IEICE TRANS. FUNDAMENTALS, VOL.E93–A, NO.6 JUNE 2010

PAPER Special Section on Discrete Mathematics and Its Applications

New Analysis Based on Correlations of RC4 PRGA with Nonzero-Bit Differences**

Atsuko MIYAJI[†*a)], ***Member and*** **Masahiro SUKEGAWA**[††], ***Nonmember***

SUMMARY RC4 is the stream cipher proposed by Rivest in 1987, which is widely used in a number of commercial products because of its simplicity and substantial security. RC4 exploits shuffle-exchange paradigm, which uses a permutation S. Many attacks have been reported so far. No study, however, has focused on correlations in the Pseudo-Random Generation (PRGA) between two permutations S and S' with some differences, nevertheless such correlations are related to an inherent weakness of shuffle-exchange-type PRGA. In this paper, we investigate the correlations between S and S' with some differences in the initial round. We show that correlations between S and S' remain before "i" is in the position where the nonzero-bit difference exists in the initial round, and that the correlations remain with non negligible probability even after "i" passed by the position. This means that the same correlations between S and S' will be observed after the 255-th round. This reveals an inherent weakness of shuffle-exchange-type PRGA.

key words: *RC4, correlation, shuffle-exchange structure, pseudo key collision*

1. Introduction

RC4 is the stream cipher proposed by Rivest in 1987, which is widely used in a number of commercial products because of its simplicity and substantial security. Though many cryptanalysis of RC4 have been proposed so far [3]–[5], [7]–[9], [11], [12], [15]–[17], it has remained secure under proper use. As a result, RC4 is widely used in many applications such as Secure Sockets Layer (SSL), Wired Equivalent Privacy (WEP), etc.

RC4 exploits shuffle-exchange paradigm, which uses a permutation $S = (S[0], \cdots, S[N-1])$ given in the initial, and outputs 8-bit data in each round by updating the permutation S, where typically each $S[i]$ $(i \in [0, N-1])$ is 8 bits and $N = 256$. In more detail, RC4 consists of two algorithms, the Key Scheduling Algorithm (KSA) and the Pseudo Random Generation Algorithm (PRGA). KSA is given a secret key with ℓ bytes (typically, $5 \leq \ell \leq 16$) and generates the initial permutation S_0, which is an input of PRGA. PRGA is given the initial permutation S_0, uses two indices i and j, (where i is a public counter but j is one element of secret state information), updates S and j, and outputs $Z = S[S[i] + S[j]]$ as a key stream at the end of each round.

There are mainly two approaches to the cryptanalysis of RC4, analysis of the weaknesses of the KSA [1], [2], [13], and analysis of the weaknesses of the PRGA [5], [7], [11], [14]. In [5], the first state recovery attack was proposed, whose computational complexity is 2^{779}. In [7], [11], [14], non-uniform distribution of the initial permutation S_0 was shown. Recently, the state key recovery attack is improved by [6], whose computational complexity is 2^{241}.

Many works, however, focus on the bias between a secret key and the initial permutation, which is an input of PRGA. Some analysis of the weaknesses of the PRGA also focus on the correlation between the first keystream output of PRGA and the secret key. We have not seen any research on correlations in PRGA between two permutations with some differences. However, such correlations should be investigated, since it is reported that sets of two keys which output either the same initial permutations or initial permutations with differences of just a few bits can be intentionally induced [10]. Furthermore, correlations between outputs of two consecutive rounds is an inherent weakness of shuffle-exchange-type PRGA.

In this paper, we focus on a shuffle-exchange structure of PRGA, where 1 swap is executed in each round. We investigate how the structure mixes the permutation S, by observing correlations between two permutations, S and S', with some differences in the initial round. The set of indices where differences exist in the initial round is represented by Diff_0. The correlations are measured over (a) the difference value of two permutations $\Delta S = S \oplus S'$, (b) the difference value of two outputs of PRGA, $\Delta Z = Z \oplus Z'$, and (c) the difference value of two indices $\Delta j = j \oplus j'$. We start with $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$. Our results, however, are easily applicable to other cases where there exist differences Diff_0 with $\#\mathsf{Diff}_0 > 2$ in the initial round.

We show theoretically that correlations between two permutations S and S', such as $\Delta Z = 0$, $\Delta j = 0$, and the hamming weight of ΔS, remain when $i < \mathsf{df}_0[1]$. Furthermore, we show that such correlations between two permutations S and S' remain with non negligible probability when $i \geq \mathsf{df}_0[1]$, thus, the same correlations between permutations will be observed when $i < \mathsf{df}_0[2]$. For example, the probability that such correlations remain when $i > \mathsf{df}_0[1]$ is greater than 30% in the cases of $\mathsf{df}_0[1] \geq 93$. We give the theoretical formulae of the probability of both outputs being equal when $i = \mathsf{df}_0[1]$. All theoretical results have been

Manuscript received September 25, 2009.
Manuscript revised December 12, 2009.

[†]The author is with Japan Advanced Institute of Science and Technology, Nomi-shi, 923-1292 Japan.
[††]The author is with NEC, Japan. This work was conducted when he was with JAIST.

*This study is partly supported by Grant-in-Aid for Scientific Research (B), 203000032.
**A preliminary version was presented at ACISP 2009.

a) E-mail: miyaji@jaist.ac.jp

DOI: 10.1587/transfun.E93.A.1066

confirmed experimentally.

This paper is organized as follows. Section 2 summarizes the known facts on RC4 together with notation. Section 3 investigates correlations in each round between two permutations S and S' with some differences in the initial round. Section 4 investigates correlations in each round between outputs of two permutations S and S'. Section 5 shows the experimental results which confirm all theories in Sects. 3 and 4. Section 6 investigates how to predict inner states.

2. Preliminary

This section presents the KSA and the PRGA of RC4, after explaining the notations used in this paper.

- S, S': permutations
- S_0, S'_0: the initial permutations of PRGA
- Diff_0: the set of indices where differences between S and S' exist in the initial round
- r: number of rounds ($r = 0$ means the initial round)
- $\mathsf{df}_0[1], \mathsf{df}_0[2]$: the positions where differences exist in the initial round
- i_r, j_r (j'_r): i and j (j') of S (S') after r rounds
- S_r (S'_r): the permutation S (S') after r rounds
- $S_r[i]$ ($S'_r[i]$): the value of S_r (S'_r) in the position i after r rounds
- ΔS_r: $S_r \oplus S'_r$
- $\Delta S_r[i]$: $S_r[i] \oplus S'_r[i]$
- $|\Delta S_r|$: the number of indices with $\Delta S_r[i] \neq 0$
- Z_r (Z'_r): the output under S (S') at the r-th round
- ΔZ_r: $Z_r \oplus Z'_r$
- Δj_r: $j_r \oplus j'_r$
- $\Delta\mathsf{State}[0], \Delta\mathsf{State}[1], \cdots$: the state differences between S and S' (j and j') in a round r. (The state differences of i are omitted since the same i is used each other.)

RC4 has a secret internal state which is a permutation of all the $N = 2^n$ possible n-bit words and index j. RC4 generates a pseudo-random stream of bits (a keystream) which, for encryption, is combined with the plaintext using XOR; decryption is performed in the same way. To generate the keystream, the cipher makes use of a secret internal state which consists of two parts (shown in Fig. 1): A key scheduling algorithm, KSA, which turns a random key (whose typical size is 40–256 bits) into an initial permutation S_0 of $\{0, \ldots, N-1\}$, and an output generation algorithm, PRGA, which uses the initial permutation to generate a pseudo-random output sequence.

The algorithm KSA consists of N loops. It initializes S to be the identity permutation, and both i and j to 0, and then repeats three simple operations: increment i, which acts as a counter, set j by using S and a secret key K with ℓ bytes where each word contains n bits, and swap two values of S in positions i and j. Finally, it outputs a random permutation $S = S_0$.

The algorithm PRGA is similar to KSA. It repeats four simple operations: increment i, which act as a counter, set j by using S and the previous j, swap two values of S in positions i and j, and output the value of S in position $S[i]+S[j]$. Each value of S is swapped at least once (possibly with itself) within any N consecutive rounds. All additions used in both KSA and PRGA are in general additions modulo N unless specified otherwise.

KSA(K)
Initialization
For $i = 0 \ldots N-1$
$S[i] = i$
$j = 0$
Scrambling:
For $i = 0 \ldots N-1$
$j = j + S[i] + K[i \pmod{\ell}]$
Swap$(S[i], S[j])$

PRGA(K)
Initialization:
$i = 0$
$j = 0$
Generation loop:
$i = i + 1$
$j = j + S[i]$
Swap$(S[i], S[j])$
Output $z = S[S[i] + S[j]]$

Fig. 1 The key scheduling algorithm and the pseudo-random generation algorithm.

3. State Analysis of Permutations with Some Differences

This section analyzes correlations between two permutations, S and S', with some differences in the initial round. The set of indices where differences exist in the initial round is represented by $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2], \cdots\}$. The indices with nonzero bit differences are arranged in order of positions that i will reach after the next round. Therefore, if nonzero bit differences exist in positions 0 and $N-1$ in the initial round, then $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\} = \{N-1, 0\}$ since i will be incremented to 1 in the first round.

3.1 Overview of Analysis

Assume that two permutations S and S' with $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$ in the initial round are given, where $(S_0[\mathsf{df}_0[1]], S_0[\mathsf{df}_0[2]]) = (a, b)$ and $(S'_0[\mathsf{df}_0[1]], S'_0[\mathsf{df}_0[2]]) = (b, a)$ (See Fig. 2). The initial state of differences between S_0 and S'_0 is:

$$\Delta\mathsf{State}[0] : (\Delta S[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_0) \wedge (\Delta j = 0).$$

We analyze the conditions in each round in which the initial state changes from the current state to another, or remains the same.

