Analysis of Burn-in period for RC4 State Transition

Goutam Paul and Souvik Ray*

Indian Statistical Institute, Kolkata 700 108, India.
goutam.paul@isical.ac.in, souvikr974@gmail.com

Abstract. The internal state of RC4 stream cipher is a permutation over $\mathbb{Z}_N$ and its state transition is effectively a transposition or swapping of two elements. How the randomness of RC4 state evolves due to its state transitions has been studied for many years. As the number of swaps increases, the state comes closer to a uniform random permutation. We call the burn-in period of RC4 state transition as the number of swaps required to make the state very close to uniform random permutation under some suitably defined distance measure. Earlier, Mantin in his Master's thesis (2001) has performed an approximate analysis of the burn-in period. In this paper, we perform a rigorous analysis of the burn-in period and in the process derive the exact distribution of the RC4 state elements at any stage.

Keywords: Bias, Burn-in, Cryptography, Random Permutation, RC4, State transition, Stream cipher.

1 Introduction

RC4, since its inception in 1987, has been the most popular software stream cipher for commercial use until recently. The internet security protocols like SSL, SSH, TLS, WEP, WPA have extensively used RC4. Due to its simple structure, it has also invited a lot of cryptanalytic efforts [21, 27, 9, 7, 5, 4, 14, 15, 12, 13, 15, 17, 19, 25, 26, 10, 3, 23, 24, 22, 2, 8, 6]. Due to the recent attacks on TLS [2, 6], the Internet Engineering Task Force (IETF) has deprecated its use in TLS 1.0 and it has been removed from the TLS 2.0 draft under preparation.

Interestingly, the usability of the *RC4-like* ciphers is re-iterated in the recent proposal of Spritz [20] from the authors of RC4. The original RC4 and its variants belong to the shuffle-exchange paradigm and the state evolutions of these ciphers have several interesting combinatorial results [15]. Thus, even if RC4 is replaced by other stream ciphers in practical protocols, RC4 and its variants, with their elegant and robust structures and nice combinatorial properties, are likely to remain model stream ciphers for both designers and cryptanalysts of the future.

* The second author worked for this paper during the winter break in 2016 in his Master of Statistics course.

The internal state of RC4 consists of an array S of size $N = 256$, which contains a permutation of $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$. A secret key k consists of l (typically 5 to 32) elements from $\mathbb{Z}_N$ and is stretched to an array K of size N is such that $K[i] = k[i \bmod l]$, $0 \leq i \leq N-1$.

Like any stream cipher, RC4 algorithm has two components: the Key Scheduling Algorithm (KSA) and the Pseudo Random Generation Algorithm (PRGA). The KSA initializes S to an identity permutation over $\mathbb{Z}_N$ and uses the secret key K to scramble S by N transpositions or swaps. The PRGA uses this scrambled S to produce one keystream output $\in \mathbb{Z}_N$ per iteration and updates S by further swaps. The algorithms are described below.

Algorithm KSA

Initialization:

For $i \leftarrow 0, \ldots, N-1$

$S[i] \leftarrow i$;

$j \leftarrow 0$;

Scrambling:

For $i \leftarrow 0, \ldots, N-1$

$j \leftarrow (j + S[i] + K[i]) \bmod N$;

Swap$(S[i], S[j])$;

Algorithm PRGA

Initialization:

$i \leftarrow 0, j \leftarrow 0$;

Output Keystream Generation Loop:

$i \leftarrow (i+1) \bmod N$;

$j \leftarrow (j + S[i]) \bmod N$;

Swap$(S[i], S[j])$;

$t \leftarrow (S[i] + S[j]) \bmod N$;

Output $z \leftarrow S[t]$;

One direction of analysis of RC4 looks into the randomness of the state S. Randomness of S is very important, because that in turn results in randomness of the keystream generated from S. In the actual RC4 KSA, the resulting state has lots of non-randomness that causes non-randomness in initial few hundred keystream outputs of the PRGA. In practice, these initial outputs are discarded to ensure randomness of the keystream. The *burn-in period* for RC4 state transition may be defined as the number of swaps needed to make the state ϵ-close to a random permutation under some distance measure for a negligibly small ϵ.

