], Sourav Sen Gupta[2], Goutam Paul[2], and Subhamoy Maitra[2]

[1] Chennai Mathematical Institute, Chennai
[2] Indian Statistical Institute, Kolkata
sarkar.santanu.bir@gmail.com, sg.sourav@gmail.com, goutam.k.paul@gmail.com, subho@isical.ac.in

Abstract. After a series of works on RC4 cryptanalysis in last few years (published in flagship cryptology conferences and journals), the most significant (and also very recent) attack on the cipher has been the discovery of vulnerabilities in the SSL/TLS protocol, by AlFardan, Bernstein, Paterson, Poettering and Schuldt. They ran extensive computations to identify significant short-term single-byte keystream biases of RC4, and utilized that knowledge in the attack. The biases identified by AlFardan et al. consist of earlier known biases of RC4, as well as some newly discovered ones.

In this paper, we attempt at proving the new, unproved or partially proved biases amongst the above-mentioned ones. The theoretical proofs of these biases not only assert a scientific justification, but also discover intricate patterns and operations of the cipher associated with these biases. For example, while attempting the proof of a bias of the first output byte towards 129, we observe that this bias occurs prominently only for certain lengths of the secret key of RC4. In addition, our findings reveal that this bias may be related to the old and unsolved problem of "anomalies" in the distribution of the state array after the Key Scheduling Algorithm. In this connection, we prove the anomaly in $S_0[128] = 127$, a problem open for more than a decade.

Other than proving the new biases, we also complete the proof for the extended keylength dependent biases in RC4, a problem attempted and partially solved by Isobe, Ohigashi, Watanabe and Morii in FSE 2013. Our new proofs and observations in this paper, along with the connection to the older results, provide a comprehensive view on the state-of-the-art literature in RC4 cryptanalysis.

Keywords: Stream cipher, RC4, Biases, Short-term, Keylength dependent, Anomaly.

1 Introduction

Over the last three decades of research in stream ciphers, several designs have been proposed and analyzed by the community. The RC4 stream cipher, 'allegedly' designed by Rivest in 1987, has sustained to be one of the most popular ciphers in this category for more than 25 years. The cipher has continued gaining its fabled popularity for its intriguing simplicity that has made it widely accepted in the community for various software and web applications.

The cipher consists of two major components, the Key Scheduling Algorithm (KSA) and the Pseudo-Random Generation Algorithm (PRGA). The internal permutation of RC4 is of N bytes, and so is the key K. The original secret key is of length typically between 5 to 32 bytes, and is repeated to form the final key K. The KSA produces the initial permutation of RC4 by scrambling an identity permutation using key K. The initial permutation S produced by the KSA acts as an input to the next procedure PRGA that generates the output keystream, as shown in Fig. 1.

Notation. For round $r \geq 1$ of RC4 PRGA, we denote the indices by i_r, j_r, the output byte by Z_r, the index location of output Z_r as t_r, and the permutations before and after the swap by S_{r-1} and S_r respectively. Thus, round r of RC4 PRGA is defined by $i_r = i_{r-1} + 1$, $j_r = j_{r-1} + S_{r-1}[i_r]$, swap $S_{r-1}[i_r] \leftrightarrow S_{r-1}[j_r]$, $t_r = S_r[i_r] + S_r[j_r]$, and $Z_r = S_r[t_r]$. After $r \geq 1$ rounds of KSA, we denote the state variables by adding a superscript K to each variable. By S_0^K and S_0, we denote the initial permutations before KSA and PRGA respectively. Note that S_0^K is the identity permutation and $S_0 = S_N^K$ is the permutation obtained right after the completion of KSA. Throughout this paper, all operations in context of RC4 are to be considered 'modulo N'.

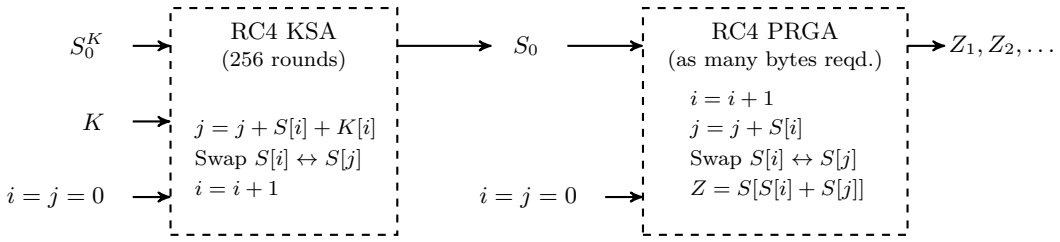

Fig. 1. Key-Scheduling Algorithm and Pseudo-Random Generation Algorithm of RC4.

1.1 Motivation of our work

In a recent paper [24] at FSE 2013, Sepehrdad, Susil, Vaudenay, and Vuagnoux have rightly claimed:

> For some people, attacking WEP is like beating a dead horse, but this horse is still running wildly in many countries all over the world. Also, some companies are selling hardware using modified versions of the WEP protocol, they claim to be secure.

IEEE WiFi security protocol WEP is based on the stream cipher RC4, and hence the same statement applies to RC4 as well. The history of RC4 cryptanalysis is more than 20 years old. However, in recent times, there is a renewed surge of interest in RC4 cryptanalysis in the cryptographic community. For example, significant cryptanalytic results on WEP and WPA have been published in Eurocrypt 2011 by Sepehrdad, Vaudenay, and Vuagnoux [26]. In only the first quarter of the current year (2013), RC4 has attracted 10 publications [1, 6, 8–10, 17, 20, 22, 24, 27]. In spite of this, many problems are still open and the cipher is not yet broken.

As a stream cipher, RC4 promises to deliver pseudo-random bytes as keystream output. Thus, any lapse in that goal creates interesting consequences towards the security of the cipher. This is the reason why statistical weaknesses like biases and their application as distinguishers have attracted the main focus of RC4 cryptanalysis to date. There have been numerous results on RC4 biases over years, and the trend still continues.

Most of the existing results are targeted towards specific short-term (involving only the initial few bytes of the output) biases and correlations [1,5–7,11,12,15,16,19,22,23,25,26], while there exist only a few important results for long-term (prominent even after discarding an arbitrary number of initial bytes of the output) biases [2, 4, 5, 7, 12, 14].

In this paper, we concentrate on the short-term traits of non-random behavior in the initial keystream bytes of RC4, especially in the first N output bytes. The prominent results on the short-term biases of RC4 include Mantin and Shamir second byte bias [15], Mironov first byte sine-curve-like distribution [16], Maitra, Paul and Sen Gupta short-term biases towards zero [11], Sen Gupta, Maitra, Paul and Sarkar proof of first byte bias [22], Sarkar second byte negative bias [20], Isobe, Ohigashi, Watanabe and Morii full broadcast attack [6], and the most recent results by AlFardan, Bernstein, Paterson, Poettering and Schuldt [1, 3].

AlFardan, Bernstein, Paterson, Poettering and Schuldt [1,3]. The most prominent attempt at identifying all possible single-byte short-term biases in the initial keystream bytes of RC4 was recently made by AlFardan, Bernstein, Paterson, Poettering and Schuldt [1, 3]. They ran extensive experiments, using more than 2^{44} random keys, to generate a list of approximately 65536 single-byte short-term biases of RC4, including the previously known ones [6, 11, 15, 16, 22]. This search provides a comprehensive list of non-random behavior of the initial keystream bytes (bytes 1 to $N = 256$) of RC4 when a 16-byte key is used.

The main goal of this analysis [1] was to exploit those in a practical attack against the SSL/TLS protocol that uses RC4 for confidentiality. The authors could use all of the above-mentioned 65536 initial short-term biases of RC4 to mount a plaintext recovery attack on the SSL/TLS protocol that recovers the first 256 bytes of the plaintext from the knowledge of only 2^{32} ciphertexts generated using random keys, with no prior plaintext knowledge. This attack by AlFardan, Bernstein, Paterson, Poettering and Schuldt [1] is undoubtedly the most extensive attack on any practical RC4 protocol to date, with far-reaching consequences. This attack alone is sufficient to highlight the practical importance of identifying, proving and exploiting short-term biases in RC4.

RC4 short-term landscape generated from the data of [1, 3]. The extensive experimental results by AlFardan, Bernstein, Paterson, Poettering and Schuldt [1] identified several non-randomnesses in the short-term output keystream of RC4. Figure 2 presents a 3D model of the probabilities $\Pr(Z_r = v)$ for $r = 1, \ldots, N$ and $v = 0, \ldots, N-1$, which we call the RC4 landscape of initial keystream bytes.

Note that this landscape is for the most practical version of RC4 that uses a 16-byte key, and is not identical for RC4 initial keystream patterns generated by secret keys of various other lengths. For example, the non-random peaks and troughs present in the 16-byte key landscape reduce to a certain extent if one uses a full length $N = 256$ bytes key.

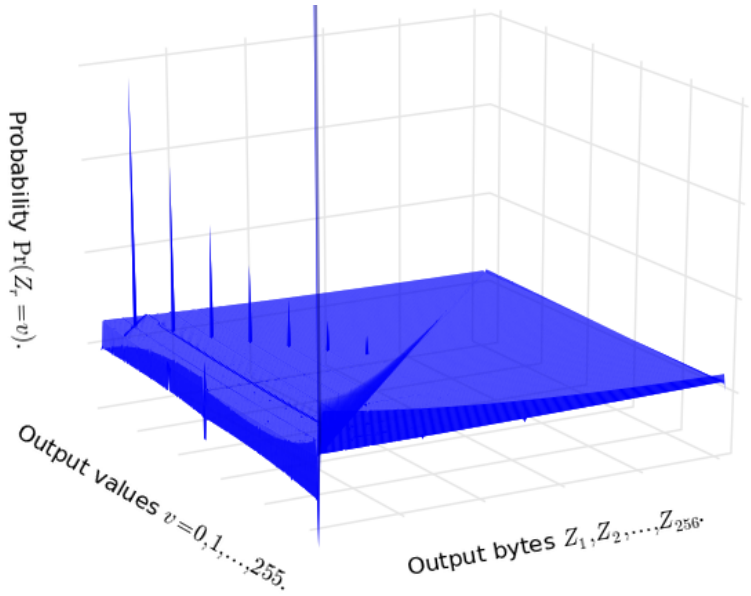

Fig. 2. The RC4 landscape of initial keystream bytes (data from [1,3]).

The visible vertical walls and spikes in Fig. 2 identify the prominent short-term bias patterns in the RC4 landscape. The main ones are for the events $Z_2 = 0$ (largest positive spike), $Z_2 = 2$ (largest negative spike), $Z_1 = v$ where $v = 0, \ldots, N-1$ (sine-curve-like vertical wall on the left side), $Z_r = 0$ (decreasing vertical wall on the right side), $Z_r = r$ (decreasing vertical wall at the center) and $Z_r = -r$ (decreasing series of spikes at the center), where $r = 1, \ldots, N$ denotes the number of the round in RC4 PRGA.

The proofs for most of these major non-random events are present in the literature. The biases in $Z_2 = 0$ and $Z_2 = 2$ have been proved by Mantin and Shamir [15] in 2001 and Sarkar [20] in 2013

respectively. The sine-curve-like pattern of Z_1 for full-length key, including the negative biases in $Z_1 = 0, 1$, have been proved by Sen Gupta, Maitra, Paul and Sarkar [22] in 2013, and the general proof for $Z_r = 0$ has been done by Maitra, Paul and Sen Gupta [11] in 2011. In 2011, Sen Gupta, Maitra, Paul and Sarkar [21] proved the $Z_r = -r$ case for $r = 16$ (keylength), and in 2013, the general pattern for $Z_r = -r$ was partially proved by Isobe, Ohigashi, Watanabe and Morii [6]. In the same paper of 2013, Isobe, Ohigashi, Watanabe and Morii [6] proved the bias pattern for $Z_r = r$, and the slightly weaker single-byte bias of $Z_3 = 131$.

