Proof of Empirical RC4 Biases and New Key Correlations

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

[1] Applied Statistics Unit, Indian Statistical Institute, Kolkata 700 108, India
[2] Dept. of Computer Science and Engg., Jadavpur University, Kolkata 700 032, India
{sg.sourav,sarkar.santanu.bir}@gmail.com, subho@isical.ac.in,
goutam.paul@ieee.org,

Abstract. In SAC 2010, Sepehrdad, Vaudenay and Vuagnoux have reported some empirical biases between the secret key, the internal state variables and the keystream bytes of RC4, by searching over a space of all linear correlations between the quantities involved. In this paper, for the first time, we give theoretical proofs for all such significant empirical biases. Our analysis not only builds a framework to justify the origin of these biases, it also brings out several new conditional biases of high order. We establish that certain conditional biases reported earlier are correlated with a third event with much higher probability. This gives rise to the discovery of new keylength-dependent biases of RC4, some as high as $50/N$, where N is the size of the RC4 permutation. The new biases in turn result in successful *keylength prediction* from the initial keystream bytes of the cipher.

Keywords: Conditional Bias, Key Correlation, Keylength Prediction, RC4.

1 Introduction

RC4 is one of the most popular stream ciphers for software applications. Designed by Ron Rivest in 1987, the algorithm of RC4 has two parts; Key Scheduling (KSA) and Pseudo-Random Generation (PRGA), presented in Table 1.

Given a secret key k of length l bytes, an array K of size N bytes is created to hold the key such that $K[y] = k[y \bmod l]$ for all $y \in [0, N-1]$. Generally, N is chosen to be 256. The first part of the cipher, KSA, uses this K to scramble an initial identity permutation $\{0, 1, \ldots, N-1\}$ to obtain a 'secret' state S. Then the PRGA generates keystream bytes from this initial state S, which are used for encrypting the plaintext. Two indices i (deterministic) and j (pseudo-random) are used in KSA as well as PRGA to point to the locations of S. All additions in the RC4 algorithm are performed modulo N.

After r (≥ 1) rounds of RC4 PRGA, we denote the variables by S_r, i_r, j_r, z_r and the output index $S_r[i_r] + S_r[j_r]$ by t_r. After r rounds of KSA, we denote the same 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 $S_0 = S_N^K$.

A. Miri and S. Vaudenay (Eds.): SAC 2011, LNCS 7118, pp. 151–168, 2012.

Table 1. The RC4 Algorithm: KSA and PRGA

Key Scheduling (KSA)

```
Input: Secret Key $K$.
Output: S-Box $S$ generated by $K$.

Initialize $S = \{0, 1, 2, \ldots, N-1\}$;
Initialize counter: $j = 0$;

for $i = 0, \ldots, N-1$ do
    $j = j + S[i] + K[i]$;
    Swap $S[i] \leftrightarrow S[j]$;
end
```

Pseudo-Random Generation (PRGA)

```
Input: S-Box $S$, output of KSA.
Output: Random stream $Z$.

Initialize the counters: $i = j = 0$;

while $TRUE$ do
    $i = i + 1$, $j = j + S[i]$;
    Swap $S[i] \leftrightarrow S[j]$;
    Output $Z = S[S[i] + S[j]]$;
end
```

Existing Results. In SAC 2010, Sepehrdad, Vaudenay and Vuagnoux [12] have reported experimental results of an exhaustive search for biases in all possible linear combinations of the state variables and the keystream bytes of RC4. In the process, they have discovered many new biases that are significantly high compared to random association. Some of these biases were further shown to be useful for key recovery in WEP [3] mode. In a recent work [13] at Eurocrypt 2011, the same authors have utilized the pool of all existing biases of RC4, including a few reported in [12], to mount a distinguishing attack on WPA [4].

In the above approach, RC4 is treated as a black box, where the secret key bytes are the *inputs*, the permutation and the index j are *internal state variables* and the keystream bytes are the *outputs*. The goal of [12] was to find out empirical correlations between the inputs, internal state and the outputs and no attempt was made to theoretically prove these biases. Finding empirical biases without any justification or proof may be useful from application point of view. However, cryptanalysis is a disciplined branch of science and a natural quest in RC4 cryptanalysis should be: *Where do all these biases come from?*

Motivation. We felt three primary reasons behind a theoretical investigation into the source and nature of these biases.

- We attempt to build a framework to analyze the biases and their origin.
- In the process of proving the existing biases, one may need to consider some additional events and thus may end up discovering new biases, leading to further insight into the cipher. We have observed some interesting events with strong biases, which have not yet been reported in the literature.
- When there is a conditional bias in the event ‘A given B’, there may be three reasons behind it: either some subset of A directly causes B or some subset of B directly causes A or another set C of different events cause both A and B. Just from empirical observation, it is impossible to infer what is the actual reason behind the bias. Only a theoretical study can shed light upon the interplay between the events. Our observations and analysis suggest that some conditional biases reported in [12] are possibly of the third kind discussed above and this provides us with some interesting new results depending on the length of the RC4 secret key.

Contribution. Our main contribution in this paper is summarized as follows.

1. In Section 2, we provide theoretical proofs for some significant empirical biases of RC4 reported in SAC 2010 [12]. In particular, we justify the reported biases of order approximately $2/N$, summarized in Table 2. Note that the authors of [12] denote the PRGA variables by primed indices. Moreover, the probabilities mentioned in the table are the ones observed in [12], and the values for 'biases at all rounds (round-dependent)' are the ones for $r = 3$. We provide general proofs and formulas for all of these biases.

Table 2. Significant biases observed in [12] and proved in this paper

Type of Bias	Label as in [12]	Event	Probability
Bias at Specific Initial Rounds	New_004	$j_2 + S_2[j_2] = S_2[i_2] + z_2$	$2/N$
	New_noz_007	$j_2 + S_2[j_2] = 6$	$2.37/N$
	New_noz_009	$j_2 + S_2[j_2] = S_2[i_2]$	$2/N$
	New_noz_014	$j_1 + S_1[i_1] = 2$	$1.94/N$
Bias at All Rounds (round-independent)	New_noz_001	$j_r + S_r[i_r] = i_r + S_r[j_r]$	$2/N$
	New_noz_002	$j_r + S_r[j_r] = i_r + S_r[i_r]$	$2/N$
Bias at All Rounds (round-dependent)	New_000	$S_r[t_r] = t_r$	$1.9/N$ at $r = 3$
	New_noz_004	$S_r[i_r] = j_r$	$1.9/N$ at $r = 3$
	New_noz_006	$S_r[j_r] = i_r$	$2.34/N$ at $r = 3$

2. In Section 3, we try to justify the bias $\Pr[S_{16}[j_{16}] = 0 \mid z_{16} = -16] = 0.038488$ observed in [12], for which the authors have commented:

   > *"So far, we have no explanation about this new bias."* [12, Section 3]

   We have observed that the implied correlation arises because both the events depend on some other event based on the length of RC4 secret key. We also prove some related correlations in this direction, in full generality for any keylength l.
3. In Section 3, we also prove an array of *new keylength-dependent conditional biases* of RC4 that are of the same or even higher magnitude. To the best of our knowledge, these are not reported in the literature [1, 2, 6, 9–15].
4. In Section 3.3, we prove a strong correlation between the length l of the secret key and the l-th output byte (typically for $5 \leq l \leq 30$), and thus propose a method to predict the keylength of the cipher by observing the keystream. As far as we know, no such significant keylength related bias exists in the RC4 literature [1, 2, 6, 9–15].

2 Proofs of Recent Empirical Observations

In this section, we investigate some significant empirical biases discovered and reported in [12]. We provide theoretical justification only for the new biases which are of the approximate order of $2/N$ or more, summarized in Table 2. In this target list, general biases refer to the ones occurring in all initial rounds of PRGA ($1 \leq r \leq N-1$), whereas the specific ones have been reported only for rounds 1 and 2 of PRGA. *We do not consider the biases reported for rounds* $0 \bmod 16$ *in this section, as they are of order* $1/N^2$ *or less.*

For the proofs and numeric probability calculations in this paper, we require [6, Theorem 6.3.1], restated as Proposition 1 below.