The transitions of state are different according to the position of i, that is, $i < \mathsf{df}_0[1]$; $i = \mathsf{df}_0[1]$ and the nonzero bit difference still exists in the position $\mathsf{df}_0[1]$; $i = \mathsf{df}_0[1]$ but the nonzero bit difference does not exist in the position

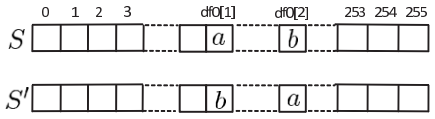

Fig. 2 ΔState[0].

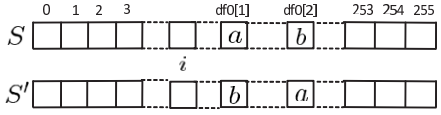

Fig. 3 Event[1].

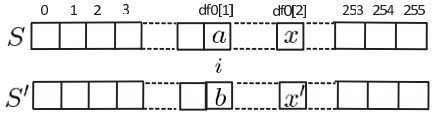

Fig. 4 Event[2].

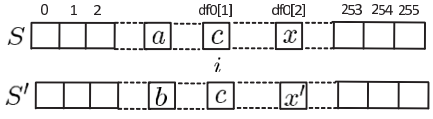

Fig. 5 Event[3].

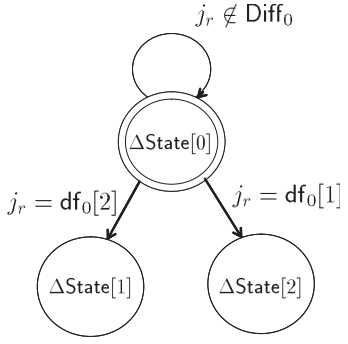

Fig. 6 State Diagram of PRGA in Event[1].

$\mathsf{df}_0[1]$, which are formalized as follows.

$$\begin{aligned}
&\mathsf{Event}[1]\colon i_r < \mathsf{df}_0[1] && \text{(Fig. 3)},\\
&\mathsf{Event}[2]\colon [i_r = \mathsf{df}_0[1]] \wedge [\Delta S_{r-1}[\mathsf{df}_0[1]] \neq 0] && \text{(Fig. 4)},\\
&\mathsf{Event}[3]\colon [i_r = \mathsf{df}_0[1]] \wedge [\Delta S_{r-1}[\mathsf{df}_0[1]] = 0] && \text{(Fig. 5)}.
\end{aligned}$$

Figures 3, 4, and 5 show each event, where $(x, x') = (b, a)$ or $x = x'$. We will see the reason for this in the following subsections. The following subsections describe each transition and the probability of its occurrence in each event. We will see that the state of differences between two permutations S and S' has the Markov property, that is, given the state in a certain round (the present state), the state in a future round (future states) is independent of past rounds.

3.2 Transitions of ΔState[0] before the Nonzero Bit Difference

This subsection shows the transitions from the initial state in Event[1] and their associated probabilities. The state diagram is given in Fig. 6.

Theorem 1: Assume that two initial permutations S and S' are in the state of differences ΔState[0] in the $(r-1)$-th round, and that Event[1] occurs in the r-th round.
(1) The state changes to the state ΔState[0] (resp. ΔState[1], resp. ΔState[2]) if $j_r \notin \mathsf{Diff}_0$ (resp. $j_r = \mathsf{df}_0[2]$, resp. $j_r = \mathsf{df}_0[1]$), where

$$\begin{aligned}
&\Delta\mathsf{State}[0]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_0] \wedge [\Delta j_r = 0],\\
&\Delta\mathsf{State}[1]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_1] \wedge [\Delta j_r = 0],\\
&\Delta\mathsf{State}[2]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_2] \wedge [\Delta j_r = 0].
\end{aligned}$$

and where $\mathsf{Diff}_1 = \{\mathsf{df}_1[1], \mathsf{df}_1[2]\} = \{\mathsf{df}_0[1], i_r\}$ and $\mathsf{Diff}_2 = \{\mathsf{df}_2[1], \mathsf{df}_2[2]\} = \{\mathsf{df}_0[2], i_r\}$.
(2) Each transition occurs with the following probabilities if j is distributed randomly:

$$\begin{aligned}
\mathsf{Prob}\,[\Delta\mathsf{State}[0]] &= \frac{N-2}{N},\\
\mathsf{Prob}\,[\Delta\mathsf{State}[1]] &= \frac{1}{N}, \quad \text{and}\\
\mathsf{Prob}\,[\Delta\mathsf{State}[2]] &= \frac{1}{N},
\end{aligned}$$

where each probability is taken over choices of S and S' in state ΔState[0] in the initial round.

proof: (1) It is clear that $j_r = j'_r$ holds in any case, since $j_r = j_{r-1} + S_{r-1}[i_r]$, $\Delta j_{r-1} = 0$, and $i_r \notin \mathsf{Diff}_0$. In the case of $j_r \notin \mathsf{Diff}_0$, $\Delta S_{r-1}[i_r] = \Delta S_{r-1}[j_r] = 0$ holds and, thus, positions of non-zero-bit differences remain the same as those in $(r-1)$-round. Therefore, ΔState[0] occurs. In the case of $j_r = \mathsf{df}_0[2]$,

$$\begin{aligned}
(S_r[i_r], S_r[j_r]) &= (S_{r-1}[j_r], S_{r-1}[i_r]) = (b, S_{r-1}[i_r]);\\
(S'_r[i_r], S'_r[j'_r]) &= (S'_{r-1}[j'_r], S'_{r-1}[i_r]) = (a, S'_{r-1}[i_r]);
\end{aligned}$$

and, thus, the non-zero-bit difference in the position $\mathsf{df}_0[2]$ moves to the current i_r. Therefore, ΔState[1] occurs. In the case of $j_r = \mathsf{df}_0[1]$,

$$\begin{aligned}
(S_r[i_r], S_r[j_r]) &= (S_{r-1}[j_r], S_{r-1}[i_r]) = (a, S_{r-1}[i_r]);\\
(S'_r[i_r], S'_r[j'_r]) &= (S'_{r-1}[j'_r], S'_{r-1}[i_r]) = (b, S'_{r-1}[i_r]);
\end{aligned}$$

and, thus, the non-zero-bit difference in the position $\mathsf{df}_0[1]$ moves to the current i_r. Therefore, ΔState[2] occurs.

(2) The probability that each state will occur follows from the above discussion. ∎

Theorem 1 implies that

- $|\Delta S_r| = 2$ and $\Delta j_r = 0$ hold as long as i_r is not equal to the position that a nonzero bit difference exits in the initial round.
- If $j_r = \mathsf{df}_0[1]$ at least once in the r-th round during $i_r < \mathsf{df}_0[1]$, then the nonzero bit difference in the position $\mathsf{df}_0[1]$ moves to the current i_r. As a result, the nonzero-bit difference that was originally in the position $\mathsf{df}_0[1]$ affects neither $|\Delta S|$ nor Δj until the $(r+N-1)$-th round.

This is the case in which Event[3] occurs.

The following corollary describes the detailed cases in which i is not equal to any position that a nonzero bit difference existed originally before the N-th round.

Corollary 1: Assume that two initial permutations S and S' in the state of differences ΔState[0] are given. Then, if either of the following events occurs, then i is not equal to any position that a nonzero bit difference exits; and both $|\Delta S_r| = 2$ and $\Delta j_r = 0$ hold until the N-th round.

$$\begin{aligned}\mathsf{Event}[4]: \quad & [j_{r_1} = \mathsf{df}_0[1] \wedge 1 \le \exists i_{r_1} < \mathsf{df}_0[1]] \\ \wedge \quad & [j_{r_2} = \mathsf{df}_0[2] \wedge i_{r_1} < \exists i_{r_2} < \mathsf{df}_0[2]] \\ \mathsf{Event}[5]: \quad & [j_{r_3} = \mathsf{df}_0[2] \wedge 1 \le \exists i_{r_3} < \mathsf{df}_0[1] - 1] \\ \wedge \quad & [j_{r_4} = \mathsf{df}_0[1] \wedge i_{r_3} < \exists i_{r_4} < \mathsf{df}_0[1]]\,.\end{aligned}$$

Note that i_{r_3} is less than $\mathsf{df}_0[1] - 1$ since $i_{r_3} < i_{r_4} < \mathsf{df}_0[1]$.

proof: Assume that Event[4] has occurred in (j_{r_1}, j_{r_2}), that is, first $j_{r_1} = \mathsf{df}_0[1]$ for $1 \le i_{r_1} < \mathsf{df}_0[1]$ has occurred. This means that ΔState[2] has occurred in the index of i_{r_1} and, thus, $\Delta S_{r_1}[x] \neq 0 \iff x \in \mathsf{Diff}_2$. Therefore, the nonzero-bit difference in the position $\mathsf{df}_0[1]$ moves to the position i_{r_1}. Next, it is assumed that $j_{r_2} = \mathsf{df}_0[2] (i_{r_1} < i_{r_2} < \mathsf{df}_0[2])$ has occurred. Then, $\Delta S_{r_2}[x] \neq 0 \iff x \in \{i_{r_1}, i_{r_2}\}$ by applying Theorem 1 to Diff_2. Thus, i is not equal to any position that a nonzero bit difference exits until the N-th round.

Assume that Event[5] has occurred in (j_{r_3}, j_{r_4}), that is, first $j_{r_3} = \mathsf{df}_0[2]$ for $1 \le i_{r_3} < \mathsf{df}_0[1] - 1$ has occurred. This means that ΔState[1] has occurred in the index of i_{r_3} and, thus, $\Delta S_{r_3}[x] \neq 0 \iff x \in \mathsf{Diff}_1$. Then, the index $\mathsf{df}_0[2]$ no longer indicates a nonzero bit difference. Next, it is assumed that $j_{r_4} = \mathsf{df}_0[1] (i_{r_3} < i_{r_4} < \mathsf{df}_0[1])$ has occurred. Then, $\Delta S_{r_4}[x] \neq 0 \iff x \in \{i_{r_3}, i_{r_4}\}$ by applying Theorem 1 to Diff_1. Thus, i is not equal to any position that a nonzero bit difference exits until the N-th round. ∎

Event[3] occurs if and only if $j_r = \mathsf{df}_0[1]$ for $1 \le \exists r < \mathsf{df}_0[1]$. The probability that Event[3] occurs, Prob [Event[3]], is computed by the next theorems.