Table 1. Identified and/or proved short-term keystream biases of RC4.

Bias in event	Type of bias	Discovered	Proved
	Isolated short-term biases		
$Z_1 = 0$	Negative	[16]	[22]
$Z_1 = 1$	Negative	[22]	[22]
$Z_1 = 129$	Negative (16-byte key)	[1]	Open
$Z_2 = 0$	Positive	[15]	[15]
$Z_2 = 2$	Negative	[1, 20]	[20]
$Z_2 = 129$	Negative	[1, 20]	Open
$Z_2 = 172$	Positive	[1]	Open
$Z_3 = 131$	Positive	[1, 6]	[6]
$Z_4 = 2$	Positive	[1]	Open
$Z_{256} = 0$	Negative	[1, 6]	Open
$Z_{257} = 0$	Positive	[6]	Open
	Patterns of short-term biases		
$Z_1 = v$	Sinusoidal ($v = 0, \ldots, 255$)	[16]	[22]
$Z_r = 0$	Positive ($r = 3, \ldots, N - 1$)	[11]	[11]
$Z_r = r$	Positive ($r = 3, \ldots, N - 1$)	[1, 6]	[6]
$Z_l = -l$	Positive (l is the keylength)	[21]	[21, 22]
$Z_{xl} = -xl$	Positive (l is the keylength)	[6]	Open (attempted in [6])

A consolidated account of the current state-of-the-art in terms of identified and/or proved short-term keystream biases of RC4 is presented in Table 1. Our motivation for this paper is to attempt proofs for all "Open" (or partially proved) problems listed in Table 1.

1.2 Contributions of our work

We can summarize the contributions of our work as follows.

- In Section 2, we prove all open isolated short-term single-byte keystream biases reported and exploited by AlFardan, Bernstein, Paterson, Poettering and Schuldt in their recent attack [1, 3] on the SSL/TLS protocol. This includes the biases in the events $Z_2 = 129$, $Z_2 = 172$, $Z_4 = 2$, $Z_{256} = 0$ and $Z_{257} = 0$.
- In Section 3, we observe that the bias of Z_1 towards 129 occurs prominently only for certain lengths of the secret key of RC4. We also discover that this bias may be related to the long-standing mysterious problem of "anomalies" in the distribution of the state array after the RC4 KSA. In this connection, we prove the anomaly in $S_0[128] = 127$, a problem open for more than a decade [13].
- In Section 4, we complete the proof for the extended keylength dependent biases in RC4, i.e., biases in the events $Z_{xl} = -xl$ for any positive integer x and keylength l. This problem was attempted and partially solved by Isobe et al. in [6]. However, the proof was left incomplete which we settle here. Note that the particular case of $x = 1$ in this class of biases reduces to the keylength-dependent biases of [22].

2 Proof of some isolated short-term biases

In this section, we prove all open isolated short-term biases of Table 1, except the case of $Z_1 = 129$. The latter case is related to the "anomaly pairs" and hence we treat it separately in Section 3.

2.1 Proof of bias in $(Z_2 = 129)$

We notice that the bias in $(Z_2 = 129)$ for $N = 256$ is a special case of the general bias in $(Z_2 = N/2 + 1)$ for any even value of N. We present the general result as follows.

Theorem 1. *Suppose that the initial permutation S_0 of RC4 PRGA is randomly chosen from the set of all permutations of $\{0, 1, \ldots, N-1\}$, where N is even. Then* $\Pr(Z_2 = N/2+1) \approx 1/N - 2/N^2$.

Proof. We consider two mutually exclusive paths from the initial state S_0.

Path 1. Consider $S_0[2] = 0$ and $S_0[1] \neq 2$. From the analysis of Mantin and Shamir [15] for the bias in $(Z_2 = 0)$, we know that $Z_2 = 0$ in this situation. Thus, $Z_2 \neq N/2 + 1$.

Path 2. Consider $S_0[2] = N/2 + 1$ and $S_0[1] \neq 2$. After the first round, $j_1 = S_0[1] = X \neq 2$, and thus $S_1[2] = N/2 + 1$ and $S_1[X] = X$. In the second round, we get $j_2 = (N/2 + 1) + X$, and let us say $S_1[j_2] = S_1[(N/2 + 1) + X] = Z$. Since S_1 is a permutation, $X = S_1[X] \neq S_1[(N/2 + 1) + X] = Z$. After the swap in the second round, we get $Z_2 = S_2[(N/2 + 1) + Z] \neq S_2[(N/2 + 1) + X] = N/2 + 1$. Figure 3 illustrates the scenario.

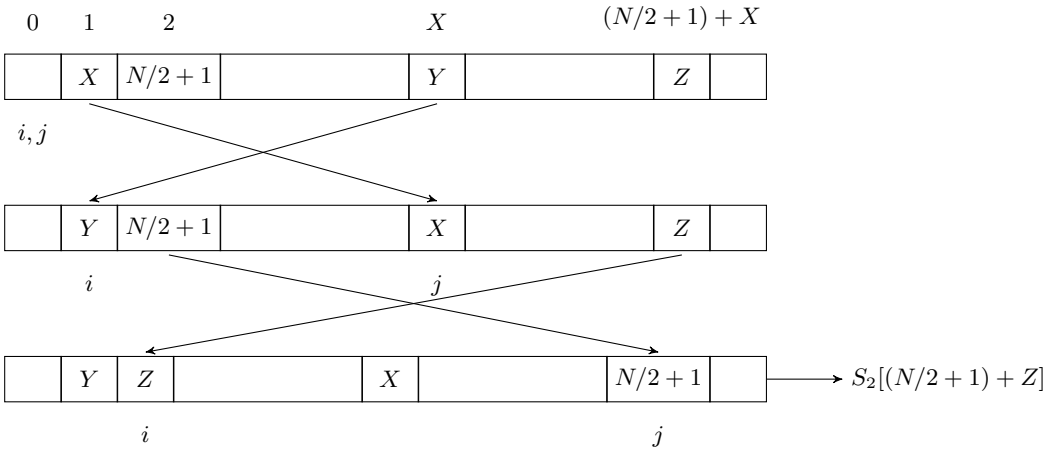

Fig. 3. The first two rounds of RC4 main cycle when $S_0[2] = N/2 + 1$ and $S_0[1] \neq 2$.

Let us denote the aforesaid mutually exclusive events as $A = (S_0[2] = 0 \wedge S_0[1] \neq 2)$ and $B = (S_0[2] = N/2 + 1 \wedge S_0[1] \neq 2)$ to obtain $\Pr(Z_2 = N/2 + 1)$ as

$$\begin{aligned}&\Pr(Z_2 = N/2 + 1 \mid A) \cdot \Pr(A) + \Pr(Z_2 = N/2 + 1 \mid B) \cdot \Pr(B) \\ &\quad + \Pr(Z_2 = N/2 + 1 \mid \overline{A} \wedge \overline{B}) \cdot \Pr(\overline{A} \wedge \overline{B}) \approx 0 + 0 + \Pr(Z_2 = N/2 + 1 \mid \overline{A} \wedge \overline{B}) \cdot (1 - 2/N).\end{aligned}$$

Assuming $\Pr(Z_2 = N/2 + 1 \mid \overline{A} \wedge \overline{B}) \approx 1/N$, due to random association, we get the desired probability as $\Pr(Z_2 = N/2 + 1) \approx (1/N) \cdot (1 - 2/N) = 1/N - 2/N^2$. □

2.2 Proof of bias in $(Z_2 = 172)$

Theorem 2. *In practical RC4 with* $N = 256$, $\Pr(Z_2 = 172) \approx 1/N + 0.28/N^2$.

Proof. We consider the following mutually exclusive paths from the initial state S_0.

Path 1. Consider $S_0[2] = 0$. If $S_0[1] \neq 2$, from the analysis of Mantin and Shamir [15] for the bias in $(Z_2 = 0)$, we know that $Z_2 = 0$ in this situation. Thus, $Z_2 \neq 172$. In case $S_0[1] = 2$, we may assume that $Z_2 = 172$ occurs with probability $1/N$. Thus, $\Pr(Z_2 = 172 \mid S_0[2] = 0) \approx 1/N^2$.

Path 2. Consider $S_0[2] = 86$. In this case, we have the following sub-paths.

1. Consider $S_0[1] = 172$. In this case, $j_1 = S_0[1] = 172$ results in a swap to produce $S_1[172] = 172$, while $S_1[2] = 86$ remains untouched. In the next round, $j_2 = j_1 + S_1[2] = 172 + 86 = 258 = 2 = i_2$ ensures that there is no swap in the S-array. Thus, $Z_2 = S_2[S_2[i_2] + S_2[j_2]] = S_1[86 + 86] = S_1[172] = 172$. Note that this path is possible for any X is $S_0[1] = X$ and $S_0[2] = X/2$, and if $X + X/2 = 2$. Thus, this path results in the modular equation $3X \equiv 4 \bmod N$, which has a unique solution $X = 172$ for $N = 256$.
2. Consider $S_0[1] \neq 172$ and $S_0[S_0[1] + 86] = 172$. In the first round, $S_1[2] = 86$ remains untouched, and $j_2 = j_1 + S_1[2] = S_0[1] + 86$ results in a swap to produce $S_2[2] = S_1[j_2] = S_1[S_0[1] + 86] = S_0[S_0[1] + 86] = 172$ and $S_2[S_0[1] + 86] = 86$. Thus, in the second round, we get $Z_2 = S_2[S_2[i_2] + S_2[j_2]] = S_2[172 + 86] = S_2[2] = 172$. Figure 4 illustrates the scenario.

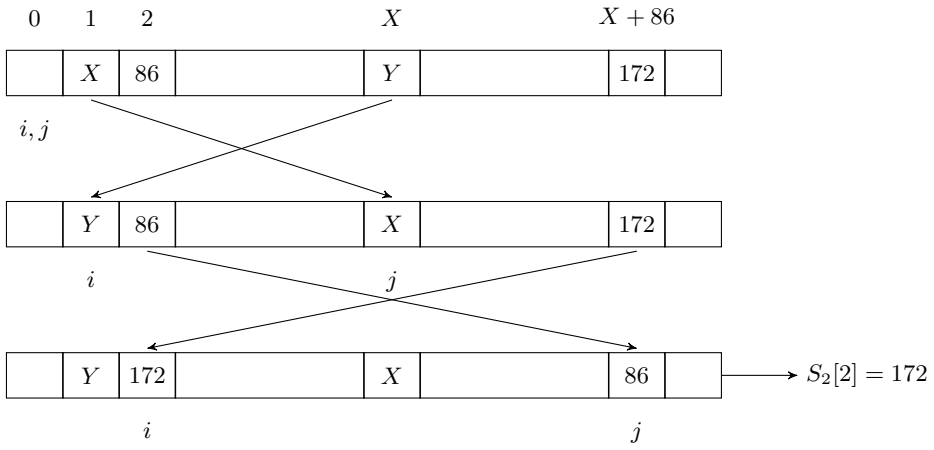

Fig. 4. The first two rounds of RC4 main cycle when $S_0[2] = 86$, $S_0[1] \neq 2, 172$ and $S_0[S_0[1] + 86] = 172$.