Proposition 1. *At the end of RC4 KSA, for* $0 \leq u \leq N-1$, $0 \leq v \leq N-1$,

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

If a pseudorandom permutation is taken as the initial state S_0 of RC4 PRGA, then we would have $\Pr(S_0[u]=v) = \frac{1}{N}$ for all $0 \leq u \leq N-1$, $0 \leq v \leq N-1$.

2.1 Bias at Specific Initial Rounds of PRGA

In this part of the paper, we prove the biases labeled New_noz_014, New_noz_007, New_noz_009 and New_004, as in [12, Fig. 3 and Fig. 4] and Table 2.

Theorem 1. *After the first round* ($r=1$) *of RC4 PRGA,*

$$\Pr(j_1 + S_1[i_1] = 2) = \Pr(S_0[1]=1) + \sum_{X \neq 1} \Pr(S_0[X] = 2-X) \cdot \Pr(S_0[1] = X)$$

Proof. Note that $j_1 = S_0[1]$ and $S_1[i_1] = S_0[j_1]$. So, in the case $j_1 = S_0[1] = 1$, we will have $j_1 + S_0[j_1] = S_0[1] + S_0[1] = 2$ with probability 1. Otherwise, the probability turns out to be $\Pr(j_1 + S_0[j_1] = 2 \ \& \ j_1 = S_0[1] \neq 1) = \sum_{X \neq 1} \Pr(X + S_0[X] = 2 \ \& \ S_0[1] = X)$. Thus, the probability $\Pr(j_1 + S_1[i_1] = 2)$ can be written as $\Pr(j_1 + S_1[i_1] = 2) = \Pr(S_0[1]=1) + \sum_{X \neq 1} \Pr(S_0[X] = 2-X) \cdot \Pr(S_0[1] = X)$, as desired. Hence the claimed result. □

Numerical Values. If we consider the practical RC4 scheme, the probabilities involving S_0 in the expression for $\Pr(j_1 + S_1[i_1] = 2)$ should be evaluated using Proposition 1, giving a total probability of approximately $1.937/N$ for $N = 256$. This closely matches the observed value $1.94/N$. If we assume that RC4 PRGA starts with a truly pseudorandom initial state S_0, the probability turns out to be approximately $2/N - 1/N^2 \approx 1.996/N$ for $N = 256$, i.e., almost twice that of a random occurrence.

Theorem 2. *After the second round* ($r=2$) *of RC4 PRGA, the following probability relations hold between the index* j_2 *and the state variables* $S_2[i_2], S_2[j_2]$.

$$\Pr\ (j_2 + S_2[j_2] = 6) \approx \Pr(S_0[1] = 2) + \sum_{X \textit{ even}, X \neq 2} (2/N) \cdot \Pr(S_0[1] = X) \quad (1)$$

$$\Pr\ (j_2 + S_2[j_2] = S_2[i_2]) \approx 2/N - 1/N^2 \quad (2)$$

$$\Pr\ (j_2 + S_2[j_2] = S_2[i_2] + z_2) \approx 2/N - 1/N^2 \quad (3)$$

Proof. In Equation (1), we have $j_2+S_2[j_2] = (j_1+S_1[2])+S_1[i_2] = S_0[1]+2\cdot S_1[2]$. In this expression, note that if $S_0[1] = 2$, then one must have the positions 1 and 2 swapped in the first round of PRGA, and thus $S_1[2] = S_0[1] = 2$ as well. This provides one path for $j_2 + S_2[j_2] = S_0[1] + 2 \cdot S_1[2] = 2 + 2 \times 2 = 6$, with probability $\Pr(S_0[1] = 2) \cdot 1 \approx \frac{1}{N}$. If on the other hand, $S_0[1] = X \neq 2$, we have $\Pr(j_2 + S_2[j_2] = 6 \;\&\; S_0[1] \neq 2) = \sum_{X\neq 2} \Pr(X + 2 \cdot S_1[2] = 6 \;\&\; S_0[1] = X)$. Note that the value of X is bound to be even and for each such value of X, the variable $S_1[2]$ can take 2 different values to satisfy the equation $2\cdot S_1[2] = 6-X$. Thus, we have $\sum_{X\neq 2} \Pr(2 \cdot S_1[2] = 6 - X \;\&\; S_0[1] = X) \approx \sum_{X \text{ even}, X\neq 2} \frac{2}{N} \cdot \Pr(S_0[1] = X)$. Combining the two disjoint cases $S_0[1] = 2$ and $S_0[1] \neq 2$, we get Equation (1).

In case of Equation (2), we have a slightly different condition $S_0[1]+2\cdot S_1[2] = S_2[i_2] = S_1[j_2] = S_1[S_0[1] + S_1[2]]$. In this expression, if we have $S_1[2] = 0$, then the left hand side reduces to $S_0[1]$ and the right hand side becomes $S_1[S_0[1] + S_1[2]] = S_1[S_0[1]] = S_1[j_1] = S_0[i_1] = S_0[1]$ as well. This provides a probability $\frac{1}{N}$ path for the condition to be true. In all other cases with $S_1[2] \neq 0$, we can approximate the probability for the condition as $\frac{1}{N}$, and hence approximate the total probability $\Pr(j_2 + S_2[j_2] = S_2[i_2])$ as $\Pr(j_2 + S_2[j_2] = S_2[i_2] \;\&\; S_1[2] = 0) + \Pr(j_2 + S_2[j_2] = S_2[i_2] \;\&\; S_1[2] \neq 0) \approx \frac{1}{N} + \left(1 - \frac{1}{N}\right) \cdot \frac{1}{N} = \frac{2}{N} - \frac{1}{N^2}$.

Finally, for Equation (3), the main observation is that this is almost identical to the condition of Equation (2) apart from the inclusion of z_2. But our first path of $S_1[2] = 0$ in the previous case also provides us with $z_2 = 0$ with probability 1 (this path was first observed by Mantin and Shamir [7]). Thus, we have $\Pr(j_2 + S_2[j_2] = S_2[i_2] + z_2 \;\&\; S_1[2] = 1) \approx \frac{1}{N} \cdot 1$. In all other cases with $S_1[2] \neq 0$, we assume the conditions to match uniformly at random, and therefore have $\Pr(j_2 + S_2[j_2] = S_2[i_2] + z_2) \approx \frac{1}{N} \cdot 1 + \left(1 - \frac{1}{N}\right) \cdot \frac{1}{N} = \frac{2}{N} - \frac{1}{N^2}$. Hence the desired results of Equations (1), (2) and (3). □

Numerical Values. In case of Equation (1), if we assume S_0 to be the practical initial state for RC4 PRGA, and substitute all probabilities involving S_0 using Proposition 1, we get the total probability equal to $2.36/N$ for $N = 256$. This value closely match the observed probability $2.37/N$. If we suppose that S_0 is pseudorandom, we will get probability $2/N - 2/N^2 \approx 1.992/N$ for Equation (1). The theoretical results are summarized in Table 3 along with the experimentally observed probabilities of [12].

Table 3. Theoretical and observed biases at specific initial rounds of RC4 PRGA

Label [12]	Event	Observed Probability [12]	Theoretical Probability S_0 of RC4	 Random S_0
New_noz_014	$j_1 + S_1[i_1] = 2$	$1.94/N$	$1.937/N$	$1.996/N$
New_noz_007	$j_2 + S_2[j_2] = 6$	$2.37/N$	$2.363/N$	$1.992/N$
New_noz_009	$j_2 + S_2[j_2] = S_2[i_2]$	$2/N$	$1.996/N$	$1.996/N$
New_noz_004	$j_2 + S_2[j_2] = S_2[i_2] + z_2$	$2/N$	$1.996/N$	$1.996/N$

2.2 Biases at All Initial Rounds of PRGA (Round-Independent)

In this section, we turn our attention to the biases labeled New_noz_001 and New_noz_002 in [12], both of which continue to persist in all initial rounds ($1 \leq r \leq N-1$) of RC4 PRGA.