Theorem 2: Assume that two initial permutations S and S' in the state of differences ΔState[0] with $\mathsf{df}_0[1] \ge 5$ are given. Then, the probability that Event[3] will occur in $\mathsf{df}_0[1] \ge 5$ is given as follows if each j is distributed randomly for any given S and S':

$$\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] \ge 5] = 1 - \left(\frac{N-1}{N}\right)^{\mathsf{df}_0[1]-1},$$

where the probability is taken over choices of S and S' with differences in Diff_0 in the initial round.

proof: Event[2], the complement of Event[3], occurs if and only if $j \neq \mathsf{df}_0[1]$ during $i < \mathsf{df}_0[1]$. Therefore, $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] \ge 5] = 1 - \left(\frac{N-1}{N}\right)^{\mathsf{df}_0[1]-1}$ if j is distributed randomly. ∎

In the case of $\mathsf{df}_0[1] < 5$, we can describe Prob [Event[3]] by the conditions of S_0 as follows:

Theorem 3: Assume that two initial permutations S and S' in the state of differences ΔState[0] with $\mathsf{df}_0[1] < 5$ are given. Then, Event[3] will occur in the following probability if $S_0[1]$, $S_0[2]$, and $S_0[3]$ are distributed randomly:
(1) *In the case of* $\mathsf{df}_0[1] = 2$, $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 2] = \mathsf{Prob}\,[S_0[1] = j_1 = 2] = \frac{1}{N}$,
(2) *In the case of* $\mathsf{df}_0[1] = 3$,

$$\begin{aligned}& \mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 3] \\ = \;& \mathsf{Prob}\,[S_0[1] = 3 \mid \mathsf{df}_0[1] = 3] + \\ & \mathsf{Prob}\,[S_0[1] \neq 2, 3 \wedge S_0[1] + S_0[2] = 3] \\ = \;& \frac{2N-3}{N(N-1)},\end{aligned}$$

(3) *In the case of* $\mathsf{df}_0[1] = 4$,

$$\begin{aligned}& \mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 4] \\ = \;& \mathsf{Prob}\,[S_0[1] = 2] + \mathsf{Prob}\,[S_0[1] = 4] + \\ & \mathsf{Prob}\,[S_0[1] = 3 \wedge S_0[2] = N-2] + \\ & \mathsf{Prob}\,[S_0[1] \neq 2, 3, 4] \\ & \wedge S_0[3] \neq 0, 1 \wedge S_0[1] + S_0[2] + S_0[3] = 4 \\ = \;& \frac{2(2N-3)}{N(N-1)},\end{aligned}$$

where the probability is taken over choices of S and S' with differences in Diff_0 in the initial round.

proof: (1) Let $\mathsf{df}_0[1] = 2$. Then, Event[3] occurs if and only if $j_1 = \mathsf{df}_0[1] = 2$, where $j_1 = j_0 + S_0[1] = S_0[1]$. Therefore, $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 2] = \mathsf{Prob}\,[S_0[1] = 2] = \frac{1}{N}$.
(2) Let $\mathsf{df}_0[1] = 3$. Then, Event[3] occurs if and only if $j_1 = \mathsf{df}_0[1] = 3$ or $j_2 = \mathsf{df}_0[1] = 3$. If $S_0[1] = 3$, then we get $j_1 = j_0 + S_0[1] = S_0[1] = 3 = \mathsf{df}_0[1]$. If $S_0[1] \neq 2$, then $S_0[1] = j_1 \neq 2$, which means that $S_0[1]$ is not swapped with $S_0[2]$ in the first round. This implies that $S_1[2] = S_0[2]$. Thus, if $[S_0[1] \neq 2, 3] \wedge [S_0[1] + S_0[2] = 3]$, we get $j_2 = j_1 + S_1[2] = S_0[1] + S_0[2] = 3 = \mathsf{df}_0[1]$. Therefore, $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 3] = \frac{1}{N} + \frac{N-2}{N(N-1)} = \frac{2N-3}{N(N-1)}$.
(3) Let $\mathsf{df}_0[1] = 4$. Then, Event[3] occurs if and only if $j_1 = \mathsf{df}_0[1] = 4$, $j_2 = \mathsf{df}_0[1] = 4$, or $j_3 = \mathsf{df}_0[1] = 4$. If $S_0[1] = 4$, then we get $j_1 = j_0 + S_0[1] = S_0[1] = 4 = \mathsf{df}_0[1]$. If $S_0[1] = 2$, then $j_1 = j_0 + S_0[1] = 2$; $S_0[1]$ is swapped with $S_0[2]$; and, we get $j_2 = j_1 + S_1[2] = j_1 + S_0[1] = 4 = \mathsf{df}_0[1]$. Note that $S_0[1]$ is swapped with $S_0[2]$ if and only if $S_0[1] = 2$. If $S_0[1] \neq 2, 4$ and $S_0[1] + S_0[2] = 4$, then we get $j_2 = j_1 + S_1[2] = S_0[1] + S_0[2] = 4 = \mathsf{df}_0[1]$.

If $S_0[1] = 3$ and $S_0[2] = N - 2$, then $j_1 = S_0[1] = 3$; and $S_0[1]$ is swapped with $S_0[3]$, which implies that $(S_1[1], S_1[3]) = (S_0[3], S_0[1])$. Then, in the 2nd round, $j_2 = j_1 + S_1[2] = 3 + S_0[2] = 1$; and $S_1[2]$ is swapped with $S_1[1]$, which implies that $S_2[3] = S_1[3] = 3$. Thus, in the 3rd round, we get $j_3 = j_2 + S_2[3] = 4$. Note that $S_0[1]$ is swapped with $S_0[3]$ if and only if $S_0[1] = 3$.

If $S_0[1] \neq 2, 3, 4$; $S_0[3] \neq 0, 1$; and $S_0[1] + S_0[2] + S_0[3] = 4$, then $S_1[2] = S_0[2]$; $S_1[3] = S_0[3]$; and $S_0[1] + S_0[2] \neq 3$. This implies that $S_1[3]$ is not swapped with $S_1[2]$ and that $S_2[3] = S_1[3]$. Thus, we get $j_3 = S_0[1] + S_0[2] + S_0[3] = 4$. To sum up all conditions, which are independent of each other, $\mathsf{Prob}\,[\mathsf{Event}[3]] = \frac{2}{N} + \frac{N-2}{N(N-1)} + \frac{1}{N(N-1)} + \frac{N-3}{N(N-1)} = \frac{2(2N-3)}{N(N-1)}$. ∎

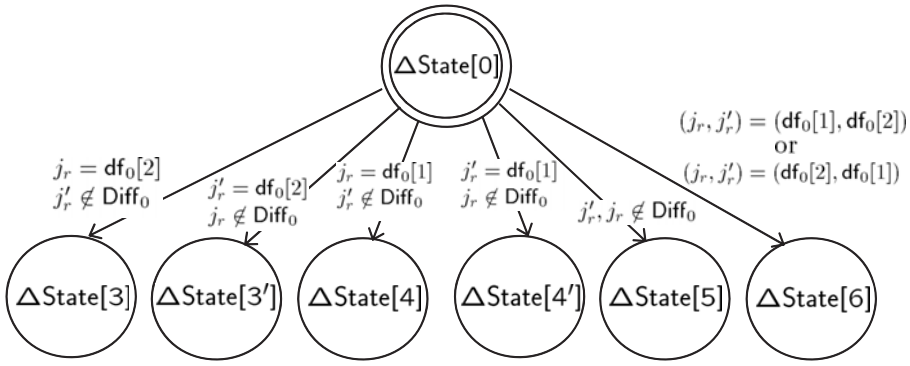

Fig. 7 State diagram of PRGA in Event[2].

3.3 Transitions of ΔState[0] on the Nonzero Bit Difference

This subsection shows each transition of the initial state ΔState[0] and the probability of its occurrence when Event[2] occurs. The state diagram is given in Fig. 7.