Let us denote the aforesaid events as $B = (S_0[2] = 86)$, $C = (S_0[1] = 172)$, and $D = (S_0[S_0[1] + 86] = 172)$. This results in

$$\begin{aligned}
&\Pr(Z_2 = 172 \mid S_0[2] = 86) = \Pr(Z_2 = 172 \mid B) \\
&= \Pr(Z_2 = 172 \mid B \wedge C) \cdot \Pr(C) + \Pr(Z_2 = 172 \mid B \wedge \overline{C}) \cdot \Pr(\overline{C}) \\
&\approx 1 \cdot (1/N) + \big(\Pr(Z_2 = 172 \mid B \wedge \overline{C} \wedge D) \cdot \Pr(D) \\
&\qquad + \Pr(Z_2 = 172 \mid B \wedge \overline{C} \wedge \overline{D}) \cdot \Pr(\overline{D})\big) \cdot (1 - 1/N) \\
&\approx (1/N) + (1 \cdot (1/N) + (1/N) \cdot (1 - 1/N)) \cdot (1 - 1/N) \approx 3/N - 3/N^2.
\end{aligned}$$

Path 3. Consider $S_0[2] = 172$. In this situation, $Z_2 = 172$ if $S_0[1] = 2$ and $S_0[4] = N - 1$, and in all other cases, $Z_2 \neq 172$. Thus, $\Pr(Z_2 = 172 \mid S_0[2] = 172) \approx 1/N^2$.

Let us combine the aforesaid paths to obtain $\Pr(Z_2 = 172)$ as

$$\begin{aligned}&\Pr(Z_2 = 172 \mid S_0[2] = 0) \cdot \Pr(S_0[2] = 0) + \Pr(Z_2 = 172 \mid S_0[2] = 86) \cdot \Pr(S_0[2] = 86) \\ &\quad + \Pr(Z_2 = 172 \mid S_0[2] = 172) \cdot \Pr(S_0[2] = 172) \\ &\approx (1/N^2) \cdot \Pr(S_0[2] = 0) + (3/N - 3/N^2) \cdot \Pr(S_0[2] = 86) + (1/N^2) \cdot \Pr(S_0[2] = 172).\end{aligned}$$

In the above equation, computing the probability terms involving S_0 using the formula of Mantin [13], we get $\Pr(Z_2 = 172) \approx 1/N + 0.28/N^2$. □

2.3 Proof of bias in $(Z_4 = 2)$

Theorem 3. *Suppose that the initial permutation S_0 of RC4 PRGA is randomly chosen from the set of all permutations of $\{0, 1, \ldots, N-1\}$, where $N = 256$. Then $\Pr(Z_4 = 2) \approx 1/N + 1/N^2$.*

Proof. We observe the main paths for this bias as follows.

Path 1. Consider $j_4 = 4$. Then, $Z_4 = S_4[S_4[4] + S_4[j_4]] = S_4[2 \cdot S_4[4]]$. We may further consider some subpaths within this case.

- Subpath 1: $S_4[4] = 2$ gives $Z_4 = S_4[4] = 2$ with probability 1. However, the event $(S_4[4] = 2 \mid j_4 = 4)$ occurs with probability approximately $2/N$, as follows.
  - If $S_0[4] = 2$ and $j_1, j_2, j_3 \neq 4$, then $S_4[4] = 2$ remains constant during the first three rounds. Thus, $\Pr(S_4[4] = 2 \mid j_4 = 4 \wedge S_0[4] = 2) \approx 1$.
  - If $S_0[1] = 2$ and $j_3 \neq 4$, it can be shown that $S_4[4] = 2$ through the first three rounds. Thus, $\Pr(S_4[4] = 2 \mid j_4 = 4 \wedge S_0[1] = 2) \approx 1$.
  - Consider $S_0[4] \neq 2$ and $S_0[1] \neq 2$. In this case, we show that $S_4[4] = 2$ and $j_4 = 4$ can not occur simultaneously. Suppose the event $(j_4 = 4 \wedge S_4[4] = 2)$ does occur. Then we have $j_3 = 0$, and hence $S_2[4] = S_3[4] = S_4[4] = 2$. As we know $S_0[4] \neq 2$, this implies $j_1 = 4$ and/or $j_2 = 4$.
    If $j_1 = 4$ and $j_2 = 4$, we get $S_2[4] = 0$, contradiction, as $S_2[4] = 2$.
    If $j_1 = 4$ and $j_2 \neq 4$, we have $S_1[4] = S_2[4] = 4$, contradiction, as $S_2[4] = 2$.
    Consider $j_1 \neq 4$ and $j_2 = 4$. Since $j_2 = S_0[1] + S_1[2]$ and $S_0[1] \neq 2$, we must have $S_2[4] = S_1[2] \neq 2$, a contradiction.

  In summary, $\Pr(S_4[4] = 2 \mid j_4 = 4) \approx 2/N$, and $\Pr(Z_4 = 2 \wedge S_4[4] = 2 \mid j_4 = 4) \approx 2/N$.
- Subpath 2: $S_4[4] = N/2 + 2$ gives $Z_4 = S_4[N + 4] = S_4[4] = N/2 + 2 \neq 2$. So, $\Pr(Z_4 = 2 \wedge S_4[4] = N/2 + 2 \mid j_4 = 4) = 0$.
- Subpath 3: $S_4[4] = 0$ gives $Z_4 = S_4[0] = 2$ with probability $1/N$. However, the event $(S_4[4] = 0 \mid j_4 = 4)$ occurs with probability $2/N$, as follows.
  - If $S_0[1] = 2$ and $S_0[3] = 0$, then $j_1 = 2$, $S_1[2] = S_0[1] = 2$, which implies $j_2 = 4$. This produces $S_2[4] = S_1[2] = 2$, and we get $j_3 = j_2 = 4$ because of $S_2[3] = S_1[3] = S_0[3] = 0$. After the third round, $S_3[4] = S_2[3] = 0$, and in the next round, $j_4 = j_3 = 4$ ensures no swap. Thus, we get both $j_4 = 4$ and $S_4[4] = S_3[4] = 0$.
  - In other situations, $S_4[4] = 0$ and $j_4 = 4$ occur due to random situation.

  In summary, $\Pr(S_4[4] = 0 \mid j_4 = 4) \approx 2/N$, and $\Pr(Z_4 = 2 \wedge S_4[4] = 0 \mid j_4 = 4) \approx 2/N^2$.
- Subpath 4: $S_4[4] \neq 0, 2, N/2 + 2$ gives $Z_4 = 2$ with probability approximately $1/N$ due to random association. Due to the previous subpaths, we know that the event $(S_4[4] \neq 0, 2, N/2 + 2 \mid j_4 = 4)$ occurs with probability $(1 - 5/N)$. Thus,

$$\Pr(Z_4 = 2 \wedge S_4[4] \neq 0, 2, N/2 + 2 \mid j_4 = 4) \approx (1/N) \cdot (1 - 5/N) = 1/N - 5/N^2.$$

Combining all the subpaths mentioned above, we get $\Pr(Z_4 = 2 \wedge j_4 = 4)$ as

$$\begin{aligned}
&\Pr(Z_4 = 2 \wedge S_4[4] = 2 \mid j_4 = 4) \cdot \Pr(j_4 = 4) \\
&\quad + \Pr(Z_4 = 2 \wedge S_4[4] = N/2 + 2 \mid j_4 = 4) \cdot \Pr(j_4 = 4) \\
&\quad + \Pr(Z_4 = 2 \wedge S_4[4] = 0 \mid j_4 = 4) \cdot \Pr(j_4 = 4) \\
&\quad + \Pr(Z_4 = 2 \wedge S_4[4] \neq 0, 2, N/2 + 2 \mid j_4 = 4) \cdot \Pr(j_4 = 4) \\
&= (2/N) \cdot (1/N) + 0 + (2/N^2) \cdot (1/N) + (1/N - 5/N^2) \cdot (1/N) = 3/N^2 - 3/N^3.
\end{aligned}$$

Path 2. Consider $j_4 \neq 4$. Then, $Z_4 = S_4[S_4[4] + S_4[j_4]] = S_4[S_4[4] + X]$, where $X = S_4[j_4] \neq S_4[4]$, say. Here we may consider two subpaths, as follows.

- Subpath 1: $S_4[4] = 2$ gives $Z_4 = S_4[2 + X] \neq S_4[4] = 2$, as $X = S_4[j_4] \neq S_4[4] = 2$ for $j_4 = 4$. Thus, $\Pr(Z_4 = 2 \wedge S_4[4] = 2 \mid j_4 \neq 4) = 0$.
- Subpath 2: $S_4[4] \neq 2$ gives $Z_4 = 2$ with due to random association. Thus, $\Pr(Z_4 = 2 \wedge S_4[4] \neq 2 \mid j_4 \neq 4) \approx 1/N \cdot (1 - 1/N) = (1/N - 1/N^2)$.

Combining the subpaths mentioned above, we have $\Pr(Z_4 = 2 \wedge j_4 \neq 4)$ as

$$\begin{aligned}
&\Pr(Z_4 = 2 \wedge S_4[4] = 2 \mid j_4 \neq 4) \cdot \Pr(j_4 \neq 4) + \Pr(Z_4 = 2 \wedge S_4[4] \neq 2 \mid j_4 \neq 4) \cdot \Pr(j_4 \neq 4) \\
&= 0 + (1/N - 1/N^2) \cdot (1 - 1/N) = 1/N - 2/N^2 + 1/N^3.
\end{aligned}$$

Adding the contributions from the two mutually exclusive paths above, we get

$$\begin{aligned}
\Pr(Z_4 = 2) &= \Pr(Z_4 = 2 \wedge j_4 = 4) + \Pr(Z_4 = 2 \wedge j_4 \neq 4) \\
&= (3/N^2 - 3/N^3) + (1/N - 2/N^2 + 1/N^3) = 1/N + 1/N^2 - 2/N^3.
\end{aligned}$$

Hence we get $\Pr(Z_4 = 2) \approx 1/N + 1/N^2$. □

2.4 Proof of bias in $(Z_{256} = 0)$

Theorem 4. *In practical RC4 with $N = 256$, $\Pr(Z_N = 0) \approx 1/N - 0.36/N^2$.*

Proof. Let us consider the following two paths.

Path 1. Consider $S_1[0] = 0$. In this case, if $j_2, \ldots, j_{N-1}$ are all non zero, then one can check that $Z_N \neq 0$. In all other cases, one may consider $\Pr(Z_N = 0 \mid S_1[0] = 0) \approx 1/N$ due to random association. Thus, $\Pr(Z_N = 0 \mid S_1[0] = 0) \approx \left(1 - (1 - 1/N)^{N-2}\right) \cdot (1/N)$.

Path 2. Consider $S_1[0] \neq 0$. In this case, we may again consider the following sub-paths, depending on the state S_{N-3}.

$$\begin{aligned}
&\Pr(Z_N = 0 \mid S_1[0] \neq 0) \\
&= \Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[0] = 0) \cdot \Pr(S_{N-3}[0] = 0 \mid S_1[0] \neq 0) \\
&\quad + \Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-2] = 0) \cdot \Pr(S_{N-3}[N-2] = 0 \mid S_1[0] \neq 0) \\
&\quad + \Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-1] = 0) \cdot \Pr(S_{N-3}[N-1] = 0 \mid S_1[0] \neq 0) \\
&\quad + \sum_{x=1}^{N-3} \Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[x] = 0) \cdot \Pr(S_{N-3}[x] = 0 \mid S_1[0] \neq 0).
\end{aligned}$$

- Case 1: If $S_{N-3}[0] = 0$ and $j_{N-2}, j_{N-1} \neq 0$, we have $S_{N-1}[0] = 0$, which implies $j_N = j_{N-1}$ and $S_{N-1}[j_{N-1}] \neq j_{N-1}$. Thus, $Z_N = S_N[S_{N-1}[j_N] + S_{N-1}[0]] = S_N[S_{N-1}[j_{N-1}]] \neq S_N[j_{N-1}] = S_N[j_N] = S_{N-1}[0] = 0$. Thus for $Z_N = 0$, we must have either $j_{N-2} = 0$ or $j_{N-1} = 0$ in this case, and in each case, $Z_N = 0$ will occur with probability $1/N$ of random association. Hence $\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[0] = 0) \approx 2/N^2$.