Theorem 3. *At any initial round* $1 \leq r \leq N-1$ *of RC4 PRGA, the following two relations hold between the indices* i_r, j_r *and the state variables* $S_r[i_r], S_r[j_r]$.

$$\Pr(j_r + S_r[j_r] = i_r + S_r[i_r]) \approx 2/N \tag{4}$$

$$\Pr(j_r + S_r[i_r] = i_r + S_r[j_r]) \approx 2/N \tag{5}$$

Proof. For both the events mentioned above, we shall take the path $i_r = j_r$. Notice that $i_r = j_r$ occurs with probability $\frac{1}{N}$ and in that case both the events mentioned above hold with probability 1. In the case where $i_r \neq j_r$, we rewrite the events as $S_r[j_r] = (i_r - j_r) + S_r[i_r]$ and $S_r[j_r] = (j_r - i_r) + S_r[i_r]$. Here we already know that $S_r[j_r] \neq S_r[i_r]$, as $j_r \neq i_r$ and S_r is a permutation. Thus in case $i_r \neq j_r$, the values of $S_r[i_r]$ and $S_r[j_r]$ can be chosen in $N(N-1)$ ways (drawing from a permutation without replacement) to satisfy the relations stated above. This gives the total probability for each event approximately as $\Pr(j_r = i_r) \cdot 1 + \sum_{j_r \neq i_r} \frac{1}{N(N-1)} = \frac{1}{N} + (N-1) \cdot \frac{1}{N(N-1)} = \frac{2}{N}$. Hence the claimed result for Equations (4) and (5). □

The probabilities for New_noz_001 and New_noz_002 proved in Theorem 3 do not vary with change in r (i.e., they continue to persist at the same order of $2/N$ at any arbitrary round of PRGA), and our theoretical results match the probabilities reported in [12, Fig. 2].

2.3 Biases at All Initial Rounds of PRGA (Round-Dependent)

Next, we consider the biases that are labeled as New_000, New_noz_004 and New_noz_006 in [12, Fig. 2]. We prove the biases for rounds 3 to 255 in RC4 PRGA, and we show that all of these decrease in magnitude with increase in r, as observed experimentally in the original paper.

Let us first prove observation New_noz_006 of [12]. This proof was also attempted in [5, Lemma 1], where the event was equivalently stated as $S_{r-1}[r] = r$. But that proof used a crude approximation which resulted in a slight mismatch of the theoretical and practical patterns in the main result of the paper [5, Fig. 2]. Our proof of Theorem 4, as follows, corrects the proof of [5, Lemma 1], and removes the mismatch in [5, Fig. 2].

Theorem 4. *For PRGA rounds* $r \geq 3$*, value of* $\Pr(S_r[j_r] = i_r)$ *is approximately*

$$\Pr(S_1[r] = r) \left[1 - \frac{1}{N}\right]^{r-2} + \sum_{t=2}^{r-1} \sum_{k=0}^{r-t} \frac{\Pr(S_1[t] = r)}{k! \cdot N} \left[\frac{r-t-1}{N}\right]^k \left[1 - \frac{1}{N}\right]^{r-3-k}$$

Before proving Theorem 4, let us first prove a necessary technical result.

Lemma 1. *After the first round of RC4 PRGA, the probability* $\Pr(S_1[t] = r)$ *is*

$$\Pr(S_1[t] = r) = \begin{cases} \sum_{X=0}^{N-1} \Pr(S_0[1] = X) \cdot \Pr(S_0[X] = r), & t = 1; \\ \Pr(S_0[1] = r) + (1 - \Pr(S_0[1] = r)) \cdot \Pr(S_0[r] = r), & t = r; \\ (1 - \Pr(S_0[1] = t)) \cdot \Pr(S_0[t] = r), & t \neq 1, r. \end{cases}$$

Proof. After the first round of RC4 PRGA, we obtain the state S_1 from the initial state S_0 through a single swap operation between the positions $i_1 = 1$ and $j_1 = S_0[i_1] = S_0[1]$. Thus, all other positions of S_0 remain the same apart from these two. This gives us the value of $S_1[t]$ as follows: $S_1[t] = S_0[S_0[1]]$ if $t = 1$, $S_1[t] = S_0[1]$ if $t = S_0[1]$, and $S_1[t] = S_0[t]$ in all other cases. Now, we can compute the probabilities $\Pr(S_1[t] = r)$ based on the probabilities for S_0, which are in turn derived from Proposition 1. We have three cases:

- Case $t = 1$. In this case, using the recurrence relation $S_1[1] = S_0[S_0[1]]$, we can write $\Pr(S_1[1] = r) = \sum_{X=0}^{N-1} \Pr(S_0[1] = X) \cdot \Pr(S_0[X] = r)$.
- Case $t = r$. In this situation, if $S_0[1] = r$, we will surely have $S_1[r] = r$ as these are the positions swapped in the first round, and if $S_0[1] \neq r$, the position $t = r$ remains untouched and $S_1[r] = r$ is only possible if $S_0[r] = r$. Thus, $\Pr(S_1[r] = r) = \Pr(S_0[1] = r) + (1 - \Pr(S_0[1] = r)) \cdot \Pr(S_0[r] = r)$.
- Case $t \neq 1, r$. In all other cases where $t \neq 1, r$, it can either take the value $S_0[1]$ with probability $\Pr(S_0[1] = t)$, or not. If $t = S_0[1]$, the value $S_0[t]$ will get swapped with $S_0[1] = t$ itself, i.e., we will get $S_1[t] = t \neq r$ for sure. Otherwise, the value $S_1[t]$ remains the same as $S_0[t]$. Hence, $\Pr(S_1[t] = r) = (1 - \Pr(S_0[1] = t)) \cdot \Pr(S_0[t] = r)$.

Combining all the above cases together, we obtain the desired result. □

Proof of Theorem 4. Let us start from the PRGA state S_1, that is, the state that has been updated once in the PRGA (we refer to the state after KSA by S_0). We know that the event $\Pr(S_1[r] = r)$ is positively biased for all r, and hence the natural path for investigation is the effect of the event $(S_1[r] = r)$ on $(S_{r-1}[r] = r)$, i.e, on $(S_r[j_r] = i_r)$. Notice that there can be two cases, as follows.

Case I. In the first case, suppose that $(S_1[r] = r)$ after the first round, and the r-th index is not disturbed for the next $r - 2$ state updates. Notice that index i varies from 2 to $r - 1$ during these period, and hence never touches the r-th index. Thus, the index r will retain its state value r if index j does not touch it. The probability of this event is $\left(1 - \frac{1}{N}\right)^{r-2}$ over all the intermediate rounds. Hence the first part of the probability is $\Pr(S_1[r] = r) \left(1 - \frac{1}{N}\right)^{r-2}$.

Case II. In the second case, suppose that $S_1[r] \neq r$ and $S_1[t] = r$ for some $t \neq r$. In such a case, only a swap between the positions r and t during rounds 2 to $r - 1$ of PRGA can make the event $(S_{r-1}[r] = r)$ possible. Notice that if t does not fall in the *path of* i, that is, if the index i does not touch the t-th location, then the value at $S_1[t]$ can only go to some position behind i, and this can never reach $S_{r-1}[r]$, as i can only go up to $(r - 1)$ during this period. Thus we must have $2 \leq t \leq r - 1$ for $S_1[t]$ to reach $S_{r-1}[r]$. Note that the way $S_1[t]$ can move to the r-th position may be either a one hop or a multi-hop route.