Theorem 4: Assume that two permutations S and S' are in the state of differences ΔState[0] in the $(r-1)$-th round. (1) The state changes to ΔState[3] (resp. ΔState[3′], resp. ΔState[4], resp. ΔState[4′], resp. ΔState[5], resp. ΔState[6]), if $[j_r = \mathsf{df}_0[2]] \wedge [j'_r \notin \mathsf{Diff}_0]$ (resp. $[j'_r = \mathsf{df}_0[2]] \wedge [j_r \notin \mathsf{Diff}_0]$, resp. $[j_r = \mathsf{df}_0[1]] \wedge [j'_r \notin \mathsf{Diff}_0]$, resp. $[j'_r = \mathsf{df}_0[1]] \wedge [j_r \notin \mathsf{Diff}_0]$, resp. $j'_r, j_r \notin \mathsf{Diff}_0$, resp. $j'_r, j_r \in \mathsf{Diff}_0$), where

$$
\begin{aligned}
&\Delta\mathsf{State}[3]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_3] \ \wedge [\Delta j_r \neq 0],\\
&\Delta\mathsf{State}[3']\colon [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_{3'}] \wedge [\Delta j_r \neq 0],\\
&\Delta\mathsf{State}[4]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_4] \ \wedge [\Delta j_r \neq 0],\\
&\Delta\mathsf{State}[4']\colon [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_{4'}] \wedge [\Delta j_r \neq 0],\\
&\Delta\mathsf{State}[5]\colon\ [\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_5] \ \wedge [\Delta j_r \neq 0],\\
&\Delta\mathsf{State}[6]\colon\ [|\Delta S_r| = 0] \qquad\qquad\quad \wedge [\Delta j_r \neq 0],
\end{aligned}
$$

where

$$
\begin{aligned}
\mathsf{Diff}_3 &= \{\mathsf{df}_3[1], \mathsf{df}_3[2]\}\\
&= \{\mathsf{df}_0[1], j'_r\} = \{i_r, j'_r\},\\
\mathsf{Diff}_{3'} &= \{\mathsf{df}_{3'}[1], \mathsf{df}_{3'}[2]\}\\
&= \{\mathsf{df}_0[1], j_r\} = \{i_r, j_r\},\\
\mathsf{Diff}_4 &= \{\mathsf{df}_4[1], \mathsf{df}_4[2], \mathsf{df}_4[3]\}\\
&= \{\mathsf{df}_0[1], \mathsf{df}_0[2], j'_r\} = \{i_r, \mathsf{df}_0[2], j'_r\},\\
\mathsf{Diff}_{4'} &= \{\mathsf{df}_{4'}[1], \mathsf{df}_{4'}[2], \mathsf{df}_{4'}[3]\}\\
&= \{\mathsf{df}_0[1], \mathsf{df}_0[2], j_r\} = \{i_r, \mathsf{df}_0[2], j_r\},\\
\mathsf{Diff}_5 &= \{\mathsf{df}_5[1], \mathsf{df}_5[2], \mathsf{df}_5[3], \mathsf{df}_5[4]\}\\
&= \{\mathsf{df}_0[1], \mathsf{df}_0[2], j_r, j'_r\} = \{i_r, \mathsf{df}_0[2], j_r, j'_r\}.
\end{aligned}
$$

(2) Each transition occurs with the following probability, if j is distributed randomly:

$$
\begin{aligned}
&\mathsf{Prob}\,[\Delta\mathsf{State}[3] \vee \Delta\mathsf{State}[3']] = \mathsf{Prob}\,[\mathsf{Event}[2]] \cdot \tfrac{2(N-2)}{N(N-1)},\\
&\mathsf{Prob}\,[\Delta\mathsf{State}[4] \vee \Delta\mathsf{State}[4']] = \mathsf{Prob}\,[\mathsf{Event}[2]] \cdot \tfrac{2(N-2)}{N(N-1)},\\
&\mathsf{Prob}\,[\Delta\mathsf{State}[5]] \qquad\qquad = \mathsf{Prob}\,[\mathsf{Event}[2]] \cdot \tfrac{(N-2)(N-3)}{N(N-1)},\\
&\mathsf{Prob}\,[\Delta\mathsf{State}[6]] \qquad\qquad = \mathsf{Prob}\,[\mathsf{Event}[2]] \cdot \tfrac{2}{N(N-1)}.
\end{aligned}
$$

proof: (1) It is clear that $j_r \neq j'_r$ in each case, since $\Delta j_r = \Delta j_{r-1} + \Delta S_{r-1}[i_r] = \Delta S_{r-1}[i_r] \neq 0$. In the case of $j_r = \mathsf{df}_0[2]$ and $j'_r \notin \mathsf{Diff}_0$, $S_{r-1}[i_r] = S_{r-1}[\mathsf{df}_0[1]] = a$ is swapped with $S_{r-1}[j_r] = b$; $S'_{r-1}[i_r] = S'_{r-1}[\mathsf{df}_0[1]] = b$ is swapped with $S'_{r-1}[j_r]$, which implies that $S'_{r-1}[\mathsf{df}_0[2]] = a$ remains the same. Thus, we get $\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_3$ after the r-th round. In the case of $j'_r = \mathsf{df}_0[2]$ and $j_r \notin \mathsf{Diff}_0$, the same also holds.

In the case of $j_r = \mathsf{df}_0[1]$ and $j'_r \notin \mathsf{Diff}_0$, $i_r = j_r = \mathsf{df}_0[1]$ occurs; $S_{r-1}[i_r] = S_{r-1}[j_r] = a$ remains the same; and $S'_{r-1}[i_r] = b$ is swapped with $S'_{r-1}[j_r]$, Thus, we get $\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_4$ after the r-th round. In the case of $j'_r = \mathsf{df}_0[1]$ and $j_r \notin \mathsf{Diff}_0$, the same also holds.

In the case of $j'_r, j_r \notin \mathsf{Diff}_0$, $S_{r-1}[i_r] = a$ (resp. $S'_{r-1}[i_r] = b$) is swapped with $S_{r-1}[j_r]$ (resp. $S'_{r-1}[j'_r]$), where nonzero-bit difference did not exist; and both $S_{r-1}[\mathsf{df}_0[2]] = b$ and $S'_{r-1}[\mathsf{df}_0[2]] = a$ still remain the same. Thus, we get $\Delta S_r[x] \neq 0 \Longleftrightarrow x \in \mathsf{Diff}_5$ after the r-th round.

In the case of $(j_r, j'_r) = (\mathsf{df}_0[1], \mathsf{df}_0[2])$, $S'_{r-1}[i_r] = S'_{r-1}[\mathsf{df}_0[1]] = b$ is swapped with $S'_{r-1}[j'_r] = S'_{r-1}[\mathsf{df}_0[2]] = a$ while both $S_{r-1}[i_r] = S_{r-1}[j_r] = a$ and $S_{r-1}[j_r] = b$ remain the same. Thus, all nonzero-bit differences disappear after swapping in the r-th round. The same also holds in the case of $(j_r, j'_r) = (\mathsf{df}_0[2], \mathsf{df}_0[1])$.
(2) The probability that each state will occur follows from the above discussion. ∎

4. Correlation between Outputs and State Transitions

This section analyzes the differences between outputs of two permutations S and S' in each transition described in Sect. 3, where two initial permutations S and S' are in the state of differences ΔState[0].

4.1 Outputs before the Nonzero-Bit Difference

This subsection investigates the correlation between outputs of two permutations in each transition before the first nonzero-bit difference (i.e. $i < \mathsf{df}_0[1]$). The states of differences of two permutations in any round $r < \mathsf{df}_0[1]$ are ΔState[0], ΔState[1], or ΔState[2] from Theorem 1. The probability that both outputs of permutations are equal,

$\mathsf{Prob}\,[\Delta Z = 0]$, is given in the next theorem.

Proposition 1: Assume that two initial permutations S and S' are in the state of differences $\Delta\mathsf{State}[0]$ in the $(r-1)$-th round, and that $\mathsf{Event}[1]$ occurs in the r-th round. Then, $\mathsf{Prob}\,[\Delta Z = 0]$ in each state is as follows:

$$\mathsf{Prob}\,[\Delta Z = 0] = \frac{N-2}{N}, \frac{2}{N(N-1)}, \text{ or } \frac{2}{N(N-1)}$$

if $\Delta\mathsf{State}[0]$, $\Delta\mathsf{State}[1]$, or $\Delta\mathsf{State}[2]$ occurs, respectively.

proof: Theorem 1 has shown that

- $\Delta j_r = 0$ and $j_r, i_r \notin \mathsf{Diff}_0$ if $\Delta\mathsf{State}[0]$,
- $\Delta j_r = 0$, $i_r \in \mathsf{Diff}_1$ and $j_r \notin \mathsf{Diff}_1$ if $\Delta\mathsf{State}[1]$,
- $\Delta j_r = 0$, $i_r \in \mathsf{Diff}_1$ and $j_r \notin \mathsf{Diff}_2$ if $\Delta\mathsf{State}[2]$.

Then, the necessary and sufficient conditions for $\Delta Z = 0$ in each state are as follows.

In $\Delta\mathsf{State}[0] : \Delta Z = 0$
$\iff [\Delta(S_r[i_r] + S_r[j_r]) = 0] \wedge [S_r[i_r] + S_r[j_r] \notin \mathsf{Diff}_0]$
$\iff S_r[i_r] + S_r[j_r] \notin \mathsf{Diff}_0$
Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{N-2}{N}$.

In $\Delta\mathsf{State}[1] : \Delta Z = 0$
$\iff [\Delta(S_r[i_r] + S_r[j_r]) \neq 0]$
$\wedge [S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j_r] \in \mathsf{Diff}_1]$
$\iff S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j_r] \in \mathsf{Diff}_1$
Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{2}{N(N-1)}$ since $\#\mathsf{Diff}_1 = 2$ and $S_r[i_r] + S_r[j_r] \neq S'_r[i_r] + S'_r[j_r]$.

In $\Delta\mathsf{State}[2] : \Delta Z = 0$
$\iff [\Delta(S_r[i_r] + S_r[j_r]) \neq 0]$
$\wedge [S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j_r] \in \mathsf{Diff}_2]$
$\iff S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j_r] \in \mathsf{Diff}_2$
Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{2}{N(N-1)}$ since $\#\mathsf{Diff}_2 = 2$ and $S_r[i_r] + S_r[j_r] \neq S'_r[i_r] + S'_r[j_r]$.

From the above, Proposition 1 follows. ∎

From Theorem 1 and Proposition 1, the probability of $\mathsf{Prob}\,[\Delta Z = 0]$ if $r < \mathsf{df}_0[1]$ (i.e. $i < \mathsf{df}_0[1]$) can be computed as follows.

Corollary 2: Assume that two initial permutations S and S' with $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$ are given. Then, $\mathsf{Prob}\,[\Delta Z = 0] = \left(\frac{N-2}{N}\right)^2 + \frac{4}{N^2(N-1)}$, if $r < \mathsf{df}_0[1]$.

4.2 Outputs on the Nonzero-Bit Difference

This subsection investigates the correlation between outputs of two permutations in each transition when $r = \mathsf{df}_0[1]$ (i.e. $i = \mathsf{df}_0[1]$). The probability that both outputs are equal, $\mathsf{Prob}\,[\Delta Z = 0]$, is given in the next theorem.