Let S_r denote the permutation after r many swaps in the KSA. That is, S_0 is the initial identity permutation at the beginning of the KSA and S_N is the permutation after the KSA is over. Similarly, let S_r^G denote the permutation after r many swaps in the PRGA. In [11, Chapter 6 and Appendix C] and later in [15], the problem of estimating $\Pr(S_N[u] = v)$, $0 \leq u \leq N-1, 0 \leq v \leq N-1$, i.e., the distribution of the values in each permutation position after the KSA has been discussed. A long derivation of these results has been presented in Mantin's thesis [11, Chapter 6 and Appendix C]. Later, Mironov [15] provided a shorter proof. However, Mironov's argument [15] does not consider the scenario of arbitrary initial distribution of S, which we have investigated here. Further, we rigorously investigate the evolution of the randomness of S as the number of swaps increases, and derive the general distribution of $S_r[u]$ and $S_r^G[u]$ for each position $u \in \mathbb{Z}_N$. This helps us to perform a tighter analysis of the *burn-in period* for RC4 state transition, that has been bounded via a loose approximation in Mantin's thesis [11, Section 6.3.4].

2 Solution for arbitrary initial state distribution

We consider an idealized model for the KSA and PRGA algorithm which is seemingly consistent with the intent of the cipher designer. All the operations made on the index j in those two algorithms are aimed at making that index random. So, the idealized model assigns to j a random index from $\{0, \ldots, N-1\}$ whenever j gets changed. This idealized models, denoted by KSA* and PRGA*, are described below.

Algorithm KSA*

Initialization:
    For $i \leftarrow 0, \ldots, N-1$
        $S[i] \leftarrow i$;
Scrambling:
    For $i \leftarrow 0, \ldots, N-1$
        $j \leftarrow Uniform\{0, \ldots, N-1\}$;
        Swap$(S[i], S[j])$;

Algorithm PRGA*

Initialization:
    $i \leftarrow 0, j \leftarrow 0$;
Output Keystream Generation Loop:
    $i \leftarrow (i+1) \bmod N$;
    $j \leftarrow Uniform\{0, \ldots, N-1\}$;
    Swap$(S[i], S[j])$;
    $t \leftarrow (S[i] + S[j]) \bmod N$;
    Output $z \leftarrow S[t]$;

All the analysis presented in this paper will be based upon this idealized model. So, it has to be kept in mind that, from now on whenever we use the notations introduced in the previous section, they are associated with this idealized model.

Let us denote

$$p_{u,v}^{(t)} := \Pr[S_t[u] = v]$$

where $u, v \in \{0, \ldots, N-1\}$ and $t \geq 0$. Initially the state array S_0 is the array of the numbers $0, \ldots, N-1$ in their usual order. In particular, for the standard RC4,

$$p_{u,v}^{(0)} = \begin{cases} 1, \text{ if } u = v; \\ 0, \text{ otherwise.} \end{cases}$$

According to Algorithm KSA*, at the t-th step during the evolution of the state array (where t lies between 1 and N), the $(t-1)$-th cell is taken and one of the cells of the total N cells is chosen uniformly at random. For simplicity of notation, let us call the index of random cell chosen as the random index and the index of the cell chosen deterministically (e.g. $(t-1)$-th cell at t-th step) as the *deterministic index*. The contents of these two cells are then interchanged. Hence, for $t \geq 1$, the state probabilities are given by the following result.

Proposition 1. *[15]*

$$p_{u,v}^{(t)} = \begin{cases} \dfrac{1}{N}, \ \textit{if } t = u+1; \\ \\ \dfrac{1}{N} p_{t-1,v}^{(t-1)} + \dfrac{N-1}{N} p_{u,v}^{(t-1)}, \ \textit{if } t \neq u+1. \end{cases}$$

Note that Mironov [15] used it for the PRGA* which we shall discuss later. We want to use this recursion to get an expression for the state array probabilities after KSA* completes, i.e. after time $t = N$. We shall pursue this objective by two methods, first one being an algebraic way, i.e. unfolding of the recursion and the later one uses combinatorial arguments.

2.1 Algebraic Derivation of State probabilities for the KSA*

We shall make use of two lemmas as follows.