- In the easiest case of single hop, we require j not to touch t until i touches t, and $j = r$ when $i = t$, and j not to touch r for the next $r-t-1$ state updates. Total probability comes to be $\Pr(S_1[t] = r)\left(1 - \frac{1}{N}\right)^{t-2} \cdot \frac{1}{N} \cdot \left(1 - \frac{1}{N}\right)^{r-t-1} = \Pr(S_1[t] = r) \cdot \frac{1}{N}\left(1 - \frac{1}{N}\right)^{r-3}$.
- Suppose that it requires $(k + 1)$ hops to reach from $S_1[t]$ to $S_{r-1}[r]$. Then the main issue to note is that the transfer will never happen if the position t swaps with any index which does not lie in the future *path of* i. Again, this path of i starts from $\frac{r-t-1}{N}$ for the first hop and decreases approximately to $\frac{r-t-1}{lN}$ at the l-th hop. We would also require j not to touch the position r for the remaining $(r - 3 - k)$ number of rounds. Combining all, we get the second part of the probability as $\Pr(S_1[t] = r)\left[\prod_{l=1}^{k} \frac{r-t-1}{lN}\right]\left[1 - \frac{1}{N}\right]^{r-3-k} = \frac{\Pr(S_1[t]=r)}{k! \cdot N}\left[\frac{r-t-1}{N}\right]^{k}\left[1 - \frac{1}{N}\right]^{r-3-k}$.

Finally, note that the number of hops $(k+1)$ is bounded from below by 1 and from above by $(r - t + 1)$, depending on the initial gap between t and r positions. Considering the sum over t and k with this consideration, we get the desired expression for $\Pr(S_{r-1}[r] = r)$. □

Remark 1. In proving Theorem 4, we use the initial condition $S_1[r] = r$ to branch out the probability paths, and not $S_0[r] = r$ as in [5, Lemma 1]. This is because the probability of $S[r] = r$ takes a leap from around $1/N$ in S_0 to about $2/N$ in S_1, and this turns out to be the actual cause behind the bias in $S_{r-1}[r] = r$.

Fig. 1 illustrates the experimental observations (averages taken over 100 million runs with 16-byte key) and the theoretical values for the distribution of $\Pr(S_r[j_r] = i_r)$ over the initial rounds $3 \le r \le 255$ of RC4 PRGA. It is evident that our theoretical formula matches the experimental observations in this case.

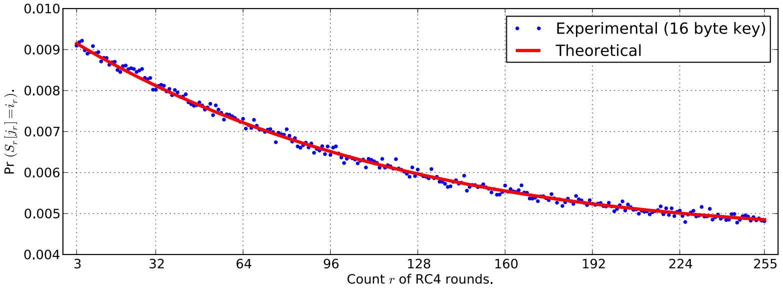

Fig. 1. Distribution of $\Pr(S_r[j_r] = i_r)$ for initial rounds $3 \le r \le 255$ of RC4 PRGA

Now let us take a look at the other two round-dependent biases of RC4, observed in [12]. We can state the related result in Theorem 5 (corresponding to observations New_noz_004 and New_000).

Theorem 5. *For PRGA rounds $r \geq 3$, the probabilities* $\Pr(S_r[i_r] = j_r)$ *and* $\Pr(S_r[t_r] = t_r)$ *are approximately*

$$\frac{r}{N^2} + \sum_{X=r}^{N-1} \tfrac{1}{N} \left[\Pr(S_1[X] = X) \left[1 - \frac{1}{N}\right]^{r-2} \right.$$
$$\left. + \sum_{t=2}^{r-1} \sum_{k=0}^{r-t} \frac{\Pr(S_1[t] = r)}{k! \cdot N} \left[\frac{r-t-1}{N}\right]^k \left[1 - \frac{1}{N}\right]^{r-3-k} \right]$$

The proof of this result is omitted for brevity, as it follows the same logic as in the proof of Theorem 4. A brief proof sketch is presented as follows. For this proof sketch, we consider the variables j_r and t_r to be pseudorandom variables that can take any value between 0 to 255 with probability $1/N$. The reader may note that this is a crude approximation, especially for small values of r, and causes minor mismatch with the experimental observations in the final result.

Proof-sketch for $\Pr(S_r[i_r] = j_r)$. For this probability computation, we first rewrite the event as $(S_{r-1}[j_r] = j_r)$ to make it look similar to $S_{r-1}[r] = r$, as in Theorem 4. The only difference is that we were concentrating on a fixed index r in Theorem 4 instead of a variable index j_r. This produces two cases.

Case I. First, suppose that j_r assumes a value $X \geq r$. In this case, the probability calculation can be split in two paths, one in which $S_1[X] = X$ is assumed, and the other in which $S_1[X] \neq X$. If we assume $S_1[X] = X$, the probability of $(S_{r-1}[X] = X)$ becomes $\Pr(S_1[X] = X) \left[1 - \frac{1}{N}\right]^{r-2}$, similar to the logic in Theorem 4. If we suppose that $S_1[t] = X$ was the initial state, then one may notice the following two sub-cases:

- The probability for this path is identical to that in Theorem 4 if $2 \leq t \leq r-1$.
- The probability is 0 in case $t \geq r$, as in this case the value X will always be behind the position of $i_r = r$, whereas $X > r$ as per assumption. That is, the value X can never reach index X from t.

Assuming $\Pr(j_r = X) = 1/N$, this gives $\sum_{X=r}^{N-1} \frac{1}{N} \left[\Pr(S_1[X] = X) \left[1 - \frac{1}{N}\right]^{r-2} \right.$ $\left. + \sum_{t=2}^{r-1} \sum_{k=0}^{r-t} \frac{\Pr(S_1[t]=r)}{k! \cdot N} \left[\frac{r-t-1}{N}\right]^k \left[1 - \frac{1}{N}\right]^{r-3-k} \right]$.

Case II. In the second case, we assume that j_r takes a value X between 0 to $r-1$. Approximately this complete range is touched by index i for sure, and may also be touched by index j. Thus, with probability approximately 1, the index $j_r = X$ is touched by either of the indices. Simplifying all complicated computations involving the initial position of value X and the exact location of index X in this case, we shall assume that the approximate value of $\Pr(S_{r-1}[X] = X)$ is $1/N$. Thus, the total contribution of Case II, assuming $\Pr(j_r = X) = 1/N$, is given by $\sum_{X=0}^{r-1} \Pr(j_r = X) \cdot \Pr(S_{r-1}[X] = X) \approx \sum_{X=0}^{r-1} \frac{1}{N} \cdot \frac{1}{N} = \frac{r}{N^2}$.

Adding the contributions of the two disjoint cases I and II, we obtain the total probability for $(S_r[i_r] = j_r)$ as desired. One may investigate Case II in more details to incorporate all intertwined sub-cases, and obtain a better closed form expression for the probability.

Proof-sketch for $\Pr(S_r[t_r] = t_r)$. In this case, notice that t_r is just another random variable like j_r, and may assume all values from 0 to 255 with approximately the same probability $1/N$. Thus we can approximate $\Pr(S_r[t_r] = t_r)$ by $\Pr(S_{r-1}[j_r] = j_r)$ with a high confidence margin to obtain the desired expression.

This approximation is particularly close for higher values of r because the effect of a single state change $S_{r-1} \to S_r$ is low in such a case. For smaller values of r, one may approximate $\Pr(S_{r-1}[t_r] = t_r)$ by $\Pr(S_{r-1}[j_r] = j_r)$ and critically analyze the effect of the r-th round of PRGA thereafter. However, in spite of the approximations we made, one may note that the theoretical values closely match the experimental observations (averages taken over 100 million runs of RC4 with 16-byte key), as shown in Fig. 2.