– Case 2: If $S_{N-3}[N-2] = 0$ and $j_{N-2} = 0$, we have $S_{N-2}[0] = 0$ and $j_{N-1} = S_{N-2}[N-1] \neq 0$. Thus, $S_{N-1}[0] = 0$ and $j_N = j_{N-1}$, which gives $Z_N = S_N[S_{N-1}[0] + S_{N-1}[j_N]] = S_N[S_{N-1}[j_{N-1}]] = S_N[S_{N-2}[N-1]] = S_N[j_{N-1}] = S_N[j_N] = S_{N-1}[0] = 0$. So, $\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-2] = 0 \wedge j_{N-2} = 0) = 1$. On the other hand, if $S_{N-3}[N-2] = 0$ and $j_{N-2} \neq 0$, then $Z_N \neq 0$ where $j_{N-1} \neq 0$ and $S_{N-1}[j_N] = 0$, and $Z_N = 0$ due to random association in all other cases. So, $\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-2] = 0 \wedge j_{N-2} \neq 0) \approx 1/N - 1/N^2$. Combining the two items as above, we get

$$\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-2] = 0) \approx 2/N - 2/N^2.$$

– Case 3: Similarly for $S_{N-3}[N-1] = 0$, it can be proved that $\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[N-1] = 0) \approx 2/N - 2/N^2$.
– Case 4: Now consider the case $S_{N-3}[x] = 0$ for $1 \leq x \leq N-3$. If $j_{N-2} \neq x$, $j_{N-1} \neq x$ and $j_N = x$, one can verify that $Z_N \neq 0$. In all other cases, $Z_N = 0$ occurs with probability $1/N$. Thus for $1 \leq x \leq N-3$, $\Pr(Z_N = 0 \mid S_1[0] \neq 0 \wedge S_{N-3}[x] = 0) \approx 1/N - 1/N^2$.

Now, let us consider the conditional events $(S_{N-3}[x] = 0 \mid S_1[0] \neq 0)$, for $0 \leq x \leq N-1$, to complete the picture. Starting with $S_1[0] \neq 0$, if $j_2, \ldots, j_{N-3}$ are all non zero, we have $S_{N-3}[0] \neq 0$ as well. So, $\Pr(S_{N-3}[0] = 0 \mid S_1[0] \neq 0) = \left(1 - (1-1/N)^{N-4}\right) \cdot (1/N) = P_A$, say. For all $x \neq 0$, we may now assume $\Pr(S_{N-3}[x] = 0 \mid S_1[0] \neq 0) \approx (1 - P_A)/(N-1) = P_B$, say. Taking into account the contributions from all four sub-cases within this path, we get

$$\begin{aligned}\Pr(Z_N = 0 \mid S_1[0] \neq 0) &= (2/N^2) \cdot P_A + (2/N - 2/N^2) \cdot P_B \\ &\quad + (2/N - 2/N^2) \cdot P_B + (1/N - 1/N^2) \cdot (1 - P_A - 2P_B) \\ &= (1/N - 1/N^2) - (1/N - 3/N^2) \cdot P_A + (2/N - 2/N^2) \cdot P_B.\end{aligned}$$

Combining the above two paths, we get $\Pr(Z_N = 0)$ as

$$\begin{aligned}&\Pr(Z_N = 0 \mid S_1[0] = 0) \cdot P(S_1[0] = 0) + \Pr(Z_N = 0 \mid S_1[0] \neq 0) \cdot P(S_1[0] \neq 0) \\ &\approx \left(1 - (1-1/N)^{N-2}\right) \cdot (1/N) \cdot (2/N) \\ &\quad + ((1/N - 1/N^2) - (1/N - 3/N^2) \cdot P_A + (2/N - 2/N^2) \cdot P_B) \cdot (1 - 2/N) \approx 1/N - 0.36/N^2,\end{aligned}$$

for $N = 256$, as in the case with practical RC4. □

2.5 Proof of bias in $(Z_{257} = 0)$

Theorem 5. *In practical RC4 with* $N = 256$, $\Pr(Z_{N+1} = 0) \approx 1/N + 0.36/N^2$.

Proof. We may write $Z_{N+1} = S_{N+1}[S_N[1] + S_N[j_{N+1}]]$, and consider the following two paths.

Path 1. Consider the case $S_N[1] = 1$, where we may write $Z_{N+1} = S_{N+1}[1 + S_N[j_{N+1}]]$. If $S_N[j_{N+1}] = 0$, we have $Z_{N+1} = S_{N+1}[1] = S_N[j_{N+1}] = 0$. Otherwise if $S_N[j_{N+1}] = X \neq 0$, we have $Z_{N+1} = S_{N+1}[1 + X] = 0$ only due to random association. Let us denote events $A = (S_N[1] = 1)$ and $B = (S_N[j_{N+1}] = 0)$ to get

$$\begin{aligned}\Pr(Z_{N+1} = 0 \mid A) &= \Pr(Z_{N+1} = 0 \mid A \wedge B) \cdot \Pr(B) + \Pr(Z_{N+1} = 0 \mid A \wedge \overline{B}) \cdot \Pr(\overline{B}) \\ &\approx 1 \cdot (1/N) + (1/N) \cdot (1 - 1/N) = 2/N - 1/N^2.\end{aligned}$$

Path 2. Consider the case $S_N[1] = X \neq 1$. Here we have $Z_{N+1} = S_{N+1}[X + S_N[j_{N+1}]]$. If $S_N[j_{N+1}] = 0$, we will get $Z_{N+1} = S_{N+1}[X] \neq S_{N+1}[1] = S_N[j_{N+1}] = 0$. Otherwise, for $S_N[j_{N+1}] = Y \neq 0$, we may have $Z_{N+1} = S_{N+1}[X + Y] = 0$ due to random association. Let us denote events $A = (S_N[1] = 1)$ and $B = (S_N[j_{N+1}] = 0)$ to get

$$\begin{aligned}\Pr(Z_{N+1} = 0 \mid \overline{A}) &= \Pr(Z_{N+1} = 0 \mid \overline{A} \wedge B) \cdot \Pr(B) + \Pr(Z_{N+1} = 0 \mid \overline{A} \wedge \overline{B}) \cdot \Pr(\overline{B}) \\ &\approx 0 + (1/N) \cdot (1 - 1/N) = 1/N - 1/N^2.\end{aligned}$$

From [22, Theorem 1], we have $\Pr(S_N[1] = 1) \approx 0.00532$ when $N = 256$. Thus,

$$\begin{aligned}\Pr(Z_{N+1} = 0) &= \Pr(Z_{N+1} = 0 \mid A) \cdot \Pr(A) + \Pr(Z_{N+1} = 0 \mid \overline{A}) \cdot \Pr(\overline{A}) \\ &\approx (2/N - 1/N^2) \cdot (0.00532) + (1/N - 1/N^2) \cdot (1 - 0.00532) \approx 1/N + 0.36/N^2,\end{aligned}$$

for $N = 256$, as in the case with practical RC4. □

3 Negative bias in $Z_1 = 129$ and the anomaly in $S_0[128] = 127$

In this section, we attempt at solving the mystery of the negative bias in $Z_1 = 129$, which was observed in [1, 3], but not in [16, 22]. We first notice that the length of the secret key used in the experiments of [1,3] was consistently $l = 16$, whereas the same for [16,22] might have been different. This hinted that the bias in $Z_1 = 129$ may be keylength dependent. Our experiments revealed that the negative bias of $Z_1 = 129$ is prominent only for keylength l equal to non-trivial factors of 256, that is, for $l = 2, 4, 8, 16, 32, 64, 128$. This behavior is depicted in Fig. 5.

Dependence of keystream biases on the secret keylength l was first proved in [22], for any keylength l, but no such pattern for specific keylengths was discovered earlier. Our experiments with these specific keylengths $l = 2, 4, \ldots, 128$ revealed that there exists another bias of the same kind, a negative bias in $S_0[128] = 127$. Figure 6 shows the keylength dependence of this bias. This bias had been pointed out quite a few years ago [13, 18] as an "anomaly" in the otherwise smooth distribution of $S_0[u] = v$, but it was never observed as a keylength dependent phenomenon.

In this section, we first settle the mysterious open issue of the $S_0[128] = 127$ anomaly, and then proceed to analyze its connection with the negative bias of $Z_1 = 129$, if any. We will require the following technical results to prove the main theorem later.

Lemma 1. *In practical RC4 with $N = 256$, for $1 \leq r \leq N$, $\Pr(S^K_{r-1}[r] = r) \approx 1/N + (1 - 1/N)^r$.*

Proof. We know that S^K_0 is the identity permutation of $\{0, \ldots, N-1\}$, and thus $S^K_0[r] = r$. This value will remain at the same index till round $(r-1)$ if none of $j^K_1, j^K_2, \ldots, j^K_{r-1}$ touches the index r, which occurs with probability $(1-1/N)^{r-1}$, or otherwise due to random association, with probability $1/N$. Hence, we get $\Pr(S^K_{r-1}[r] = r) \approx (1-1/N)^{r-1} \cdot 1 + (1-(1-1/N)^{r-1}) \cdot (1/N) = 1/N + (1-1/N)^r$.

Lemma 2. *In practical RC4 with $N = 256$, $\Pr(S^K_{127}[128] = -K[128]) \approx 0.4/N$ if and only if l, the length of the RC4 secret key, is a non-trivial factor of $N = 256$.*

Proof. Let us consider the following two paths.

Path 1. Consider $S^K_{127}[128] = 128$. In this case, we surely require $K[128] = -128 = 128$ (modulo $N = 256$). Now, if $l = 2, 4, \ldots, 128$, then $K[128] = K[0] = 128$. This implies $j^K_1 = j^K_0 + S^K_0[0] + K[0] = 0 + 0 + 128 = 128$, which in turn results in $S^K_1[0] = 128$ and $S^K_1[128] = 0$ after swap in the first round. As i^K does not touch index locations 0 or 128 during rounds 2 to 127, we can not have $S^K_{127}[128] = 128$, a contradiction. If l does not divide 128, then $K[128]$

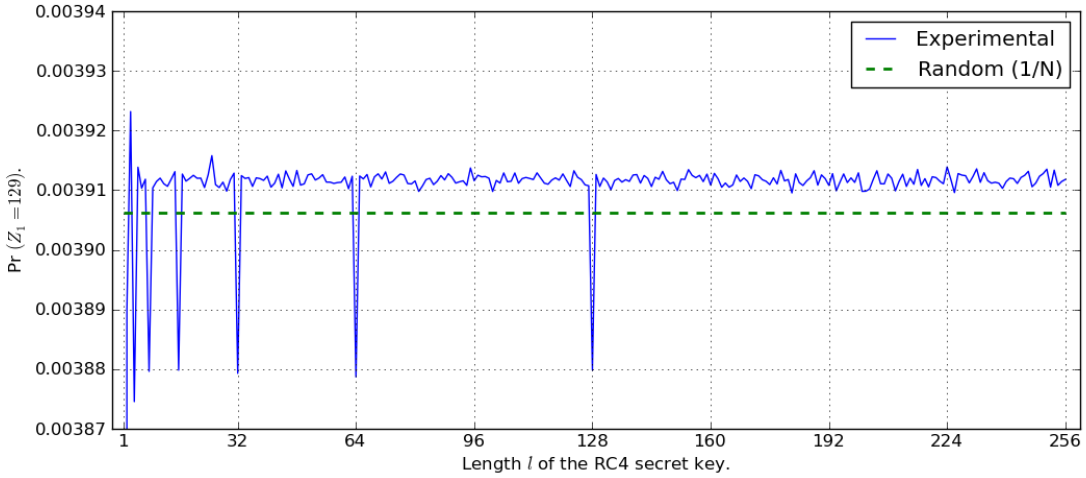

Fig. 5. Bias in the event $(Z_1 = 129)$ for keylength $1 \leq l \leq 256$.