Lemma 1.

$$p_{u,v}^{(t)} = \sum_{j=0}^{t-1} \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j} p_{j,v}^{(0)} + \Big(\frac{N-1}{N}\Big)^{t} p_{u,v}^{(0)}, \ \textit{if } 0 \le t < u+1.$$

Proof. We shall prove this by induction on t. For $t = 1$ and any $u > 0, v$, it holds true directly from Proposition 1. Now suppose it holds true for all $t = 1, \ldots, k-1$, where $2 \le k \le N-1$. Then by Proposition 1, for any $u > k-1, v$, we have

$$\begin{aligned} p_{u,v}^{(k)} &= \frac{1}{N} p_{k-1,v}^{(k-1)} + \frac{N-1}{N} p_{u,v}^{(k-1)} \\ &= \frac{1}{N}\Big[\sum_{j=0}^{k-2} \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j} p_{j,v}^{(0)} + \Big(\frac{N-1}{N}\Big)^{k-1} p_{k-1,v}^{(0)}\Big] + \frac{N-1}{N}\Big[\sum_{j=0}^{k-2} \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j} p_{j,v}^{(0)} + \Big(\frac{N-1}{N}\Big)^{k-1} p_{u,v}^{(0)}\Big] \\ &= \sum_{j=0}^{k-1} \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j} p_{j,v}^{(0)} + \Big(\frac{N-1}{N}\Big)^{k} p_{u,v}^{(0)} \end{aligned}$$

and hence by induction hypothesis our lemma is proved. □

Lemma 2.

$$p_{u,v}^{(t)} = \sum_{k=0}^{t-1} \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{k} p_{k,v}^{(0)} - \Big[\sum_{k=0}^{u} \Big(\frac{N-1}{N}\Big)^{k} p_{k,v}^{(0)}\Big] \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{t-u-1} + \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{t-u-1}, \ \textit{if } u+1 \le t \le N.$$

Proof. We again prove this by induction on t. It holds true for $t = 1$ as the only u we have to check for is $u = 0$. Now suppose it holds true for $t = 1, \ldots, k-1$ where $2 \le k \le N$. Then by Proposition 1 and Lemma 1 we proved before, we have for $k > u+1$

$$
\begin{aligned}
p_{u,v}^{(k)} &= \frac{1}{N}p_{k-1,v}^{(k-1)} + \frac{N-1}{N}p_{u,v}^{(k-1)} \\
&= \frac{1}{N}\Big[\sum_{j=0}^{k-2}\frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j}p_{j,v}^{(0)} + \Big(\frac{N-1}{N}\Big)^{k-1}p_{k-1,v}^{(0)}\Big] + \\
&\quad \frac{N-1}{N}\Big[\sum_{j=0}^{k-2}\frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j}p_{j,v}^{(0)} - \Big[\sum_{j=0}^{u}\Big(\frac{N-1}{N}\Big)^{j}p_{j,v}^{(0)}\Big]\frac{1}{N}\Big(\frac{N-1}{N}\Big)^{k-u-2} + \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{k-u-2}\Big] \\
&= \sum_{j=0}^{k-1}\frac{1}{N}\Big(\frac{N-1}{N}\Big)^{j}p_{j,v}^{(0)} - \Big[\sum_{j=0}^{u}\Big(\frac{N-1}{N}\Big)^{j}p_{j,v}^{(0)}\Big]\frac{1}{N}\Big(\frac{N-1}{N}\Big)^{k-u-1} + \frac{1}{N}\Big(\frac{N-1}{N}\Big)^{k-u-1}
\end{aligned}
$$

And for $k = u+1$, it is obvious by Proposition 1. Hence, the lemma is proved. □

So, putting $t = N$ in Lemma 2, we get the following result.

Theorem 1.

$$
p_{u,v}^{(N)} = \sum_{j=0}^{N-1} pq^{j}p_{j,v}^{(0)} - \Big[\sum_{j=0}^{u} q^{j}p_{j,v}^{(0)}\Big]pq^{N-u-1} + pq^{N-u-1}, \ \ \forall \ 0 \le u, v \le N-1,
$$

where $Np := 1$ *and* $q := 1-p$.

Notice that, the above expression holds true for arbitrary initial state array distribution. Putting values for the initial state probabilities, which are stated earlier we get,

$$
p_{u,v}^{(N)} = \begin{cases} pq^{v} + pq^{N-u-1}, \text{ if } u < v; \\ pq^{v} + (1-pq^{v})pq^{N-u-1}, \text{ if } u \ge v. \end{cases}
$$

This concludes our algebraic proof.