Fig. 2 illustrates the experimental observations (averages taken over 100 million runs with 16-byte key) and the theoretical values for the distributions of $\Pr(S_r[i_r] = j_r)$ and $\Pr(S_r[t_r] = t_r)$ over the initial rounds $3 \le r \le 255$ of RC4 PRGA. It is evident that our theoretical formulas approximately match the experimental observations in both the cases; the cause of the little deviation is explained in the proof sketch above.

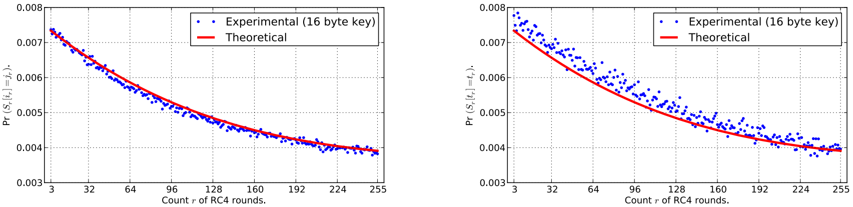

Fig. 2. Distributions of $\Pr(S_r[i_r] = j_r)$ and $\Pr(S_r[t_r] = t_r)$ for initial rounds $3 \le r \le 255$ of RC4 PRGA

Apart from the biases proved so far, all other unconditional biases reported in [12] are of order $1/N^2$ or less, and we omit their analysis in this paper. The next most significant bias reported in [12] was a new conditional bias arising from a set of correlations in RC4 PRGA. A careful study of this new bias gives rise to several related observations and results related to the KSA as well, as presented in the next section.

3 Biases Based on Keylength

In SAC 2010, Sepehrdad, Vaudenay and Vuagnoux [12] discovered several correlations in PRGA using DFT based approach. A list of such biases was presented in [12, Fig. 10], and the authors commented that:

> "After investigation, it seems that all the listed biases are artifact of a new conditional bias which is $\Pr[S'_{16}[j'_{16}] = 0 \mid z_{16} = -16] = 0.038488$."

However, the authors also admitted that

> *"So far, we have no explanation about this new bias."*

In our notation, the above event is denoted as $\Pr(S_{16}[j_{16}] = 0 \mid z_{16} = -16)$. While exploring this conditional bias and related parameters of RC4 PRGA, we could immediately observe two things:

1. The number 16 in the result comes from the keylength that is consistently chosen to be 16 in [12] for most of the experimentation. In its general form, the conditional bias should be stated as (crude approximation):

$$\Pr\left(S_l[j_l] = 0 \mid z_l = -l\right) \approx \frac{10}{N}. \tag{6}$$

   It is surprising why this natural observation could not be identified earlier.
2. Along the same line of investigation, we could find a family of related conditional biases, stated in their general form as follows (crude approximations):

$$\Pr(z_l = -l \mid S_l[j_l] = 0) \approx 10/N \tag{7}$$

$$\Pr(S_l[l] = -l \mid S_l[j_l] = 0) \approx 30/N \tag{8}$$

$$\Pr(t_l = -l \mid S_l[j_l] = 0) \approx 30/N \tag{9}$$

$$\Pr(S_l[j_l] = 0 \mid t_l = -l) \approx 30/N \tag{10}$$

Note that bias (7) follows almost immediately from bias (6), and biases (10) and (9) are related in a similar fashion. Moreover, bias (8) implies bias (9) as $t_l = S_l[l] + S_l[j_l] = -l$ under the given condition. However, we investigate even further to study the bias caused in z_l due to the state variables.

3.1 Dependence of Conditional Biases on RC4 Secret Key

We found that all of the aforementioned conditional biases between the two events under consideration are related to the following third event that is dependent on the values and the length of the RC4 secret key.

$$\sum_{i=0}^{l-1} K[i] + \frac{l(l-1)}{2} \equiv -l \pmod N$$

We shall henceforth denote the above event by $(f_{l-1} = -l)$, following the notation of Paul and Maitra [9], and this event is going to constitute the base for most of the conditional probabilities we consider hereafter. We consider $\Pr(f_{l-1} = -l) \approx \frac{1}{N}$, assuming that f_{l-1} can take any value modulo N uniformly at random.

Extensive experimentation with different keylengths (100 million runs for each keylength $1 \le l \le 256$) revealed strong bias in all of the following events:

$$\begin{array}{ll} \Pr(S_l[j_l] = 0 \mid f_{l-1} = -l), & \Pr(S_l[l] = -l \mid f_{l-1} = -l), \\ \Pr(t_l = -l \mid f_{l-1} = -l), & \Pr(z_l = -l \mid f_{l-1} = -l). \end{array}$$

Each of the correlations (6), (7), (8), (9), and (10) is an artifact of these common keylength-based correlations in RC4 PRGA. In this section, we discuss and justify all these conditional biases.

To prove our observations in this paper, we shall require the following existing results from the literature of key-correlation in RC4. These are the correlations observed by Roos [11] in 1995, which were later proved by Paul and Maitra [9].

Proposition 2. [9, Lemma 1] *If index j is pseudorandom at each KSA round, we have* $\Pr\left(j^K_{y+1} = f_y\right) \approx \left(1 - \frac{1}{N}\right)^{1+\frac{y(y+1)}{2}} + \frac{1}{N}$.

Proposition 3. [9, Corollary 1] *On completion of KSA in the RC4 algorithm,* $\Pr(S_0[y] = f_y) = \Pr(S^K_N[y] = f_y) \approx \left(1 - \frac{y}{N}\right) \cdot \left(1 - \frac{1}{N}\right)^{\frac{y(y+1)}{2}+N} + \frac{1}{N}$.

Proposition 4. [9, Corollary 1] *On completion of KSA,* $\Pr(S_0[S_0[y]] = f_y) \approx \left[\frac{y}{N} + \frac{1}{N}\left[1 - \frac{1}{N}\right]^{2-y} + \left[1 - \frac{y}{N}\right]^2 \left[1 - \frac{1}{N}\right]\right] \left[1 - \frac{1}{N}\right]^{\frac{y(y+1)}{2}+2N-4}$ *for* $0 \le y \le 31$.

Note that in each of the above statements,

$$f_y = S^K_0\left[\sum_{x=0}^{y} S^K_0[x] + \sum_{x=0}^{y} K[x]\right] = \sum_{x=0}^{y} x + \sum_{x=0}^{y} K[x] = \sum_{x=0}^{y} K[x] + \frac{y(y+1)}{2}.$$

3.2 Proof of Keylength-Dependent Conditional Biases

In this section, we will prove the four main conditional biases that we have observed. Each depends on the event $(f_{l-1} = -l)$, and can be justified as follows. In each of the following theorems, the notation '$x : A \xrightarrow{\alpha} B$' denotes that the value x transits from position A to position B with probability α.