Proposition 2: Assume that two initial permutations S and S' are in the state of differences $\Delta\mathsf{State}[0]$ in the $(r-1)$-th round, and that $\mathsf{Event}[2]$ occurs in the r-th round. Then, $\mathsf{Prob}\,[\Delta Z = 0]$ in each state is as follows:

$\mathsf{Prob}\,[\Delta Z = 0] =$

$$\begin{cases} \frac{2}{N(N-1)} & \text{if } \Delta\mathsf{State}[3] \vee \Delta\mathsf{State}[3'] \\ \frac{N-3}{N(N-2)} + \frac{3}{N(N-1)} & \text{if } \Delta\mathsf{State}[4] \vee \Delta\mathsf{State}[4'] \\ \frac{N-4}{N(N-3)} + \frac{4}{N(N-1)} & \text{if } \Delta\mathsf{State}[5] \\ 0 & \text{if } \Delta\mathsf{State}[6] \end{cases}$$

proof: Let c and $c' \in [0, N-1]$ be values in positions of j_r and j'_r before swapping in the r-th round, that is, $(c, c') = (S_{r-1}[j_r], S'_{r-1}[j'_r])$. On the other hand, $(a, b) = (S_{r-1}[\mathsf{df}_0[1]], S_{r-1}[\mathsf{df}_0[2]]) = (S'_{r-1}[\mathsf{df}_0[2]], S'_{r-1}[\mathsf{df}_0[1]])$. (See Fig. 2). Theorem 4 has shown that:

$\Delta\mathsf{State}[3]$: $(S_r[i_r], S_r[j_r]) = (b, a) \wedge (S'_r[i_r], S'_r[j'_r]) = (c', b)$ (i.e. $c = b$ and $c' \neq a, b$);
$\Delta\mathsf{State}[4]$: $(S_r[i_r], S_r[j_r]) = (a, a) \wedge (S'_r[i_r], S'_r[j'_r]) = (c', b)$ (i.e. $a = c$ and $c' \neq a, b$);
$\Delta\mathsf{State}[5]$: $(S_r[i_r], S_r[j_r]) = (c, a) \wedge (S'_r[i_r], S'_r[j'_r]) = (c', b)$ (i.e. $c' \neq c$ and $c', c \neq a, b$);
$\Delta\mathsf{State}[6]$: $\Delta S_r = 0$, $(S_r[i_r], S_r[j_r]) = (a, a) \wedge (S'_r[i_r], S'_r[j'_r]) = (a, b)$ (i.e. $i_r = j_r$);
or $\Delta S_r = 0$, $(S_r[i_r], S_r[j_r]) = (b, a) \wedge (S'_r[i_r], S'_r[j'_r]) = (b, b)$ (i.e. $i_r = j'_r$).

Therefore, the necessary and sufficient conditions of $\Delta Z = 0$ in each state are as follows.

In $\Delta\mathsf{State}[3] : \Delta Z = 0$
$\iff [S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j'_r] \in \mathsf{Diff}_3]$
$\wedge [\Delta(S_r[i_r] + S_r[j_r]) \neq 0]$
$\iff [(a + b, c' + b) = (\mathsf{df}_0[1], j'_r), (j'_r, \mathsf{df}_0[1])]$.
Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{2}{N(N-1)}$ since $a + b \neq c' + b$.
The same reasoning holds in the case of $\Delta\mathsf{State}[3']$.
In $\Delta\mathsf{State}[4] : \Delta Z = 0$
$\iff [[\Delta(S_r[i_r] + S_r[j_r]) = 0] \wedge [S_r[i_r] + S_r[j_r] \notin \mathsf{Diff}_4]]$
$\bigvee [[\Delta(S_r[i_r] + S_r[j_r]) \neq 0] \wedge$
$[S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j'_r] \in \mathsf{Diff}_4]$
$\wedge S_r[S_r[i_r] + S_r[j_r]] = S'_r[S'_r[i_r] + S'_r[j'_r]]]$
$\iff [2a = c' + b \wedge 2a \notin \mathsf{Diff}_4] \bigvee$
$[(2a, c' + b) = (i_r, \mathsf{df}_0[2]), (j'_r, i_r), (\mathsf{df}_0[2], i_r)]$,
where the former condition of S_r corresponds to $[2a = c' + b \wedge 2a \notin \mathsf{Diff}_4]$ and the latter corresponds to $[(2a, c' + b) = (i_r, \mathsf{df}_0[2]), (j'_r, i_r), (\mathsf{df}_0[2], i_r)]$. Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{N-3}{N(N-2)} + \frac{3}{N(N-1)}$. The same reasoning holds in the case of $\Delta\mathsf{State}[4']$.
In $\Delta\mathsf{State}[5] : \Delta Z = 0$
$\iff [[\Delta(S_r[i_r] + S_r[j_r]) = 0] \wedge [S_r[i_r] + S_r[j_r] \notin \mathsf{Diff}_5]]$
$\bigvee [[\Delta(S_r[i_r] + S_r[j_r]) \neq 0] \wedge$
$[S_r[i_r] + S_r[j_r], S'_r[i_r] + S'_r[j'_r] \in \mathsf{Diff}_5] \wedge$
$S_r[S_r[i_r] + S_r[j_r]] = S'_r[S'_r[i_r] + S'_r[j'_r]]]$
$\iff [c + a = c' + b \wedge c + a \notin \mathsf{Diff}_5] \bigvee$
$[(a + c, b + c') = (\mathsf{df}_0[1], j_r), (j_r, \mathsf{df}_0[2]),$
$(\mathsf{df}_0[2], j'_r), (j'_r, \mathsf{df}_0[1])]$,
where the former condition of S_r corresponds to $[c + a = c' + b \wedge c + a \notin \mathsf{Diff}_5]$ and the latter corresponds to $[(a + c, b + c') =$
$(\mathsf{df}_0[1], j_r), (j_r, \mathsf{df}_0[2]), (\mathsf{df}_0[2], j'_r), (j'_r, \mathsf{df}_0[1])]$. Thus, $\mathsf{Prob}\,[\Delta Z = 0] = \frac{N-4}{N(N-3)} + \frac{4}{N(N-1)}$.

In $\Delta\mathsf{State}[6]$: $\mathsf{Prob}\,[\Delta Z = 0] = 0$.
Because $\Delta(S_r[i_r] + S_r[j_r]) \neq 0$ and $\Delta S_r = 0$.

From the above, the proposition follows. ∎

The probability $\mathsf{Prob}\,[\Delta Z = 0]$ when $i = \mathsf{df}_0[1]$ follows immediately from Theorem 4 and Proposition 2.

Corollary 3: Assume that two permutations S and S' in the $(r-1)$-th round are in $\Delta\mathsf{State}[0]$ and $\mathsf{Event}[2]$ occurs in the r-th round. Then, the probability that both outputs are equal in the r-th round, $\mathsf{Prob}\,[\Delta Z = 0]$, is given as follows:

$$\begin{aligned}&\mathsf{Prob}\,[\Delta Z = 0]\\ =\;&\mathsf{Prob}\,[\mathsf{Event}[2]] \cdot \left(\tfrac{N^2-4N+2}{N^2(N-1)} + \tfrac{2(2N-1)(N-2)}{N^2(N-1)^2}\right)\end{aligned}$$

By using Corollaries 2 and 3 and $\mathsf{Prob}\,[\mathsf{Event}[3]]$, the probability $\mathsf{Prob}\,[\Delta Z = 0]$ in the round $r = \mathsf{df}_0[1]$ is computed as follows.

Theorem 5: Assume that two initial permutations S and S' with
$\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$ are given. Then, the probability $P_1 = \mathsf{Prob}\,[\Delta Z = 0]$ in the round $r = \mathsf{df}_0[1]$ is given as

$$\begin{aligned}P_1 =& P_2 \cdot \left((\tfrac{N-2}{N})^2 + \tfrac{4}{N^2(N-1)}\right)\\ &+(1-P_2)\cdot\left(\tfrac{N^2-4N+2}{N^2(N-1)} + \tfrac{2(2N-1)(N-2)}{N^2(N-1)^2}\right),\\ =& P_2 \cdot \left((\tfrac{N-2}{N})^2 - \tfrac{N^2-4N-2}{N^2(N-1)} - \tfrac{2(2N-1)(N-2)}{N^2(N-1)^2}\right)\\ &+\tfrac{N^2-4N+2}{N^2(N-1)} + \tfrac{2(2N-1)(N-2)}{N^2(N-1)^2},\end{aligned}$$

where $P_2 = \mathsf{Prob}\,[\mathsf{Event}[3]]$.

proof: The state of differences between two permutations has the Markov property. Therefore, the probability $\mathsf{Prob}\,[\Delta Z = 0]$ in $r = \mathsf{df}_0[1]$ is determined only by the state in the r-th round, where either $\mathsf{Event}[2]$ or $\mathsf{Event}[3]$ occurs. Theorem 5 follows from this fact. ∎

Remarks 1: 1. The second term of

$$\left(\frac{N^2-4N+2}{N^2(N-1)} + \frac{2(2N-1)(N-2)}{N^2(N-1)^2}\right)$$

can be dealt with as an error term if $\mathsf{df}_0[1]$ is large, which will be discussed in Sect. 5.
2. $P_2 = \mathsf{Prob}\,[\mathsf{Event}[3]]$ can be computed explicitly by Theorems 2 and 3. Thus, $P_1 = \mathsf{Prob}\,[\Delta Z = 0]$ in the round $r = \mathsf{df}_0[1]$ can be explicitly estimated for each $\mathsf{df}_0[1]$.