2.2 Combinatorial Derivation of the State probabilities for the KSA*

The combinatorial proof will be only for $t = N$ and assuming that, initially the state array is a fixed permutation of the numbers $\{0, \ldots, N-1\}$, rather than a probability distribution. Let us denote the initial permutation by $(\beta_0, \ldots, \beta_{N-1})$,i.e. the u-th cell of the state array initially contains the numbers β_u, for $u = 0, \ldots, N-1$. We shall prove the result through a series of lemmas. These lemmas may seem somewhat similar to the lemmas used to prove similar result in [18], but their ultimate result does not match with others. After our proof is finished we shall try to point out some flaws in their analysis which have caused the different result.

Another motivation of our alternative derivation is to get rid of the unnecessary idea of "relative position" used in [11, Chapter 6 and Appendix C]. We give a direct proof which is easy to follow.

Lemma 3. *If the event* $(S_r[k] = v)$ *occurs for some* $0 \le k < r$, *then after* t*-th step,* v *will be in one of the cells among* 0 *to* $t-1$, *for all* $t \ge r$.

Proof. We shall show by induction on t. Clearly holds true for $t = r$. Suppose holds true for $t = r, \ldots, s$, where $r \leq s \leq N-1$. Now at $(s+1)$-th step, if the current position of v remains unchanged, then we are done. Otherwise, notice that the deterministic index at $(s+1)$-th step is s and by induction hypothesis v is not in s-th cell after s-th step. So, the only way the position of v can be changed is by choosing the current cell of v as the random index and then after the $(s+1)$-th step, v will end up in the s-th cell and therefore the induction hypothesis holds true. □

Lemma 4.

$$p^{(v+1)}_{v+1,\beta_v} = pq^v, \forall\ v \geq 0.$$

Proof. Notice that, the deterministic index at the $(v+1)$-th step is v. So, v has not been chosen as a deterministic index up to step v. We consider two situations. In one situation, v has not been chosen as a random index up to step v, and then to get β_v at the $(v+1)$-th cell after $(v+1)$-th step, the random index at the $(v+1)$-th step should be $(v+1)$. This whole event has probability pq^v.

In the other situation, v is chosen as a random index at r-th step for the first time, where $1 \leq r \leq v$. Then the event $(S_r[r-1] = \beta_v)$ occurs. And therefore by Lemma 3, β_v will be in one of the cells among 0 to v after $(v+1)$-th step. Therefore, this situation does not contribute to the event $(S_{v+1}[v+1] = \beta_v)$. Hence, the lemma is proved. □

Lemma 5.

$$p^{(t)}_{u,\beta_v} = pq^v,\ \forall\ 0 \leq v \leq t-1 < u \leq N-1.$$

Proof. Note that up to the t-th step of the evolution of the state array, the deterministic indices are 0 to $t-1$. So, for any index v which is less than t, the probability of β_v being in any of the cells between t and $N-1$ after t-th step is equal, and therefore

$$p^{(t)}_{u,\beta_v} = p^{(t)}_{t,\beta_v},\ \forall\ v \leq t-1 < u.$$

Using the recursion relation of the state probabilities and the above relation, we have for all $v \leq t-2 < t-1 < u$,

$$p^{(t)}_{u,\beta_v} = pp^{(t-1)}_{t-1,\beta_v} + qp^{(t-1)}_{u,\beta_v} = pp^{(t-1)}_{u,\beta_v} + qp^{(t-1)}_{u,\beta_v} = p^{(t-1)}_{u,\beta_v}. \tag{1}$$

Thus, by Lemma 4 and (1)

$$p^{(t)}_{u,\beta_v} = p^{(v+1)}_{u,\beta_v} = p^{(v+1)}_{v+1,\beta_v} = pq^v\ ,\ \forall\ v \leq t-1 < u,$$

which concludes our proof. □

Lemma 6.