Theorem 6. *Suppose that l is the length of the secret key used in the RC4 algorithm. Given $f_{l-1} = \sum_{i=0}^{l-1} K[i] + l(l-1)/2 = -l$, we have*

$$\Pr(S_l[j_l] = 0) \approx \frac{1}{N} + \left[1 - \frac{l}{N}\right]\left[1 - \frac{1}{N}\right]^{N+l-2}\left[\left[1 - \frac{1}{N}\right]^{1+\frac{l(l+1)}{2}} + \frac{1}{N}\right]$$

$$\Pr(S_{l-2}[l-1] = -l) \approx \frac{1}{N} + \left[1 - \frac{1}{N}\right]^{l-1}\left[\left[1 - \frac{l-1}{N}\right]\left[1 - \frac{1}{N}\right]^{N+\frac{l(l-1)}{2}} + \frac{1}{N}\right]$$

Proof. For proving the first conditional bias, we need to trace the value 0 over KSA and the first l rounds of PRGA. We start from $S^K_0[0] = 0$, as the initial state S^K_0 of KSA is the identity permutation in RC4. The following gives the trace pattern for 0 through the complete KSA and l initial rounds of PRGA. We shall discuss some of the transitions in details.

$$0 : S^K_0[0] \xrightarrow{1} S^K_1[K[0]] \xrightarrow{p_1} S^K_l[K[0]] \xrightarrow{p_2} S^K_{l+1}[l] \xrightarrow{p_3} S_{l-1}[l] \xrightarrow{1} S_l[j_l]$$

Here $p_1 = \left(1 - \frac{l}{N}\right)\left(1 - \frac{1}{N}\right)^{l-1}$ denotes the probability that index $K[0]$ is not touched by i^K and j^K in the first l rounds of KSA, $p_2 = \left(1 - \frac{1}{N}\right)^{1+\frac{l(l+1)}{2}} + \frac{1}{N}$

denotes the probability $\Pr(j_{l+1}^K = f_l = K[0])$ (using Proposition 2) such that 0 is swapped from $S_l^K[K[0]]$ to $S_{l+1}^K[l]$, and $p_3 = \left(1 - \frac{1}{N}\right)^{N-2}$ denotes the probability that the location $S_{l+1}^K[l]$ containing 0 is not touched by i^K, j^K in the remaining $N - l - 1$ rounds of KSA or by i, j in the first $l - 1$ rounds of PRGA. So, this path gives a total probability of $p_1p_2p_3$. If this path does not hold, we assume that the event $(S_l[j_l] = 0)$ still holds at random, with probability $1/N$. Thus, the total probability is obtained as

$$\Pr(S_l[j_l] = 0) = p_1p_2p_3 + (1 - p_1p_2p_3) \cdot \frac{1}{N} = \frac{1}{N} + \left(1 - \frac{1}{N}\right) p_1p_2p_3.$$

We do a similar propagation tracking for the value $f_{l-1} = -l$ to prove the second result, and the main path for this tracking looks as follows.

$$-l : S_0^K[-l] \xrightarrow{p_4} S_0[l-1] \xrightarrow{p_5} S_{l-2}[l-1]$$

Here we get $p_4 = \Pr(S_0[l-1] = f_{l-1}) = \left(1 - \frac{l-1}{N}\right)\left(1 - \frac{1}{N}\right)^{N + \frac{l(l-1)}{2}} + \frac{1}{N}$ using Proposition 3 directly, and $p_5 = \left(1 - \frac{1}{N}\right)^{l-2}$ denotes the probability that the index $(l-1)$, containing $-l$, is not touched by i, j in the first $l - 2$ rounds of PRGA. Similar to the previous proof, the total probability can be calculated as

$$\Pr(S_{l-2}[l-1] = -l) = p_4p_5 + (1 - p_4p_5) \cdot \frac{1}{N} = \frac{1}{N} + \left(1 - \frac{1}{N}\right) p_4p_5.$$

We get the claimed results by substituting p_1, p_2, p_3 and p_4, p_5 appropriately. □

Numerical Values. If we substitute $l = 16$, the most common keylength for RC4, and $N = 256$, we get the probabilities of Theorem 6 of magnitude

$$\Pr(S_l[j_l] = 0 \mid f_{l-1} = -l) \approx \Pr(S_{l-2}[l-1] = -l \mid f_{l-1} = -l) \approx 50/N.$$

These are, to the best of our knowledge, *the best known key-dependent conditional biases in RC4 PRGA till date.* The estimates closely match the experiments we performed over 100 million runs with 16-byte keys. In the next theorem, we look at a few natural consequences of these biases.

Theorem 7. *Suppose that l is the length of the RC4 secret key. Given that $f_{l-1} = \sum_{i=0}^{l-1} K[i] + l(l-1)/2 = -l$, the probabilities $\Pr(S_l[l] = -l \mid f_{l-1} = -l)$ and $\Pr(t_l = -l \mid f_{l-1} = -l)$ are approximately*

$$\frac{1}{N} + \left(1 - \frac{1}{N}\right) \cdot \left[\frac{1}{N} + \left[1 - \frac{l}{N}\right]\left[1 - \frac{1}{N}\right]^{N+l-2} \left[\left[1 - \frac{1}{N}\right]^{1+\frac{l(l+1)}{2}} + \frac{1}{N}\right]\right]$$
$$\cdot \left[\frac{1}{N} + \left[1 - \frac{1}{N}\right]^{l-1} \left[\left[1 - \frac{1}{N}\right]^{N-l} + \frac{1}{N}\right]\right]$$

Proof. Before proving the path for the target events, let us take a look at rounds $l-1$ and l of RC4 PRGA when $S_{l-2}[l-1] = -l$ and $S_{l-1}[l] = 0$. In this situation, we have the following propagation for the value $-l$.

$$-l : S_{l-2}[l-1] \xrightarrow{1} S_{l-1}[j_{l-1}] = S_{l-1}[j_l] \xrightarrow{1} S_l[l]$$

In the above path, the equality holds because $j_l = j_{l-1} + S_{l-1}[l] = j_{l-1} + 0$ as per the conditions. Again, we have $S_l[j_l] = S_{l-1}[l] = 0$, implying $t_l = S_l[l] + S_l[j_l] = -l + 0 = -l$ as well. This explains the same expression for the probabilities of the two events in the statement.

Note that we require both the events $(S_l[j_l] = 0 \mid f_{l-1} = -l)$ and $(S_{l-2}[l-1] = -l \mid f_{l-1} = -l)$ to occur simultaneously, and need to calculate the joint probability. Also note that there is a significant overlap between the tracking paths of these two events, as they both assume that the first l positions of the state S_0^K are not touched by j^K in the first l rounds of KSA (refer to the proof of Theorem 6 of this paper and proofs of [9, Theorem 1, Corollary 1] for details). In other words, if we assume the occurrence of event $(S_l[j_l] = 0 \mid f_{l-1} = -l)$ (with probability p_6, as derived in Theorem 6, say), then the precondition for $(S_{l-2}[l-1] = -l \mid f_{l-1} = -l)$ will be satisfied, and thus the modified conditional probability is $\Pr(S_{l-2}[l-1] = -l \mid S_l[j_l] = 0 \ \& \ f_{l-1} = -l) = \frac{1}{N} + \left[1 - \frac{1}{N}\right]^{l-1} \left[\left[1 - \frac{1}{N}\right]^{N-l} + \frac{1}{N}\right] = p_7$, say. Now, we can compute the joint probability of the two events as

$$\Pr(S_l[l] = -l \mid f_{l-1} = -l) = p_6 p_7 + (1 - p_6 p_7) \cdot \frac{1}{N} = \frac{1}{N} + \left(1 - \frac{1}{N}\right) \cdot p_6 p_7.$$

Substituting the values of p_6 and p_7, we obtain the desired result. Event $(t_l = -l)$ follows immediately from $(S_l[l] = -l)$, with the same conditional probability. □

Numerical Values. Substituting $l = 16$ and $N = 256$, we get the probabilities of Theorem 7 of the magnitude $\Pr(S_l[l] = -l \mid f_{l-1} = -l) = \Pr(t_l = -l \mid f_{l-1} = -l) \approx 20/N$. These estimates closely match our experimental results taken over 100 million runs of RC4 with 16-byte keys.

Conditional Bias in Output. We could also find that the bias in $(z_l = -l)$ is caused due to the event $f_{l-1}[l]$, but in a different path than the one we have discussed so far. We prove the formal statement next as Theorem 8.