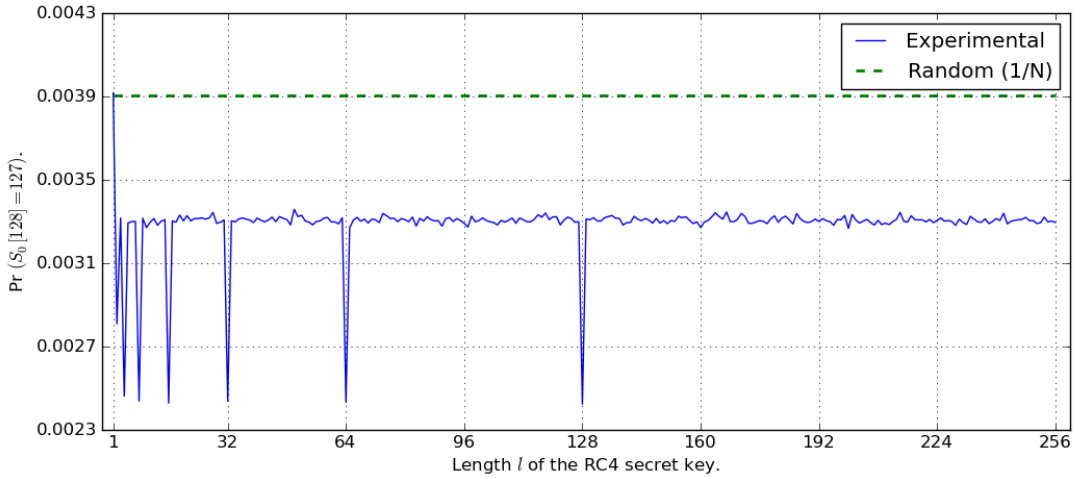

Fig. 6. Bias in the event $(S_0[128] = 127)$ for keylength $1 \leq l \leq 256$.

may not be equal to $K[0]$, and in this case $S_{127}^K[128] = 128$ may occur with probability $1/N$. In summary, $\Pr(S_{127}^K[128] = -K[128] \mid S_{127}^K[128] = 128) = 0$ if $l = 2, 4, \ldots, 128$. Otherwise, $\Pr(S_{127}^K[128] = -K[128] \mid S_{127}^K[128] = 128) \approx 1/N$.

Path 2. In case $S_{127}^K[128] \neq 128$, there is no special behavior dependent on the keylength l, and we may assume that $\Pr(S_{127}^K[128] = -K[128] \mid S_{127}^K[128] \neq 128) \approx 1/N$.

Combining the two paths, we get

$$\begin{aligned}\Pr(S_{127}^K[128] = -K[128]) &= \Pr(S_{127}^K[128] = -K[128] \mid S_{127}^K[128] = 128) \cdot \Pr(S_{127}^K[128] = 128) \\ &\quad + \Pr(S_{127}^K[128] = -K[128] \mid S_{127}^K[128] \neq 128) \cdot \Pr(S_{127}^K[128] = 128) \\ &\approx 0 \cdot (156/N) + (1/N) \cdot (1 - 156/N) \approx 0.4/N,\end{aligned}$$

if $l = 2, 4, \ldots, 128$, where $\Pr(S_{127}^K[128] = 128) \approx 156/N$ is by Lemma 1 with $r = 128$. For all other values of l, we get $\Pr(S_{127}^K[128] = -K[128]) \approx (1/N) \cdot (156/N) + (1/N) \cdot (1 - 156/N) = 1/N$. □

Theorem 6. *In practical RC4 with $N = 256$, $\Pr(S_0[128] = S_N^K[128] = 127) \approx 0.63/N$ if and only if l, the length of the RC4 secret key, is a non-trivial factor of $N = 256$.*

Proof. Let us first compute $\Pr(S_{128}^K[128] = 127)$, using the following paths.

Path 1. Consider $S_{127}^K[128] = -K[128]$. In this case, $j_{128} = j_{127} + S_{127}^K[128] + K[128] = j_{127}$. So, $S_{128}^K[128] = S_{127}^K[j_{128}] = S_{127}^K[j_{127}] = S_{126}^K[127]$. Now, by Lemma 1 with $r = 127$, we get $\Pr(S_{126}^K[127] = 127) \approx 156/N$. Thus, $\Pr(S_{128}^K[128] = 127 \mid S_{127}^K[128] = -K[128]) \approx 156/N$.

Path 2. Consider $S_{127}^K[128] \neq -K[128]$. In this case, $S_{128}^K[128] = S_{126}^K[X]$ for some $X \neq 127$. Thus by normalization over the probability values $\Pr(S_{126}^K[X] = 127)$ for $X \neq 127$, we get $\Pr(S_{128}^K[128] = 127 \mid S_{127}^K[128] \neq -K[128]) \approx (1 - 156/N)/(N-1) \approx 0.4/N$.

Combining the two paths as above, we get

$$\begin{aligned}\Pr(S_{128}^K[128] = 127) &= \Pr(S_{128}^K[128] = 127 \mid S_{127}^K[128] = -K[128]) \cdot \Pr(S_{127}^K[128] = -K[128]) \\ &\quad + \Pr(S_{128}^K[128] = 127 \mid S_{127}^K[128] \neq -K[128]) \cdot \Pr(S_{127}^K[128] \neq -K[128]) \\ &\approx (156/N) \cdot (0.4/N) + (0.4/N) \cdot (1 - 0.4/N) \approx 0.64/N,\end{aligned}$$

if $l = 2, 4, \ldots, 128$. For all other values of l, we get $\Pr(S_{128}^K[128] = 127) \approx (156/N) \cdot (1/N) + (0.4/N) \cdot (1 - 1/N) \approx 1/N$. In both cases, the value of $\Pr(S_{127}^K[128] = -K[128])$ comes from Lemma 2.

Once we have $S_{128}^K[128] = 127$, we know that $S_0[128] = S_N^K[128] = 127$ if none of $j_{129}, \ldots, j_N$ touches the index 128. If otherwise $S_{128}^K[128] \neq 127$ and the value 127 is in any index less than 128, then $S_N^K[128] \neq 127$. If $S_{128}^K[128] \neq 127$ and the value 127 is in any index I greater than 128, then $S_N^K[128] = 127$ may occur due to the following association.

- Indices $j_{129}, \ldots, j_{I-1}$ do not touch location I before $i = I$.
- When $i = I$, we have j equal to 128, so that the appropriate swap occurs.
- Moreover, none of $j_{I+1}, \ldots, j_N$ touches the location 128 after the previous event.

This path entails a approximate probability $(1/N) \cdot (1 - 1/N)^{127}$ for each I, and the total probability of the aforesaid association, over $I = 129, \ldots, 255$, becomes approximately $0.24/N$. Thus,

$$\begin{aligned}\Pr(S_N^K[128] = 127) &= \Pr(S_{128}^K[128] = 127) \cdot (1 - 1/N)^{128} + \Pr(S_{128}^K[128] \neq 127) \cdot (0.24/N) \\ &\approx (0.64/N) \cdot (155/N) + (1 - 0.64/N) \cdot (0.24/N) \approx 0.63/N,\end{aligned}$$

if $l = 2, 4, \ldots, 128$. For other values of l, we get $\Pr(S_0[128] = S_N^K[128] = 127)$ following the value predicted by the distribution of $S_0[u] = v$ by Mantin [13, 15]. Hence the "anomaly". □

The theoretical results regarding the anomaly in $S_0[128] = 127$, as above, closely match with the experimental results, both from our own experiments, as well as that reported in the literature [18]. This settles a long-standing mysterious issue in RC4 literature, and hints at the possibility that all "anomalies" or deviations of probabilities in the distribution of S_0 from that predicted by Mantin [13], may actually result from intricate keylength dependence in the cipher.

Experimentally, we find that $\Pr(Z_1 = 129 \mid S_0[128] = 127) \approx 1/N - 0.5/N^2$ and $\Pr(Z_1 = 129 \mid S_0[128] \neq 127) \approx 1/N - 2/N^2$. Thus, using the anomaly, one may estimate $\Pr(Z_1 = 129)$ as

$$\begin{aligned}&\Pr(Z_1 = 129 \mid S_0[128] = 127) \Pr(S_0[128] = 127) + \Pr(Z_1 = 129 \mid S_0[128] \neq 127) \Pr(S_0[128] \neq 127) \\ &\approx (1/N - 0.5/N^2) \cdot (0.63/N) + (1/N - 2/N^2) \cdot (1 - 0.63/N) \approx 1/N - 2/N^2.\end{aligned}$$

It may be the case that the anomaly is not directly influencing the bias in $Z_1 = 129$. However, investigation and proof of all causal paths towards proving the negative bias in $Z_1 = 129$ remains an interesting open question, which requires an independent rigorous analysis.

4 Complete proof of generalized keylength dependent biases

In [22, Section 2], Sen Gupta et al. presented a family of biases in RC4 that are dependent on the length of the secret key. The most important of those biases was a keylength-distinguisher based on the positive bias in the event $(Z_l = -l)$, where l is the length of RC4 secret key in bytes.

Subsequently, in [6, Section 3.4], Isobe et al. observed that similar bias also exists in the class of events $(Z_{xl} = -xl)$ for any positive integer x. In an attempt to prove these biases, they explored certain paths involving the expression $f_y = y(y+1)/2 + \sum_{x=0}^{y} K[x]$. However, they could not prove all the paths and substituted experimental values to compute what they referred as *semitheoretical values*. They also commented the following.

> Since semitheoretical value are partially based on experimental results, we can not claim that the proof of these bias are given.

We observe that instead of following the approach of [6], if one follows the approach in [22], then the theoretical derivation of the generalized keylength-dependent biases become much simpler. In this section, we generalize all the keylength-dependent biases of [22] for any keylength $l \in [3, N-1]$ and any integer $x \in [1, \lfloor \frac{N}{l} \rfloor]$ and thereby complete the proof of extended keylength distinguisher that was left open in [6]. As a result, the biases in [22] become special cases of our results here with $x = 1$. Note that though the general proof follows the same approach as in [22], the extension is not obvious. A general proof always imply the special cases, but the converse need not be true. We experimentally verified all the intermediate claims and assumptions related to the events involving "xl" and we found them to be consistent with our theoretical claims. We present the general theorems below and present the proofs in Appendix A for the sake of completeness.

All the biases that we are interested in are related to $(S^K_{xl+1}[xl-1] = -xl \wedge S^K_{xl+1}[xl] = 0)$, where x is an integer between 1 and $\lfloor \frac{N}{l} \rfloor$. So we first derive the probability for this event in Lemma 3.

Lemma 3. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor \frac{N}{l} \rfloor$, we have* $\Pr(S^K_{xl+1}[xl-1] = -xl \wedge S^K_{xl+1}[xl] = 0) \approx \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right) \alpha_{x,l}$, *where* $\alpha_{x,l} = \frac{1}{N}\left(1 - \frac{3}{N}\right)^{xl-2}\left(1 - \frac{xl+1}{N}\right)$.

Theorem 7. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor \frac{N}{l} \rfloor$, we have* $\Pr(S_{xl}[xl] = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) = \Pr(t_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) \approx \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right)\beta_{x,l}$, *where* $\beta_{x,l} = \frac{1}{N}\left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)^{N-3}\left(1 - \frac{3}{N}\right)^{xl-2}\left(1 - \frac{xl+1}{N}\right)$.

Theorem 8. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor \frac{N}{l} \rfloor$, we have* $\Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) \approx \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right)\gamma_{x,l}$, *where*
$\gamma_{x,l} = \frac{1}{N^2}\left(1 - \frac{xl+1}{N}\right)\sum_{u=xl+1}^{N-1}\left(1 - \frac{1}{N}\right)^{u}\left(1 - \frac{2}{N}\right)^{u-xl}\left(1 - \frac{3}{N}\right)^{N-u+2xl-4}$.