5. Experimental Results and New Bias

This section shows experimental results of Theorems 2, 3, and 5, and Corollary 2 in Sects. 3 and 4. All experiments were conducted under the following conditions: execute KSA of RC4 with $N = 256$ for 10^8 randomly chosen keys of 16 bytes, generate the initial permutation S_0, and set another initial permutation S'_0 with Diff_0. Experiments are executed over the following sets of Diff_0: $\mathsf{df}_0[1] = 2, \cdots, 255$[†] for Theorems 2 and 3; and $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\} = \{2-254, 255\}$, $\{2, 3-255\}$, and $\{3, 4-255\}$ for Theorem 5 and Corollary 2. The percentage absolute error ϵ of experimental results compared with theoretical results is computed by

$$\epsilon = \frac{|\text{ experimental value} - \text{theoretical value }|}{\text{experimental value}} \times 100(\%),$$

which is also used in [14].

5.1 Experimental Results of Event[3]

Figure 8 shows experimental results of $\mathsf{Prob}\,[\mathsf{Event}[3]]$ and its associated percentage absolute error, where the theoretical value is computed by Theorems 2 and 3. The horizontal axis represents $\mathsf{df}_0[1] = 2, \cdots, 255$. The left side of vertical axis represents $\mathsf{Prob}\,[\mathsf{Event}[3]]$, and the right side represents the percentage absolute error. Table 1 shows the cases of $\mathsf{df}_0[1] \leq 6$ in detail.

Only the cases of $2 \leq \mathsf{df}_0[1] \leq 6$ give the percentage absolute error $\epsilon \geq 5$, and, thus, our theoretical formulae closely match the experimental results if $\mathsf{df}_0[1] > 6$. The

[†]$\mathsf{Event}[3]$ does not depend on $\mathsf{df}_0[2]$. See Theorems 2 and 3.

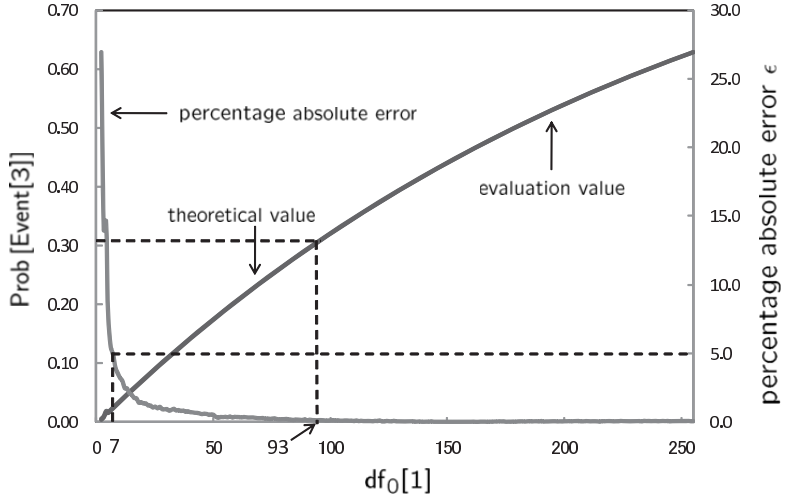

Fig. 8 Experimental results and ϵ of $\mathsf{Event}[3]$.

initial permutation S_0, that is the output of KSA, has a great influence on Event[3] when $\mathsf{df}_0[1]$ is small. Our results indicate that the bias in S_0 is propagated to Prob [Event[3]] as the bias in S_0 has been reported in [7], [11], [14].

Figure 8 also indicates that the nonzero bit difference in the position $\mathsf{df}_0[1]$ moves to another position until $i = \mathsf{df}_0[1]$ with Prob [Event[3]] $> 30\%$ when $\mathsf{df}_0[1] \geq 93$. In this case, the correlations between S and S' such as $\Delta j = 0$ and $|\Delta S| = 2$ remain the same until $i = \mathsf{df}_0[2]$.

5.2 Experimental Results of Outputs

Figure 9 shows experimental results of Prob $[\Delta Z = 0]$ in $r = \mathsf{df}_0[1] - 1$, $\mathsf{df}_0[1]$, and $\mathsf{df}_0[1] + 1$, and percentage absolute error in $r = \mathsf{df}_0[1]$ (i.e. $i = \mathsf{df}_0[1]$), where the theoretical value is computed by Theorem 5. The horizontal axis represents $\mathsf{df}_0[1] = 2, \cdots, 253$. The left side of vertical axis represents Prob $[\Delta Z = 0]$, and the right side represents the percentage absolute error. By using the experimental results, we investigate each case of outputs before or on the nonzero-bit difference.

Outputs before the nonzero-bit difference:
Let us discuss Prob $[\Delta Z = 0]$ in $r = \mathsf{df}_0[1] - 1$ (i.e. $i = \mathsf{df}_0[1] - 1$) for $\mathsf{df}_0[1] = 2, \cdots, 254$. The probability is theoretically estimated in Corollary 2. Our theoretical and experimental results indicate that both outputs of two permutations are coincident with a high probability Prob $[\Delta Z = 0] > 0.98$ during $i < \mathsf{df}_0[1]$[†].

Let us discuss[††] Prob $[\Delta Z = 0]$ in $r = \mathsf{df}_0[1] + 1$ for $\mathsf{df}_0[1] = 2, \cdots, 253$, where $\mathsf{df}_0[1] + 1 = i < \mathsf{df}_0[2]$. Actually, it corresponds to the case in which i is before the nonzero bit difference $\mathsf{df}_0[2]$ since $\mathsf{df}_0[1] + 1$ is an index of nonzero bit difference when $i = \mathsf{df}_0[1] + 1$ from the fact of $\mathsf{df}_0[1] + 1 < \mathsf{df}_0[2]$.

Our experimental results show that Prob $[\Delta Z = 0]$ in the round $\mathsf{df}_0[1]+1$ is almost the same as in the round $\mathsf{df}_0[1]$, which reflects the results in Theorem 1. To sum up, we see that it is highly probable that both outputs of permutations are coincident as long as i does not indicate the index of nonzero bit difference in the current round.

Outputs on the nonzero-bit difference:
Let us discuss Prob $[\Delta Z = 0]$ in $r = \mathsf{df}_0[1]$, where there exists originally[†††] a nonzero-bit difference. Prob $[\Delta Z = 0]$ is estimated theoretically in Theorem 5. From the fact that the percentage absolute error $\epsilon < 1$ holds in $2 \leq \forall \mathsf{df}_0[1] \leq 254$, we see that our theoretical formulae closely match the experimental results in any Diff_0.

Let us discuss the relation between two events of $\Delta Z = 0$ and Event[3] in $r = \mathsf{df}_0[1]$. Figures 8 and 9 show that $\mathsf{df}_0[1]$ satisfying Prob $[\Delta Z = 0] > 30\%$ is almost the same as $\mathsf{df}_0[1]$ satisfying Prob [Event[3]] $> 30\%$. Theorem 5 also indicates that $P_1 =$ Prob $[\Delta Z = 0]$ in the round $\mathsf{df}_0[1]$ deeply affects $P_2 =$ Prob [Event[3]]. Here, we compare the estimation of P_2 by using P_1 with that of P_2 by using the theoretical probability of P_2 in Theorems 2 and 3. Figure 10 shows the comparison between P_1 and P_2 for $2 \leq \mathsf{df}_0[1] \leq 255$, where two percentage absolute errors are listed, $\epsilon_1 = \frac{|P_2 - P_1|}{P_2}$ and $\epsilon_2 = \frac{|P_2 - (theoretical)\mathsf{Prob}[\mathsf{Event}[3]]|}{P_2}$ for experimental values P_1 and P_2. The horizontal axis represents $\mathsf{df}_0[1] = 2, \cdots, 254$. The left side of vertical axis represents Prob $[\Delta Z = 0]$, and

Table 1 Experimental results with $\epsilon \geq 5$ of Event[3].

$\mathsf{df}_0[1]$	Theoretical value	Experimental value	$\epsilon(\%)$
2	0.003906	0.005350	26.991
3	0.007797	0.009069	14.027
4	0.015548	0.018221	14.667
5	0.015534	0.016751	7.265
6	0.019379	0.020501	5.472

[†]Similar experimental results to $i = \mathsf{df}_0[1] - 1$ hold during $i < \mathsf{df}_0[1] - 1$.
[††]The case of $\mathsf{df}_0[1] = 254$ is omitted since i indicates the second nonzero bit difference $\mathsf{df}_0[2] = 255$.
[†††]If Event[3] has occurred in the round $r < \mathsf{df}_0[1]$, then $\mathsf{df}_0[1]$ is not an index of nonzero bit difference.

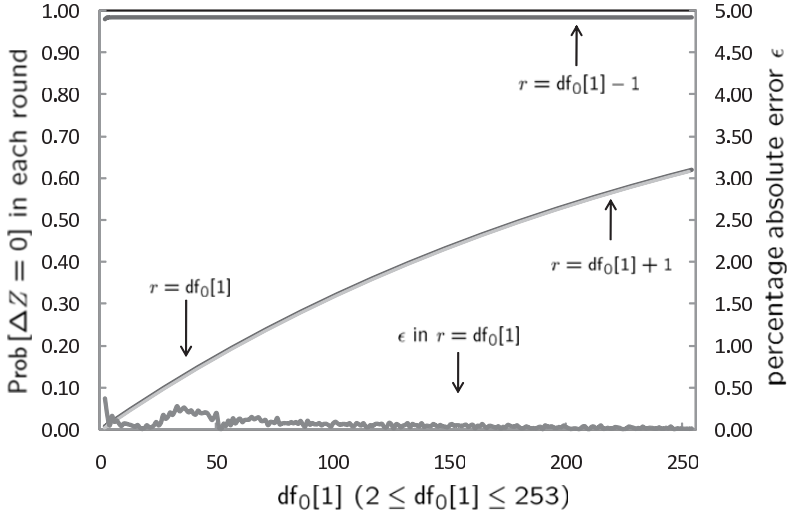

Fig. 9 Prob $[\Delta Z = 0]$ ($\mathsf{df}_0[2] = 255$).