Theorem 8. *Suppose that l is the length of the secret key of RC4. Given that $f_{l-1} = \sum_{i=0}^{l-1} K[i] + l(l-1)/2 = -l$, the probability $\Pr(z_l = -l)$ is approximately*

$$\frac{1}{N} + \left[1 - \frac{1}{N}\right] \cdot \left[\frac{1}{N} + \left[1 - \frac{l}{N}\right]\left[1 - \frac{1}{N}\right]^{N+l-2} \left[\left[1 - \frac{1}{N}\right]^{1+l} + \frac{1}{N}\right]\right]$$
$$\cdot \left[\frac{1}{N} + \left[1 - \frac{1}{N}\right]^{l+1} \Pr(S_0[S_0[l-1]] = f_{l-1})\right]$$

Proof. The proof is similar to that of Theorem 7 as both require $S_l[j_l] = S_{l-1}[l] = 0$ to occur first. Note that if $S_l[j_l] = S_{l-1}[l] = 0$, we will always have

$$z_l = S_l[S_l[l] + S_l[j_l]] = S_l[S_{l-2}[l-1] + 0] = S_l[S_{l-2}[l-1]].$$

Thus the basic intuition is to use the path $S_0[S_0[l-1]] = f_{l-1} = -l$ to get

$$-l : S_0[S_0[l-1]] \xrightarrow{p_8} S_{l-2}[S_{l-2}[l-1]] \xrightarrow{p_9} S_l[S_{l-2}[l-1]]$$

In the above expression, $p_8 = \left(1 - \frac{1}{N}\right)^{l-2}$ and $p_9 = \left(1 - \frac{1}{N}\right)^2$ denote the probabilities of j not touching the state index that stores the value $-l$. This introduces a probability $\left(1 - \frac{1}{N}\right)^l$. Thus $\Pr(S_l[S_{l-2}[l-1]] = -l \mid f_{l-1} = -l)$ is cumulatively given by $\frac{1}{N} + \left[1 - \frac{1}{N}\right]^{l+1} \Pr(S_0[S_0[l-1]] = f_{l-1}) = p_{10}$, say. Note that one of the preconditions to prove [9, Theorem 4] is that the first $(l-1)$ places of state S_0^K remain untouched by j^K for the first $l-1$ rounds of KSA. This partially matches with the precondition to prove $\Pr(S_l[j_l] = 0 \mid f_{l-1} = -l)$ (see Theorem 6), where we require the same for first l places over the first l rounds of KSA. Thus we derive the formula for $\Pr(S_l[j_l] = 0 \mid S_0[S_0[l-1]] = -l \;\&\; f_{l-1} = -l)$ by modifying the result of Theorem 6 as $\frac{1}{N} + \left[1 - \frac{l}{N}\right] \left[1 - \frac{1}{N}\right]^{N+l-2} \left[\left[1 - \frac{1}{N}\right]^{1+l} + \frac{1}{N}\right] = p_{11}$, say. The final probability for $(z_l = -l \mid f_{l-1} = -l)$ can now be computed as

$$\Pr(z_l = -l \mid f_{l-1} = -l) = p_{10}p_{11} + (1 - p_{10}p_{11}) \cdot \frac{1}{N} = \frac{1}{N} + \left(1 - \frac{1}{N}\right) \cdot p_{10}p_{11}.$$

Substituting appropriate values for p_{10} and p_{11}, we get the desired result. □

Let us consider $\Pr(z_l = -l \mid S_l[j_l] = 0) = \Pr(S_l[S_{l-2}[l-1]] = -l \mid S_l[j_l] = 0)$. From the proof of Theorem 8, it is evident that the events $(S_l[S_{l-2}[l-1]] = -l)$ and $(S_l[j_l] = 0)$ have no obvious connection. Yet, there exists a strong correlation between them, possibly due to some hidden events that cause them to co-occur with a high probability. We found that one of these hidden events is $(f_{l-1} = -l)$.

From the proofs of Theorems 6 and 8, we know that both the aforementioned events depend strongly on $(f_{l-1} = -l)$, but along two different paths, as follows.

$$\begin{aligned} 0 &: S_0^K[0] \xrightarrow{1} S_1^K[K[0]] \xrightarrow{p_1} S_l^K[K[0]] \xrightarrow{p_2} S_{l+1}^K[l] \xrightarrow{p_3} S_{l-1}[l] \xrightarrow{1} S_l[j_l] \\ -l &: S_0^K[S_0^K[l-1]] \xrightarrow{p_{12}} S_0[S_0[l-1]] \xrightarrow{p_8} S_{l-2}[S_{l-2}[l-1]] \xrightarrow{p_9} S_l[S_{l-2}[l-1]] \end{aligned}$$

Here p_{12} depends on the probability $\Pr(S_0[S_0[l-1]] = f_{l-1})$ from Proposition 4. Using these two paths, one may obtain the value of $\Pr(z_l = -l \;\&\; S_l[j_l] = 0)$ as

$$\begin{aligned} &\Pr(z_l = -l \;\&\; S_l[j_l] = 0) \\ &\quad = \Pr(f_{l-1} = -l) \cdot \Pr(S_l[S_{l-2}[l-1]] = -l \;\&\; S_l[j_l] = 0 \mid f_{l-1} = -l) \\ &\quad\quad + \Pr(f_{l-1} \neq -l) \cdot \Pr(S_l[S_{l-2}[l-1]] = -l \;\&\; S_l[j_l] = 0 \mid f_{l-1} \neq -l). \end{aligned}$$

As before, $\Pr(f_{l-1} = -l)$ can be taken as $1/N$. If one assumes that the aforementioned two paths are independent, the probabilities from Theorems 6 and 8 can be substituted in the above expression. If one further assumes that the events occur uniformly at random if $f_{l-1} \neq -l$, the values of $\Pr(S_l[j_l] = 0 \mid z_l = -l)$ and $\Pr(z_l = -l \mid S_l[j_l] = 0)$ turn out to be approximately $5/N$ each (for $l = 16$).

However, our experiments show that the two paths mentioned earlier are *not entirely independent*, and we obtain $\Pr(z_l = -l \ \& \ S_l[j_l] = 0 \mid f_{l-1} = -l) \approx 5/N$. Moreover, the events are *not uniformly random* if $f_{l-1} \neq -l$; rather they are considerably biased for a range of values of f_{l-1} around $-l$ (e.g., for values like $-l+1$, $-l+2$ etc.). These hidden paths contribute towards the probability $\Pr(f_{l-1} \neq -l) \Pr(z_l = -l \ \& \ S_l[j_l] = 0 \mid f_{l-1} \neq -l) \approx 5/N^2$. Through a careful treatment of the dependences and all the hidden paths, one would be able to justify the above observations, and obtain

$$\Pr(S_l[j_l] = 0 \mid z_l = -l) \approx \Pr(z_l = -l \mid S_l[j_l] = 0) \approx 10/N.$$

Similar techniques for analyzing dependences and hidden paths would work for all correlations reported in Equations 6, 7, 8, 9 and, 10.

We now shift our focus to $\Pr(z_l = -l \mid f_{l-1} = -l)$ and its implications.

Numerical Values. First of all, notice that the value of $\Pr(z_l = -l \mid f_{l-1} = -l)$ depends on the value of $\Pr(S_0[S_0[l-1]] = f_{l-1})$. Proposition 4 gives an explicit formula for $\Pr(z_l = -l \mid f_{l-1} = -l)$ for l up to 32. As l increases beyond 32, one may check by experimentation that this probability converges approximately to $1/N$. Thus, for $1 \leq l \leq 32$, one can use the formula from Proposition 4, and for $l > 32$, one may replace $\Pr(S_0[S_0[l-1]] = f_{l-1})$ by $1/N$ to approximately compute the distribution of $(z_l = -l \mid f_{l-1} = -l)$ completely. In fact, after the state recovery attack by Maximov and Khovratovich [8], that is of time complexity around 2^{241}, choosing a secret key of length $l > 30$ is not meaningful. The value of $\Pr(z_l = -l \mid f_{l-1} = -l)$ for some typical values of l are

$$12/N \text{ for } l = 5 \qquad 11/N \text{ for } l = 8 \qquad 7/N \text{ for } l = 16 \qquad 2/N \text{ for } l = 30.$$

In the list above, each conditional probability is quite high in magnitude compared to the natural probability of random occurrence. We try to exploit this bias in the next section to predict the length of RC4 secret key.