Theorem 9. *For any keylength $l \in [3, N-1]$ and any integer $x \in [1, \lfloor \frac{N}{l} \rfloor]$, the probability* $\Pr(S_{xl}[j_{xl}] = 0)$ *is given by*
$\delta_{x,l} \approx \Pr(S_1[xl] = 0)\left(1 - \frac{1}{N}\right)^{xl-2} + \sum_{y=2}^{xl-1}\sum_{w=0}^{xl-y}\frac{\Pr(S_1[y]=0)}{w! \cdot N}\left(\frac{xl-y-1}{N}\right)^{w}\left(1 - \frac{1}{N}\right)^{xl-3-w}$.

Theorem 10. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor \frac{N}{l} \rfloor$, we have* $\tau_{x,l} = \Pr(t_{xl} = -xl) \approx \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right)\beta_{x,l} + (1 - \delta_{x,l})\frac{1}{N}$, *where $\beta_{x,l}$ is given in Theorem 7 and $\delta_{x,l}$ is given in Theorem 9.*

Theorem 11. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor \frac{N}{l} \rfloor$,*

$$\Pr(Z_{xl} = -xl) \ \approx \ \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right)\gamma_{x,l} + (1 - \delta_{x,l})\frac{1}{N},$$

where $\gamma_{x,l}$ is given in Theorem 8 and $\delta_{x,l}$ is given in Theorem 9.

In Figure 7, we compare the experimental values of $(Z_{xl} = -xl)$, obtained from the data of [1,3], with our theoretical values derived from Theorem 11, for keylength $l = 16$ and $x = 1, 2, \ldots, 15$. We have obtained similar results for other keylengths as well, and some figures are in Appendix A.1.

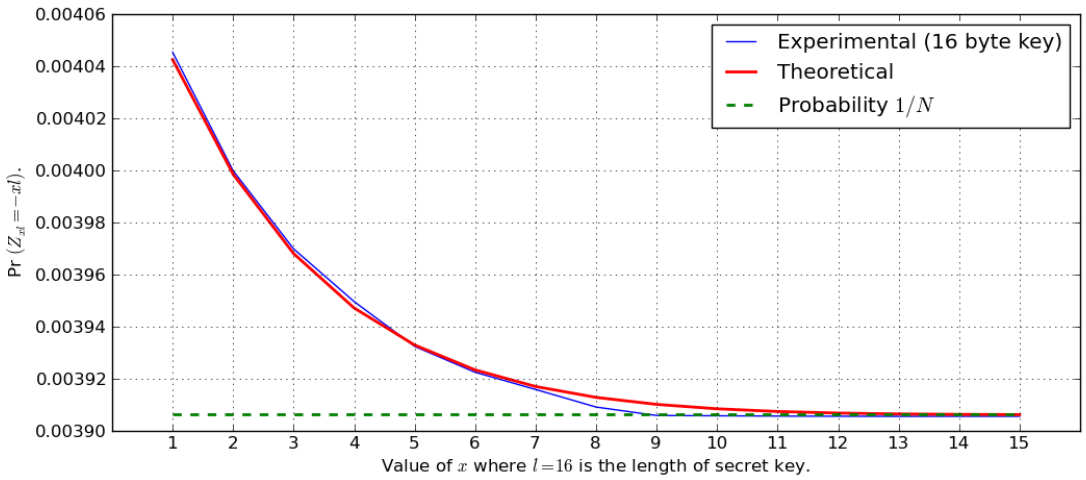

Fig. 7. Bias in the event $(Z_{xl} = -xl)$ for keylength $l = 16$ and $x = 1, 2, \ldots, 15$.

5 Conclusion

We have proved almost all open short-term single-byte biases that have been exploited in the recent TLS attack [1,3]. We have also given complete proof of 'extended keylength biases' from [6]. Table 2 compares the experimental data of [1, 3, 6] to our theoretical results.

Table 2. Proved short-term single-byte keystream biases of RC4.

Bias in event	Discovered	Theoretical proof (this paper)	Experimental [1, 3, 6]
$Z_2 = 129$	[1, 20]	$1/N - 2/N^2$	$1/N - 1.82/N^2$
$Z_2 = 172$	[1]	$1/N + 0.28/N^2$	$1/N + 0.2/N^2$
$Z_4 = 2$	[1]	$1/N + 1/N^2$	$1/N + 0.8/N^2$
$Z_{256} = 0$	[1, 6]	$1/N - 0.36/N^2$	$1/N - 0.38/N^2$
$Z_{257} = 0$	[6]	$1/N + 0.36/N^2$	$1/N + 0.35/N^2$
$Z_{xl} = -xl$	[6]	Theorem 11	Figure 7 for $l = 16$

We also attempted the proof of the bias in $Z_1 = 129$, but could not settle it completely. However, we discovered that this bias is a new 'keylength dependent' bias of RC4, which is prominent only for certain keylengths $l = 2, 4, 8, \ldots, 128$. In the process, we tried to relate it with the long-standing open issue of 'anomalies' in RC4 initial state, and could prove an important anomaly regarding the bias in $S_0[128] = 127$. Our work reveals that a thorough analysis of the "anomaly pairs" is necessary, not only for their independent theoretical interest, but also to investigate their potential implications towards keystream biases.

References

1. Nadhem AlFardan, Dan Bernstein, Kenny Paterson, Bertram Poettering, and Jacob Schuldt. On the security of RC4 in TLS. Published online at `http://www.isg.rhul.ac.uk/tls/`, 2013. Presented at the FSE 2013 invited talk [3] by Dan Bernstein. Accepted in USENIX 2013.
2. Riddhipratim Basu, Shirshendu Ganguly, Subhamoy Maitra, and Goutam Paul. A complete characterization of the evolution of RC4 pseudo random generation algorithm. *J. Mathematical Cryptology*, 2(3):257–289, 2008.
3. Daniel Bernstein. Failures of secret-key cryptography. Invited talk at FSE 2013. Session chaired by Bart Preneel.
4. Scott R. Fluhrer and David A. McGrew. Statistical analysis of the alleged RC4 keystream generator. In Bruce Schneier, editor, *FSE*, volume 1978 of *Lecture Notes in Computer Science*, pages 19–30. Springer, 2000.
5. Jovan Dj. Golic. Linear statistical weakness of alleged RC4 keystream generator. In Walter Fumy, editor, *EUROCRYPT*, volume 1233 of *Lecture Notes in Computer Science*, pages 226–238. Springer, 1997.
6. Takanori Isobe, Toshihiro Ohigashi, Yuhei Watanabe, and Masakatu Morii. Full plaintext recovery attack on broadcast RC4. In *Fast Software Encryption (FSE)*, 2013. To appear.
7. Robert J. Jenkins Jr. ISAAC and RC4. Published on the Internet at `http://burtleburtle.net/bob/rand/isaac.html`, 1996.
8. Jing Lv and Dongdai Lin. L-P states of RC4 stream cipher. *IACR Cryptology ePrint Archive*, 2013:266, 2013.
9. Jing Lv, Bin Zhang, and Dongdai Lin. Distinguishing attacks on RC4 and a new improvement of the cipher. *IACR Cryptology ePrint Archive*, 2013:176, 2013.
10. Subhamoy Maitra, Goutam Paul, Santanu Sarkar, Michael Lehmann, and Willi Meier. New results on generalization of Roos-type biases and related keystream of RC4. In *Africacrypt*, 2013. To appear.
11. Subhamoy Maitra, Goutam Paul, and Sourav Sengupta. Attack on broadcast RC4 revisited. In Antoine Joux, editor, *FSE*, volume 6733 of *Lecture Notes in Computer Science*, pages 199–217. Springer, 2011.
12. Subhamoy Maitra and Sourav Sen Gupta. A stronger glimpse of RC4. *IACR Cryptology ePrint Archive*, 2013, 2013.
13. Itsik Mantin. Analysis of the stream cipher RC4. Master's thesis, The Weizmann Institute of Science, Israel, 2001. Available at `http://www.wisdom.weizmann.ac.il/~itsik/{RC4}/{RC4}.html`.
14. Itsik Mantin. Predicting and distinguishing attacks on RC4 keystream generator. In Ronald Cramer, editor, *EUROCRYPT*, volume 3494 of *Lecture Notes in Computer Science*, pages 491–506. Springer, 2005.
15. Itsik Mantin and Adi Shamir. A practical attack on broadcast RC4. In Mitsuru Matsui, editor, *FSE*, volume 2355 of *Lecture Notes in Computer Science*, pages 152–164. Springer, 2001.
16. Ilya Mironov. (not so) random shuffles of RC4. In Moti Yung, editor, *CRYPTO*, volume 2442 of *Lecture Notes in Computer Science*, pages 304–319. Springer, 2002.
17. Mohammad Ali Orumiehchiha, Josef Pieprzyk, Elham Shakour, and Ron Steinfeld. Cryptanalysis of RC4(n,m) stream cipher. *IACR Cryptology ePrint Archive*, 2013:178, 2013.
18. Goutam Paul, Subhamoy Maitra, and Rohit Srivastava. On non-randomness of the permutation after RC4 key scheduling. In Serdar Boztas and Hsiao feng Lu, editors, *AAECC*, volume 4851 of *Lecture Notes in Computer Science*, pages 100–109. Springer, 2007.
19. Andrew Roos. A class of weak keys in the RC4 stream cipher. Two posts in sci.crypt, message-id `43u1eh$1j3@hermes.is.co.za` and `44ebge$llf@hermes.is.co.za`, 1995. Available at `http://www.impic.org/papers/WeakKeys-report.pdf`.
20. Santanu Sarkar. Further non-randomness in RC4, RC4A and VMPC. In *International Workshop on Coding and Cryptography (WCC)*, 2013.
21. Sourav Sen Gupta, Subhamoy Maitra, Goutam Paul, and Santanu Sarkar. Proof of empirical RC4 biases and new key correlations. In Ali Miri and Serge Vaudenay, editors, *Selected Areas in Cryptography*, volume 7118 of *Lecture Notes in Computer Science*, pages 151–168. Springer, 2011.
22. Sourav Sen Gupta, Subhamoy Maitra, Goutam Paul, and Santanu Sarkar. (Non–)random sequences from (non–)random permutations – analysis of RC4 stream cipher. *Journal of Cryptology*, 2013. To appear. Published online in December 2012, with DOI: 10.1007/s00145-012-9138-1.
23. Pouyan Sepehrdad. *Statistical and Algebraic Cryptanalysis of Lightweight and Ultra-Lightweight Symmetric Primitives*. Phd thesis no. 5415, École Polytechnique Fédérale de Lausanne (EPFL), 2012. Available at `http://lasecwww.epfl.ch/~sepehrdad/Pouyan_Sepehrdad_PhD_Thesis.pdf`.
24. Pouyan Sepehrdad, Petr Susil, Serge Vaudenay, and Martin Vuagnoux. Smashing WEP in a passive attack. In *Fast Software Encryption (FSE)*, 2013. To appear.
25. Pouyan Sepehrdad, Serge Vaudenay, and Martin Vuagnoux. Discovery and exploitation of new biases in RC4. In Alex Biryukov, Guang Gong, and Douglas R. Stinson, editors, *Selected Areas in Cryptography*, volume 6544 of *Lecture Notes in Computer Science*, pages 74–91. Springer, 2010.
26. Pouyan Sepehrdad, Serge Vaudenay, and Martin Vuagnoux. Statistical attack on RC4 - distinguishing WPA. In Kenneth G. Paterson, editor, *EUROCRYPT*, volume 6632 of *Lecture Notes in Computer Science*, pages 343–363. Springer, 2011.