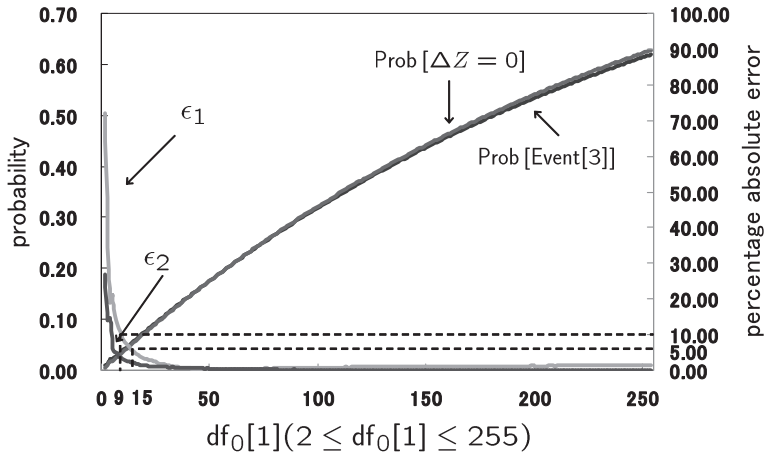

Fig. 10 Comparison of Prob [Event[3]] and Prob [$\Delta Z = 0$].

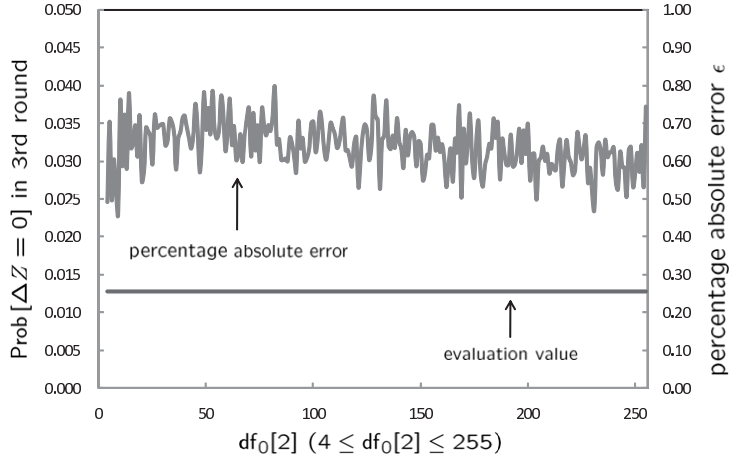

Fig. 11 Prob [$\Delta Z = 0$] ($\mathsf{df}_0[1] = 3$).

the right side represents the percentage absolute error. Experimental results show that $\epsilon_1 < 5$ (resp. 10) if $\mathsf{df}_0[1] > 15$ (resp. $\mathsf{df}_0[1] > 9$). From these theoretical and experimental results, we see that the observable event $\Delta Z = 0$ can indicate that the internal event Event[3] occurs with extremely high probability.

Figure 11 shows experimental results of Prob [$\Delta Z = 0$] in the round $\mathsf{df}_0[1] = 3$ in each case of $4 \le \mathsf{df}_0[2] \le 255$ ($\mathsf{df}_0[1] = 3$), and percentage absolute error. The horizontal axis represents $\mathsf{df}_0[2]$. The left side of vertical axis represents Prob [$\Delta Z = 0$], and the right side represents the percentage absolute error. The percentage absolute error $\epsilon < 0.8$ holds in $4 \le \forall \mathsf{df}_0[2] \le 255$. We see that our theoretical formulae closely match the experimental results independent of another nonzero-bit difference $\mathsf{df}_0[2]$.

5.3 Experimental Results of Biases in $S_0[1]$ and $S_0[2]$

Let us discuss Event[3] when $\mathsf{df}_0[1] = 3$ in detail, where the error $\epsilon > 10$ (Table 1). Theorem 3 says that both $S_0[1]$ and $S_0[2]$ determine Event[3], that is, Event[3] $\Longleftrightarrow [S_0[1] = 3] \bigvee [S_0[1] \neq 2, 3 \wedge S_0[1] + S_0[2] = 3]$. Here we investigate the bias in $S_0[1]$ and $S_0[2]$ from the point of view of Event[3].

Figure 12 shows experimental results concerning the occurrence of $S_0[1]$ with $0 \le S_0[1] \le 255$, and the percentage absolute error, where the theoretical value (a random association) of occurrence of each $S_0[1]$ is $\frac{1}{N} = 3.906 \times 10^{-3}$. Figure 13 shows experimental results concerning the occurrence of $S_0[2]$ when $S_0[1] = 3$, and the percentage absolute error, where the theoretical value (a random association) of occurrence of each $(S_0[1] = 3, S_0[2])$ is $\frac{1}{N(N-1)} = 1.532 \times 10^{-5}$. The horizontal axis represents $S_0[1]$ or $S_0[2]$. The left side of vertical axis represents each probability, and the right side represents each percentage absolute error.

These experimental results indicate a non-uniform distribution of $S_0[1]$ and $S_0[2]$ when $S_0[1] = 3$. Tables 2 and 3 show some cases that indicate a non-uniform distribution as follows:

$$\begin{aligned}
&\text{Prob}\,[S_0[1] = 3] = 5.303 \times 10^{-3} > 3.906 \times 10^{-3},\\
&\text{Prob}\,[S_0[1] = 3 \wedge S_0[2] = x] > 2.0 \times 10^{-5}\\
&\qquad\qquad > 1.532 \times 10^{-5} \text{ for } \forall x \le 135,\\
&\text{Prob}\,[S_0[1] = 3 \wedge 0 \le S_0[2] \le 128] = 3.05299 \times 10^{-3}
\end{aligned}$$

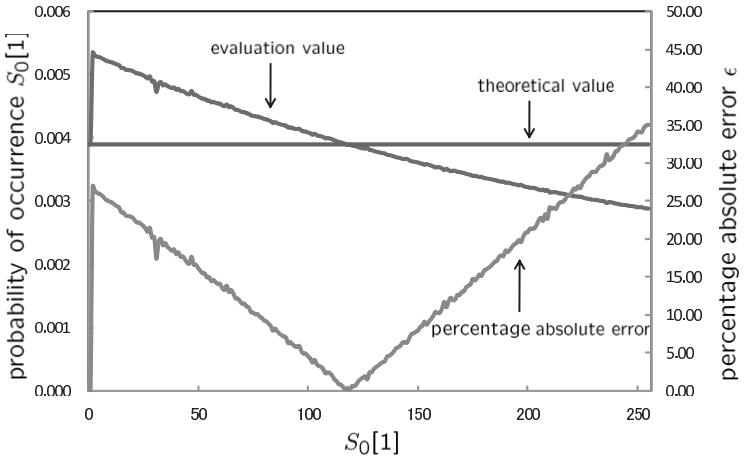

Fig. 12 Occurrence of $S_0[1]$.

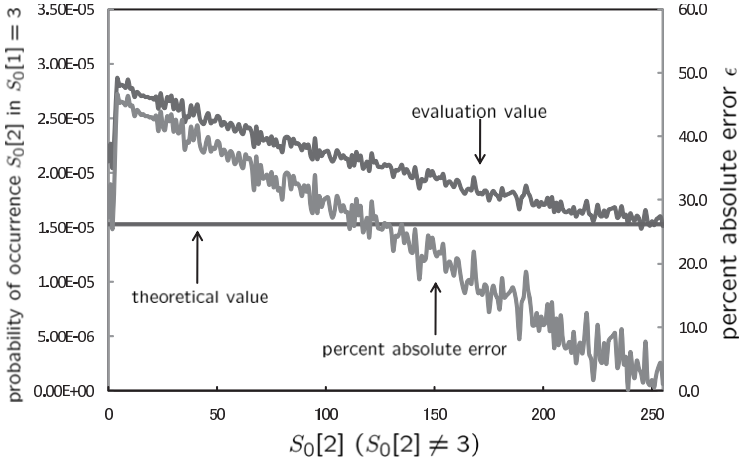

Fig. 13 Occurrence of $S_0[2]$ when $S_0[1] = 3$.

Table 2 Probability of occurrence $S_0[1]$.

$S_0[1]$	Probability of occurrence $S_0[1]$									
0–9	0.0039	0.0039	0.0054	0.0053	0.0053	0.0053	0.0053	0.0052	0.0052	0.0052
10–19	0.0052	0.0052	0.0052	0.0052	0.0052	0.0051	0.0051	0.0051	0.0051	0.0051
20–29	0.0051	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050	0.0049	0.0050	0.0049
30–39	0.0049	0.0047	0.0049	0.0049	0.0048	0.0049	0.0048	0.0048	0.0048	0.0048

Table 3 Probability of occurrence $S_0[2]$ in $S_0[1] = 3$.

$S_0[2]$	Probability of occurrence $S_0[2]$ in $S_0[1] = 3$						
0–6	0.0000211	0.0000227	0.0000207	–	0.0000286	0.0000280	0.0000281
7–13	0.0000280	0.0000278	0.0000286	0.0000277	0.0000278	0.0000270	0.0000274
14–20	0.0000273	0.0000270	0.0000271	0.0000270	0.0000270	0.0000269	0.0000269
108–114	0.0000216	0.0000213	0.0000213	0.0000206	0.0000216	0.0000207	0.0000219
115–121	0.0000212	0.0000216	0.0000204	0.0000207	0.0000210	0.0000202	0.0000218
122–128	0.0000210	0.0000211	0.0000206	0.0000206	0.0000205	0.0000208	0.0000206

$$> 1.9531 \times 10^{-3}.$$

These non-uniform distribution will be used for a new cryptanalytic analysis in Sect. 6.

6. A New Cryptanalytic Analysis

Here we investigate how to analyze the internal state of S or j. Assume that two permutations S and S' with $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$ in the initial round are given, and that both

outputs of PRGA are observable.