3.3 Keylength Prediction from Keystream

The huge conditional bias proved in Theorem 8 hints that there may be a related unconditional bias present in the event $z_l = -l$ as well. In fact, New_007 in [12, Fig. 5] reports a bias in $(z_i = -i)$ for $i = 0 \bmod 16$. The reported bias for $i = 16$ is $1.0411/N$. Notice that almost all experiments of [12] used the keylength $l = 16$, which encourages our speculation for an unconditional bias in $(z_l = -l)$ for any general keylength l of RC4 secret key. Systematic investigation in this direction reveals the following result.

Theorem 9. *Suppose that l is the length of the secret key of RC4. The probability* $\Pr(z_l = -l)$ *is given by*

$$\Pr(z_l = -l) \approx \frac{1}{N} + [N \cdot \Pr(z_l = -l \mid f_{l-1} = -l) - 1] \cdot \frac{1}{N^2}.$$

Proof. We provide a quick sketch of the proof to obtain a crude approximation of this bias in z_l. Notice that we already have a path $(z_l = -l \mid f_{l-1} = -l)$ with probability calculated in Theorem 8. If we assume that for all other values of $f_{l-1} \neq -l$, the output z_l can take the value $-l$ uniformly at random, we have

$$\begin{aligned}\Pr(z_l = -l) &\approx \Pr(f_{l-1} = -l) \cdot \Pr(z_l = -l \mid f_{l-1} = -l) \\ &\quad + \Pr(f_{l-1} \neq -l) \cdot \Pr(z_l = -l \mid f_{l-1} \neq -l) \\ &= \frac{1}{N} \cdot \Pr(z_l = -l \mid f_{l-1} = -l) + \left(1 - \frac{1}{N}\right) \cdot \frac{1}{N} \\ &= \frac{1}{N} + [N \cdot \Pr(z_l = -l \mid f_{l-1} = -l) - 1] \cdot \frac{1}{N^2}.\end{aligned}$$

Thus we obtain the desired result. □

Numerical Values. We have a closed form expression for $\Pr(z_l = -l \mid f_{l-1} = -l)$ from Theorem 8 in cases where $1 \leq l \leq 32$ (using Proposition 4). We have also calculated some numerical values of this probability for $l = 5, 8, 16, 30$ and $N = 256$. Using those numeric approximations, the value of $\Pr(z_l = -l)$ is

$$1/N + 11/N^2 \text{ for } l = 5 \qquad 1/N + 10/N^2 \text{ for } l = 8$$

$$1/N + 6/N^2 \text{ for } l = 16 \qquad 1/N + 2/N^2 \text{ for } l = 30$$

Predicting the Keylength. The lower bound for $\Pr(z_l = -l)$ within the typical range of keylength ($5 \leq l \leq 30$) is approximately $1/N + 1/N^2$, which is quite high and easily detectable. In experiments with 100 million runs and different keylengths, we have found that the probabilities are even higher than those mentioned above. This helps us in predicting the length of the secret key from the output, as follows.

1. Find the output byte z_x biased towards $-x$. This requires $O(N^3)$ many samples as the bias is $O(1/N^2)$. A 'sample' in this case means the observation of keystream bytes z_x for all $5 \leq x \leq 30$ for a specific key. The bias is computed by examining these keystream bytes with different keys, which are all of the same length l, say.
2. Check if the probability $\Pr(z_x = -x)$ is equal or greater than the value proved in Theorem 9.
3. If the above statements hold for some $5 \leq x \leq 30$, the keylength can be accurately predicted as $l = x$.

Although the bias in $z_l = -l$ has been noticed earlier in the literature for specific keylengths, no attempts have been made for its generalization. Moreover, to the best of our knowledge, the prediction of keylength from the keystream has never been attempted. We have performed extensive experiments with varying keylengths to verify the practical feasibility of the prediction technique. This prediction technique proves to be successful for all keylengths within the typical usage range $5 \leq l \leq 30$. As already pointed out in Section 3.2, choosing a secret key of length $l > 30$ is not recommended. So, our *keylength prediction* effectively works for all practical values of the keylength.

4 Conclusion

In the paper [12] of SAC 2010, several empirical observations relating a few RC4 variables have been reported, and here we prove all the significant ones. In the process, we provide a framework for justifying such non-random events in their full generality. Our study identifies and proves a family of new key correlations beyond those observed in [12]. These, in turn, result in keylength dependent biases in initial keystream bytes of RC4, enabling effective keylength prediction.

Acknowledgments. The authors would like to thank the anonymous reviewers for their feedback that helped in improving the presentation of this paper.

References

1. Fluhrer, S.R., Mantin, I., Shamir, A.: Weaknesses in the Key Scheduling Algorithm of RC4. In: Vaudenay, S., Youssef, A.M. (eds.) SAC 2001. LNCS, vol. 2259, pp. 1–24. Springer, Heidelberg (2001)
2. Klein, A.: Attacks on the RC4 stream cipher. Designs, Codes and Cryptography 48(3), 269–286 (2008)
3. LAN/MAN Standard Committee. ANSI/IEEE standard 802.11b: Wireless LAN Medium Access Control (MAC) and Physical Layer (phy) Specifications (1999)
4. LAN/MAN Standard Committee. ANSI/IEEE standard 802.11i: Amendment 6: Wireless LAN Medium Access Control (MAC) and Physical Layer (phy) Specifications. Draft 3 (2003)
5. Maitra, S., Paul, G., Sen Gupta, S.: Attack on Broadcast RC4 Revisited. In: Joux, A. (ed.) FSE 2011. LNCS, vol. 6733, pp. 199–217. Springer, Heidelberg (2011)
6. Mantin, I.: Analysis of the stream cipher RC4. Master's Thesis, The Weizmann Institute of Science, Israel (2001), `http://www.wisdom.weizmann.ac.il/~itsik/RC4/Papers/Mantin1.zip`
7. Mantin, I., Shamir, A.: A Practical Attack on Broadcast RC4. In: Matsui, M. (ed.) FSE 2001. LNCS, vol. 2355, pp. 152–164. Springer, Heidelberg (2002)
8. Maximov, A., Khovratovich, D.: New State Recovery Attack on RC4. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 297–316. Springer, Heidelberg (2008)
9. Paul, G., Maitra, S.: On biases of permutation and keystream bytes of RC4 towards the secret key. Cryptography Communications 1, 225–268 (2009)
10. Paul, G., Rathi, S., Maitra, S.: On Non-negligible bias of the first output byte of RC4 towards the first three bytes of the secret key. Designs, Codes and Cryptography 49(1-3), 123–134 (2008)
11. Roos, A.: A class of weak keys in the RC4 stream cipher. Two posts in sci.crypt, message-id `43u1eh$1j3@hermes.is.co.za`, `44ebge$llf@hermes.is.co.za` (1995), `http://marcel.wanda.ch/Archive/WeakKeys`
12. Sepehrdad, P., Vaudenay, S., Vuagnoux, M.: Discovery and Exploitation of New Biases in RC4. In: Biryukov, A., Gong, G., Stinson, D.R. (eds.) SAC 2010. LNCS, vol. 6544, pp. 74–91. Springer, Heidelberg (2011)
13. Sepehrdad, P., Vaudenay, S., Vuagnoux, M.: Statistical Attack on RC4. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 343–363. Springer, Heidelberg (2011)
14. Vaudenay, S., Vuagnoux, M.: Passive–Only Key Recovery Attacks on RC4. In: Adams, C., Miri, A., Wiener, M. (eds.) SAC 2007. LNCS, vol. 4876, pp. 344–359. Springer, Heidelberg (2007)
15. Wagner, D.: My RC4 weak keys. Post in sci.crypt, message-id `447o1l$cbj@cnn.Princeton.EDU`. (September 26, 1995), `http://www.cs.berkeley.edu/~daw/my-posts/my-rc4-weak-keys`