27. Joachim Striömbergson and Simon Josefsson. The perils of repeating patterns: Observation of some weak keys in RC4. *IACR Cryptology ePrint Archive*, 2013:241, 2013.

A Proofs of the results in Section 4

We first list some existing results that will be needed in our proofs.

Proposition 1. *[13, Theorem 6.2.1] At the end of RC4 KSA, for* $0 \le u \le N-1$, $0 \le v \le N-1$,

$$\Pr(S_0[u] = v) = \begin{cases} \frac{1}{N}\left(\left(\frac{N-1}{N}\right)^v + \left(1 - \left(\frac{N-1}{N}\right)^v\right)\left(\frac{N-1}{N}\right)^{N-u-1}\right), & \text{if } v \le u; \\ \frac{1}{N}\left(\left(\frac{N-1}{N}\right)^{N-u-1} + \left(\frac{N-1}{N}\right)^v\right), & \text{if } v > u. \end{cases}$$

Proposition 2. *[22, Lemma 1] After the first round of RC4 PRGA, for* $0 \le u \le N-1$, $0 \le v \le N-1$, *the probability* $\Pr(S_1[u] = v)$ *is:*

$$\Pr(S_1[u] = v) = \begin{cases} \Pr(S_0[1] = 1) + \sum_{X \ne 1} \Pr(S_0[1] = X \ \wedge \ S_0[X] = 1), & u = 1, v = 1; \\ \sum_{X \ne 1, v} \Pr(S_0[1] = X \ \wedge \ S_0[X] = v), & u = 1, v \ne 1; \\ \Pr(S_0[1] = u) + \sum_{X \ne u} \Pr(S_0[1] = X \ \wedge \ S_0[u] = u), & u \ne 1, v = u; \\ \sum_{X \ne u, v} \Pr(S_0[1] = X \ \wedge \ S_0[u] = v), & u \ne 1, v \ne u. \end{cases}$$

Proposition 3. *[22, Theorem 1] In RC4 PRGA, for* $3 \le u \le N-1$ *and* $0 \le v \le N-1$,

$$\begin{aligned} \Pr(S_{u-1}[u] = v) \approx & \Pr(S_1[u] = v)\left(1 - \frac{1}{N}\right)^{u-2} \\ & + \sum_{y=2}^{u-1} \sum_{w=0}^{u-y} \frac{\Pr(S_1[y] = v)}{w! \cdot N} \left(\frac{u-y-1}{N}\right)^w \left(1 - \frac{1}{N}\right)^{u-3-w}. \end{aligned}$$

Now we present complete proofs of all the results in Section 4.

Lemma 3. *Suppose that* l *is the length of the secret key of RC4. Then for* $1 \le x \le \lfloor \frac{N}{l} \rfloor$, *we have*

$$\Pr(S^K_{xl+1}[xl-1] = -xl \ \wedge \ S^K_{xl+1}[xl] = 0) \approx \frac{1}{N^2} + \left(1 - \frac{1}{N^2}\right)\alpha_{x,l},$$

where $\alpha_{x,l} = \frac{1}{N}\left(1 - \frac{3}{N}\right)^{xl-2}\left(1 - \frac{xl+1}{N}\right)$.

Proof. The major path that leads to the target event is as follows.

- In the first round of the KSA, when $i^K_1 = 0$ and $j^K_1 = K[0]$, the value 0 is swapped into the index $S^K[K[0]]$ with probability 1.
- The index $j^K_1 = K[0] \notin \{xl-1, xl, -xl\}$, so that the values $xl-1, xl, -xl$ at these indices respectively are not swapped out in the first round of the KSA. We as well require $K[0] \notin \{1, \ldots, xl-2\}$, so that the value 0 at index $K[0]$ is not touched by these values of i^K during the next $xl-2$ rounds of the KSA. This happens with probability $\left(1 - \frac{xl+1}{N}\right)$.

– From round 2 to $xl-1$ (i.e., for $i^K = 1$ to $xl-2$) of the KSA, none of $j_2^K, \ldots, j_{xl-1}^K$ touches the three indices $\{xl, -xl, K[0]\}$. This happens with probability $\left(1-\frac{3}{N}\right)^{xl-2}$.
– In round xl of the KSA, when $i_{xl}^K = xl-1$, j_{xl}^K becomes $-xl$ with probability $\frac{1}{N}$, thereby moving $-xl$ into index $xl-1$.
– In round $xl+1$ of the KSA, when $i_{xl+1}^K = xl$, j_{xl+1}^K becomes $j_{xl}^K + S_{xl}^K[xl] + K[xl] = -xl + xl + K[0] = K[0]$, and as discussed above, this index contains the value 0. Hence, after the swap, $S_{xl+1}^K[xl] = 0$. Since $K[0] \neq xl-1$, we have $S_{xl+1}^K[xl-1] = -xl$.

Considering the above events to be independent, the probability that all of above occur together is given by $\alpha_{x,l} = \frac{1}{N}\left(1-\frac{3}{N}\right)^{xl-2}\left(1-\frac{xl+1}{N}\right)$. If the above path does not occur, then the target event happens due to random association with probability $\frac{1}{N^2}$, thus contributing a probability of $(1-\alpha_{x,l})\frac{1}{N^2}$. Adding the two contributions, the result follows. □

Theorem 7. *Suppose that l is the length of the secret key of RC4. Then for $1 \leq x \leq \lfloor \frac{N}{l} \rfloor$, we have*

$$\Pr(S_{xl}[xl] = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) = \Pr(t_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) \approx \frac{1}{N^2} + \left(1-\frac{1}{N^2}\right)\beta_{x,l},$$

where $\beta_{x,l} = \frac{1}{N}\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)^{N-3}\left(1-\frac{3}{N}\right)^{xl-2}\left(1-\frac{xl+1}{N}\right)$.

Proof. From the proof of Lemma 3, consider the major path with probability $\alpha_{x,l}$ for the event $(S_{xl+1}^K[xl-1] = -xl \ \wedge \ S_{xl+1}^K[xl] = 0)$. For the remaining $N-xl-1$ rounds of the KSA and for the first $xl-2$ rounds of the PRGA (i.e., for a total of $N-3$ rounds), none of the values of j^K (corresponding to the KSA rounds) or j (corresponding to the PRGA rounds) should touch the indices $\{xl-1, xl\}$. This happens with a probability of $\left(1-\frac{2}{N}\right)^{N-3}$.

Now, in round $xl-1$ of PRGA, $i_{xl-1} = xl-1$, from where the value $xl-1$ moves to index j_{xl-1} due to the swap. In the next round, $i_{xl} = xl$ and $j_{xl} = j_{xl-1} + S_{xl-1}[xl] = j_{xl-1}$, provided the value 0 at index xl had not been swapped out by j_{xl-1}, the probability of which is $1-\frac{1}{N}$. So during the next swap, the value $-xl$ moves from index j_{xl} to index xl and the value 0 moves from index xl to j_{xl}. The probability of the above major path leading to the event $(S_{xl}[xl] = -xl \ \wedge \ S_{xl}[j_{xl}] = 0)$ is given by $\beta_{x,l} = \alpha_{x,l}\left(1-\frac{2}{N}\right)^{N-3}\left(1-\frac{1}{N}\right)$. If this path does not occur, then there is always a chance of $\frac{1}{N^2}$ for the target event to happen due to random association. Adding the two contributions and substituting the value of $\alpha_{x,l}$ from Lemma 3, the result follows.

Further, as $t_{xl} = S_{xl}[xl] + S_{xl}[j_{xl}]$, the event $(S_{xl}[xl] = -xl \ \wedge \ S_{xl}[j_{xl}] = 0)$ is equivalent to the event $(t_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0)$, and hence the result. □

Theorem 8. *Suppose that l is the length of the secret key of RC4. Then for $1 \leq x \leq \lfloor \frac{N}{l} \rfloor$, we have*

$$\Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) \approx \frac{1}{N^2} + \left(1-\frac{1}{N^2}\right)\gamma_{x,l},$$

where $\gamma_{x,l} = \frac{1}{N^2}\left(1-\frac{xl+1}{N}\right)\sum_{u=xl+1}^{N-1}\left(1-\frac{1}{N}\right)^u\left(1-\frac{2}{N}\right)^{u-xl}\left(1-\frac{3}{N}\right)^{N-u+2xl-4}$.

Proof. From the PRGA update rule, we have $j_{xl} = j_{xl-1} + S_{xl-1}[xl]$. Hence, $S_{xl}[j_{xl}] = S_{xl-1}[xl] = 0$ implies $j_{xl} = j_{xl-1}$ as well as $Z_{xl} = S_{xl}[S_{xl}[xl] + S_{xl}[j_{xl}]] = S_{xl}[S_{xl-1}[j_{xl}] + 0] = S_{xl}[S_{xl-1}[j_{xl-1}]] = S_{xl}[S_{xl-2}[xl-1]]$. Thus, the event $(Z_{xl} = -xl \wedge S_{xl}[j_{xl}] = 0)$ is equivalent to the event $(S_{xl}[S_{xl-2}[xl-1]] = -xl \ \wedge \ S_{xl-1}[xl] = 0)$.

From the proof of Lemma 3, consider the major path with probability α_{xl} for the joint event $(S_{xl+1}^K[xl-1] = -xl \wedge S_{xl+1}^K[xl] = 0)$. This constitutes the first part of our main path leading to the target event. The second part, having probability $\alpha'_{x,l}$, can be constructed as follows.

– For an index $u \in [xl+1, N-1]$, we have $S_u^K[u] = u$. This happens with probability $\left(1-\frac{1}{N}\right)^u$.
– For the KSA rounds $xl+2$ to u, the j^K values do not touch the indices $xl-1$ and xl. This happens with probability $\left(1-\frac{2}{N}\right)^{u-xl-1}$.
– In round $u+1$ of KSA, when $i_{u+1}^K = u$, j_{u+1}^K becomes $xl-1$ with probability $\frac{1}{N}$. Due to the swap, the value u moves to $S_{u+1}^K[xl-1]$ and the value $-xl$ moves to $S_{u+1}^K[u] = S_{u+1}^K[S_{u+1}^K[xl-1]]$.
– For the remaining $N-u-1$ rounds of the KSA and for the first $xl-1$ rounds of the PRGA, none of the j^K or j values should touch the indices $\{xl-1, S[xl-1], xl\}$. This happens with a probability of $\left(1-\frac{3}{N}\right)^{N-u+xl-2}$.
– So far, we have $(S_{xl-1}[S_{xl-2}[xl-1]] = -xl \ \wedge \ S_{xl-1}[xl] = 0)$. Now, we should also have $j_{xl} \notin \{xl-1, S[xl-1]\}$ for $S_{xl}[S_{xl-2}[xl-1]] = S_{xl-1}[S_{xl-2}[xl-1]] = -xl$. The probability of this condition is $\left(1-\frac{2}{N}\right)$.

Assuming all the individual events in the above path to be mutually independent, we get $\alpha'_{x,l} = \frac{1}{N}\sum_{u=xl+1}^{N-1}\left(1-\frac{1}{N}\right)^u\left(1-\frac{2}{N}\right)^{u-xl}\left(1-\frac{3}{N}\right)^{N-u+xl-2}$. Thus, the probability of the entire path is given by $\gamma_{x,l} = \alpha_{x,l}\cdot\alpha'_{x,l} = \frac{1}{N^2}\left(1-\frac{xl+1}{N}\right)\sum_{u=xl+1}^{N-1}\left(1-\frac{1}{N}\right)^u\left(1-\frac{2}{N}\right)^{u-xl}\left(1-\frac{3}{N}\right)^{N-u+2xl-4}$.