Then, by observing both outputs Z and Z' of PRGA, we can recognize the index of the first nonzero-bit difference from the first round in which both outputs are not equal. This is investigated in Sect. 5.2. Therefore, if neither $\mathsf{df}_0[1]$ nor $\mathsf{df}_0[2]$ are known, the first nonzero-bit difference is probabilistically predictable.

Consider the case of $\mathsf{df}_0[1] = 2$. By checking whether $\Delta Z = 0$ in the 2nd round, we can recognize whether $\mathsf{Event}[3]$ has occurred. If $\mathsf{Event}[3]$ has occurred, then $S_0[1] = 2$ holds from Theorem 3. The experimental result shows $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 2] = 0.005350$ (see Table 1). However, if we try to predict $S_0[1]$ from a random association, then the probability is $1/256 = 0.003906$. Therefore, one can guess $S_0[1]$ with an additional advantage of $\frac{0.005350-0.003906}{0.003906} \times 100 = 36.9\ \%$.

Consider the case of $\mathsf{df}_0[1] = 3$. By checking whether $\Delta Z = 0$ in the 3rd round, we can recognize whether $\mathsf{Event}[3]$ has occurred. Let us discuss how to predict both $S_0[1]$ and $S_0[2]$. If $\mathsf{Event}[3]$ has occurred, then $[S_0[1] = 3] \vee [S_0[1] \neq 2, 3 \wedge S_0[1] + S_0[2] = 3]$ holds, from Theorem 3. In the case of $S_0[1] = 3$, the experimental results show that $\mathsf{Prob}\,[\mathsf{Event}[3] \mid \mathsf{df}_0[1] = 3] = 0.009069$ (see Table 1) and $\mathsf{Prob}\,[S_0[1] = 3] = 0.0053$ (see Table 2). On the other hand, we predict $S_0[2]$ with the probability $1/255$. Therefore, we can predict $(S_0[1], S_0[2])$ with the probability $0.0053 \times 1/255 = 2.078431 \times 10^{-5}$. In the case of $[S_0[1] \neq 2, 3 \wedge S_0[1] + S_0[2] = 3]$, if $S_0[1]$ is predicted, then $S_0[2]$ can be predicted promptly. We find that $\mathsf{Prob}\,[\mathsf{Event}[3] \wedge [S_0[1] \neq 2, 3] \wedge [S_0[1] + S_0[2] = 3]] = (0.009069 - 0.0053) \times 1/254 = 1.483858 \times 10^{-5}$. Therefore, we can predict $(S_0[1], S_0[2])$ with the probability 1.483858×10^{-5}. Taking both together, the probability to predict $(S_0[1], S_0[2])$ is $2.078431 \times 10^{-5} + 1.483858 \times 10^{-5} = 3.562289 \times 10^{-5}$. On the other hand, if we try to predict $(S_0[1], S_0[2])$ from a random association, then the probability is $1/256 \times 1/255 = 1.531863 \times 10^{-5}$. Therefore, one can guess $(S_0[1], S_0[2])$ with an additional advantage of $\frac{3.562289-1.531863}{1.531863} \times 100 = 132.54\ \%$.

Further Discussion

Here we discuss how we generalize our analysis to RC4. In this paper, we start with $\mathsf{Diff}_0 = \{\mathsf{df}_0[1], \mathsf{df}_0[2]\}$. However, our results can be generalized to cases where there exist differences Diff_0 with $\#\mathsf{Diff}_0 > 2$ in the initial round. Then, we could apply our analysis to any given two S and S' as follow. Set the first index, whose values of both S and S' differ each other, to $\mathsf{df}_0[1]$. That is, the following holds: $S_0[i] = S'_0[i]$ for $(0 \leq i < \mathsf{df}_0[1])$ and $S_0[\mathsf{df}_0[1]] \neq S'_0[\mathsf{df}_0[1]]$. Then, by applying our discussion to the above case, we could compute the probability that both outputs are equal to each other in $r < \mathsf{df}_0[1]$ theoretically. Furthermore, by observing two outputs, we could predict inner states whether $\mathsf{Event}[3]$ has occurred or not. Then, in the special case of a small $\mathsf{df}_0[1]$, we could guess inner states such as S and j.

7. Conclusion

In this paper, we have investigated, for the first time, correlations between two permutations, S and S', with some differences in the initial round. We have shown that correlations between two permutations S and S' remain before "i" is in the position where the nonzero-bit difference exists in the initial round, and that the correlations remain with non negligible probability even after "i" passed by the position. All theoretical results have been confirmed experimentally.

Our results imply that the same correlations between two permutations will be observed with non negligible probability after the 255-th round. This reveals a new inherent weakness of shuffle-exchange-type PRGA. We have also investigated how to predict inner states such as S and j by using observable two outputs Z and shown its additional advantage compared with prediction from a random association.

Acknowledgments

The authors express our gratitude to anonymous referees for invaluable comments.

References

[1] E. Biham and Y. Carmeli, "Efficient reconstruction of RC4 keys from internal states," Proc. FSE 2008, LNCS, vol.5086, pp.270–188, Springer-Verlag, 2008.
[2] S.R. Fluhrer, I. Mantin, and A. Shamir, "Weaknesses in the key scheduling algorithm of RC4," Proc. 8th Workshop on Selected Areas in Cryptography, LNCS, vol.2259, pp.1–24, Springer-Verlag, 2001.
[3] S.R. Fluhrer and D.A. McGrew, "Statistical analysis of the alleged RC4 keystream generator," Proc. FSE 2001, LNCS, vol.1978, pp.19–30, Springer-Verlag, 2001.
[4] J. Golic, "Linear statistical weakness of alleged RC4 keystream generator," Proc. EUROCRYPT 1997, LNCS, vol.1233, pp.226–238, Springer-Verlag, 1997.
[5] L.R. Knudsen, W. Meier, B. Preneel, V. Rijmen, and S. Verdoolaege, "Analysis methods for (alleged) RC4," Proc. ASIACRYPT 1998, LNCS, vol.3494, pp.327–341, Springer-Verlag, 1998.
[6] A. Maximov and D. Khovratovich, "Analysis of the stream cipher RC4," Proc. CRYPTO 2008, LNCS, vol.5157, pp.297–316, Springer-Verlag, 2008.
[7] I. Mantin, "Analysis of the stream cipher RC4," Master's Thesis, The Weizmann Institute of Science, Israel, 2001.
[8] I. Mantin and A. Shamir, "A practical attack on broadcast RC4," Proc. FSE 2001, LNCS, vol.2355, pp.87–104, Springer-Verlag, 2002.
[9] I. Mantin, "Predicting and distinguishing attacks on RC4 keystream generator," Proc. EUROCRYPT 2005, LNCS, vol.3494, pp.491–506, Springer-Verlag, 2005.
[10] M. Matsui, "Key collisions of the RC4 stream cipher," Proc. FSE 2009, LNCS, vol.5665, pp.38–50, Springer-Verlag, 2009.
[11] I. Mironov, "(Not so) random shuffles of RC4," Proc. CRYPTO 2002, LNCS, vol.2442, pp.304–319, Springer-Verlag, 2002.
[12] S. Mister and S.E. Tavares, "Cryptanalysis of RC4-like ciphers," Proc. SAC 1998, LNCS, vol.1556, pp.131–143, Springer-Verlag, 1998.
[13] G. Paul and S. Maitra, "RC4 state information at any stage reveals

the secret key," Cryptology ePrint Archive, Report 2007/208, Available from: http://eprint.iacr.org/2007/208.pdf
[14] G. Paul, S. Maitra, and R. Srivastava, "On non-randomness of the permutation after RC4 key scheduling," Posted to AAECC2007, Available from: http://eprint.iacr.org/2007/305.pdf
[15] S. Paul and B. Preneel, "A new weakness in the RC4 keystream generator and an approach to improve the security of the cipher," Proc. FSE2004, LNCS, vol.3017, pp.245–259, Springer-Verlag, 2004.
[16] G. Paul, S. Rathi, and S. Maitra, "On non-negligible bias of the first output byte of RC4 towards the first three bytes of the secret key," Des. Codes and Cryptogr., vol.49, pp.123–134, Springer-Verlag, 2008.
[17] V. Tomasevic and S. Bojanic, "Reducing the state space of RC4 stream cipher," Proc. ICCS2004, LNCS, vol.3036, pp.644–647, Springer-Verlag, 2004.

Atsuko Miyaji received the B.Sc., the M.Sc., and the Dr.Sci. degrees in mathematics from Osaka University, Osaka, Japan in 1988, 1990, and 1997 respectively. She joined Panasonic Co., LTD from 1990 to 1998 and engaged in research and development for secure communication. She was an associate professor at the Japan Advanced Institute of Science and Technology (JAIST) in 1998. She has joined the computer science department of the University of California, Davis since 2002. She has been a professor at the Japan Advanced Institute of Science and Technology (JAIST) since 2007 and the director of Library of JAIST since 2008. Her research interests include the application of number theory into cryptography and information security. She received the IPSJ Sakai Special Researcher Award in 2002, the Standardization Contribution Award in 2003, Engineering Sciences Society: Certificate of Appreciation in 2005, the AWARD for the contribution to CULTURE of SECURITY in 2007, IPSJ/ITSCJ Project Editor Award in 2007, 2008, and 2009, the Director-General of Industrial Science and Technology Policy and Environment Bureau Award in 2007, Editorial Committee of Engineering Sciences Society: Certificate of Appreciation in 2007, and DoCoMo Mobile Science Awards in 2008. She is a member of the International Association for Cryptologic Research, the Institute of Electronics, Information and Communication Engineers, the Information Processing Society of Japan, and the Mathematical Society of Japan.

Masahiro Sukegawa received the B.Eng. degree from the Tokyo University of Science in 2006 and the M. Info. Sci. degree from the Japan Advanced Institute of Science and Technology in 2009.