If this path does not occur, then there is always a chance of $\frac{1}{N^2}$ for the target event to happen due to random association. Adding the two contributions, we get the result. □

Theorem 9. *For any keylength $l \in [3, N-1]$ and any integer $x \in [1, \lfloor\frac{N}{l}\rfloor]$, the probability* $\Pr(S_{xl}[j_{xl}] = 0)$ *is given by*

$$\delta_{x,l} \approx \Pr(S_1[xl] = 0)\left(1-\frac{1}{N}\right)^{xl-2} + \sum_{y=2}^{xl-1}\sum_{w=0}^{xl-y}\frac{\Pr(S_1[y]=0)}{w!\cdot N}\left(\frac{xl-y-1}{N}\right)^w\left(1-\frac{1}{N}\right)^{xl-3-w}.$$

Proof. Note that $S_{xl}[j_{xl}]$ is assigned the value of $S_{xl-1}[xl]$ due to the swap in round xl. Hence, by substituting $u = xl$ and $v = 0$ in Proposition 3, we get the result. □

Theorem 10. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor\frac{N}{l}\rfloor$, we have*

$$\tau_{x,l} = \Pr(t_{xl} = -xl) \ \approx \ \frac{1}{N^2} + \left(1-\frac{1}{N^2}\right)\beta_{x,l} + (1-\delta_{x,l})\frac{1}{N},$$

where $\beta_{x,l}$ is given in Theorem 7 and $\delta_{x,l}$ is given in Theorem 9.

Proof. We can write $\Pr(t_{xl} = -xl) = \Pr(t_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) + \Pr(t_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] \neq 0)$, where the first term is given by Theorem 7. When $S_{xl}[j_{xl}] \neq 0$, the event $(t_{xl} = -xl)$ can be assumed to occur due to random association. Hence the second term can be computed as $\Pr(S_{xl}[j_{xl}] \neq 0)\cdot\Pr(t_{xl} = -xl \mid S_{xl}[j_{xl}] \neq 0) \approx (1-\delta_{x,l})\frac{1}{N}$. Adding the two terms, we get the result. □

By dividing the joint probabilities $\Pr(S_{xl}[xl] = -xl \wedge S_{xl}[j_{xl}] = 0)$ and $\Pr(t_{xl} = -xl \wedge S_{xl}[j_{xl}] = 0)$ of Theorem 7, and $\Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0)$ of Theorem 8 by the appropriate marginals $\delta_{x,l} = \Pr(S_{xl}[j_{xl}] = 0)$ of Theorem 9 and $\tau_{x,l} = \Pr(t_{x,l} = -xl)$ of Theorem 10, we get theoretical values of the following conditional biases

1. $\Pr(S_{xl}[xl] = -xl \mid S_{xl}[j_{xl}] = 0) = \Pr(t_{xl} = -xl \mid S_{xl}[j_{xl}] = 0)$.
2. $\Pr(S_{xl}[j_{xl}] = 0 \mid t_{xl} = -xl)$.
3. $\Pr(Z_{xl} = -xl \mid S_{xl}[j_{xl}] = 0)$.

Theorem 11. *Suppose that l is the length of the secret key of RC4. Then for $1 \le x \le \lfloor\frac{N}{l}\rfloor$, we have*

$$\Pr(Z_{xl} = -xl) \ \approx \ \frac{1}{N^2} + \left(1-\frac{1}{N^2}\right)\gamma_{x,l} + (1-\delta_{x,l})\frac{1}{N},$$

where $\gamma_{x,l}$ is given in Theorem 8 and $\delta_{x,l}$ is given in Theorem 9.

Proof. We can write $\Pr(Z_{xl} = -xl) = \Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0) + \Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] \neq 0)$, where the first term is given by Theorem 8. When $S_{xl}[j_{xl}] \neq 0$, the event $(Z_{xl} = -xl)$ can be assumed to occur due to random association. Hence the second term can be computed as $\Pr(S_{xl}[j_{xl}] \neq 0) \cdot \Pr(Z_{xl} = -xl \mid S_{xl}[j_{xl}] \neq 0) \approx (1 - \delta_{x,l})\frac{1}{N}$. Adding the two terms, we get the result. □

By dividing the joint probability $\Pr(Z_{xl} = -xl \ \wedge \ S_{xl}[j_{xl}] = 0)$ of Theorem 8 by $\Pr(Z_{xl} = -xl)$ as given above, we get the theoretical value of $\Pr(S_{xl}[j_{xl}] = 0 \mid Z_{xl} = -xl)$.

A.1 Supporting figures for some keylengths other than 16-bytes

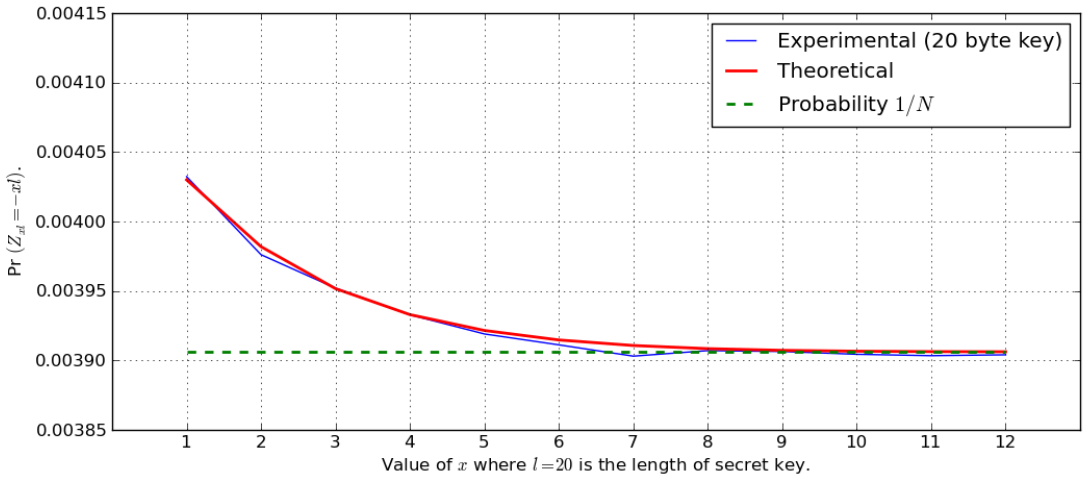

Fig. 8. Bias in the event $(Z_{xl} = -xl)$ for keylength $l = 20$ and $x = 1, 2, \ldots, 12$.

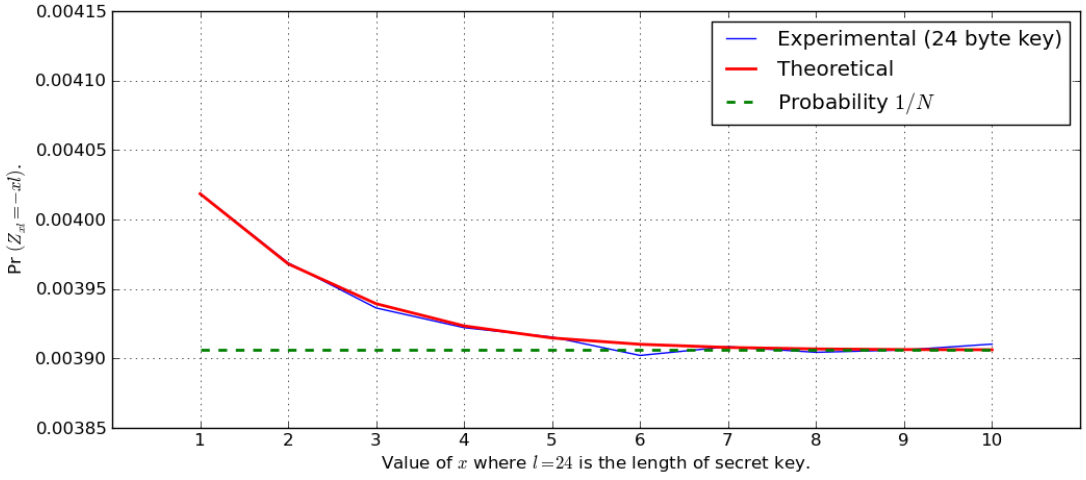

Fig. 9. Bias in the event $(Z_{xl} = -xl)$ for keylength $l = 24$ and $x = 1, 2, \ldots, 10$.

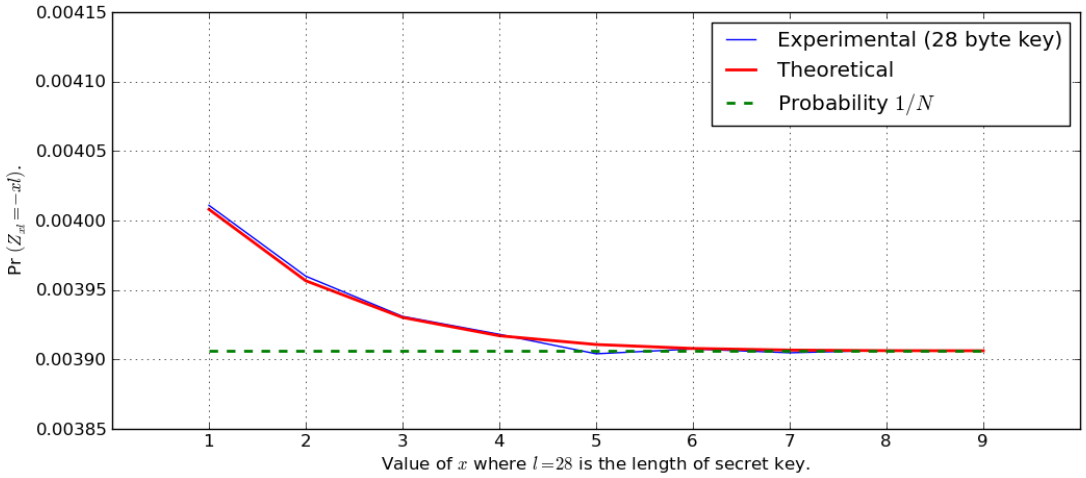

Fig. 10. Bias in the event $(Z_{xl} = -xl)$ for keylength $l = 28$ and $x = 1, 2, \ldots, 9$.

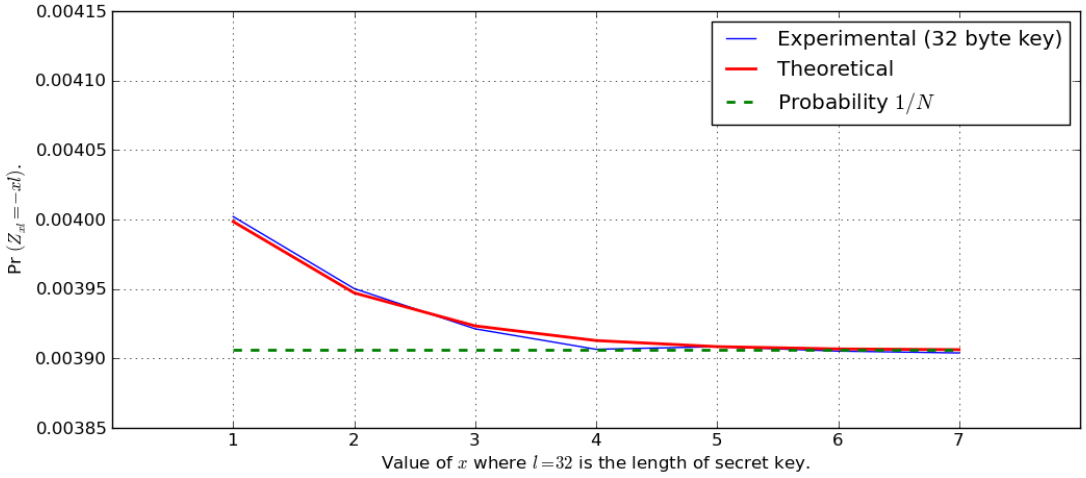

Fig. 11. Bias in the event $(Z_{xl} = -xl)$ for keylength $l = 32$ and $x = 1, 2, \ldots, 7$.