$$\Pr[S_N[u] = v | S_t[j] = v] = pq^{j-t},\ \forall\ v \geq 0,\ 0 \leq u < t \leq j,$$

provided $P[S_t[j] = v]$ *is positive.*

Proof. Consider the situation when the event $(S_t[j] = v)$ has occurred. Now we shall do a rearrangement of the state array. Think of the first t cells of the state array, i.e., cells from 0 to $t-1$. Notice that u-th cell is among them. Take this t cells together and put them at the end of the array. The j-th and the u-th cell of the old state array now becomes the $(j-t)$-th and the $(N-t+u)$-th cell of the new state array respectively. The next operations on the old state array making $t, \ldots, N-1$ as deterministic indices can be thought of same kind of operations on the new state array with $0, \ldots, N-t-1$ as deterministic indices. Now, in the new state array v was initially in $(j-t)$-th cell and the event $(S_N[u] = v)$ in the old state array implies that v will go to the $(N-t+u)$-th cell of the new state array after $N-t$ operations, which by Lemma 5 has probability pq^{j-t}. (Since that result is valid for any initial permutations.) This completes the proof of this lemma. □

Lemma 7.

$$\Pr[S_N[u] = \beta_v | S_{u+1}[u] = \beta_v] = q^{N-u-1}, \ \ \forall\, 0 \le u, v \le N-1.$$

Proof. After $(u+1)$-th step, β_v is in u-th cell. Suppose, after this $(u+1)$-th step, u is chosen as the random index for the first time during the r-th step, $r > u+1$. Then the event $(S_r[r-1] = \beta_v)$ has occurred. Finally, after the N-th step β_v ends up in u-th cell, hence consider the last step during which the position of β_v has been changed, let it be t-th step. Hence, the events $(S_{t-1}[u] \neq \beta_v), (S_t[u] = \beta_v), \ldots, (S_N[u] = \beta_v)$ have occurred, and $t > r > u+1$. Therefore, at t-th step, u was not the deterministic index, but the content of u-th cell has been changed in this step which implies u was the random index at t-th step. Therefore, the event $(S_{t-1}[t-1] = \beta_v)$ has occurred. But Lemma 3 tells that, as the event $(S_r[r-1] = \beta_v)$ has occurred, β_v will be in one of the cells among 0 to $t-2$ after $(t-1)$-th step. Therefore, to get β_v at the u-th cell after N-th step, u should not be chosen as the random index in any of the following steps. Hence, $\Pr(S_N[u] = \beta_v | S_{u+1}[u] = \beta_v) = q^{N-u-1}$. □

Lemma 8.

$$\Pr(S_N[u] = \beta_v | S_{u+1}[r] = \beta_v) = 0, \ \ \forall\, 0 \le v \le N-1, 0 \le r < u \le N-1,$$

provided $\Pr[S_{u+1}[r] = \beta_v] > 0$.

Proof. Suppose, after the $(u+1)$-th step, β_v is in the r-th cell, where $0 \le r < u$. Finally, after the N-th step, β_v ends up in the u-th cell; hence consider the last step during which the position of β_v has been changed and let it be the t-th step. Hence, the events $(S_{t-1}[u] \neq \beta_v), (S_t[u] = \beta_v), \ldots, (S_N[u] = \beta_v)$ have occurred, and $t > u+1$. Therefore, at t-th step, u was not the deterministic index, but the content of u-th cell has been changed in this step which implies u was the random index at t-th step. Therefore, the event $(S_{t-1}[t-1] = \beta_v)$ has occurred. But Lemma 3 tells that, as the event $(S_{u+1}[r] = \beta_v)$ has occurred, β_v will be in one of the cells among 0 to $t-2$ after $(t-1)$-th step. This leads to $\Pr[S_N[u] = \beta_v | S_{u+1}[r] = \beta_v] = 0$. □

We are now ready with our equipments and shall complete the proof of the main result in two steps.

Theorem 2.

$$p_{u,\beta_v}^{(N)} = pq^v + pq^{N-u-1}, \ \ \forall\, 0 \le u < v \le N-1.$$

Proof. Consider the cell where β_v is after $(u+1)$-th step. v is not a deterministic index up to $(u+1)$-th step. Notice that, if v is not chosen as a random index up to the $(u+1)$-th step, then β_v is in v-th cell after $(u+1)$-th step. Now, if v chosen as a random index in the r-th step where $1 \le r \le u+1$, then the event $(S_r[r-1] = \beta_v)$ occurs which guarantees by Lemma 3 that β_v will be in one of the cells among 0 to u after $(u+1)$-th step. Therefore, the only possible positions of β_v after $(u+1)$-th step are $0, \ldots, u, v$. This argument also indicate that $\Pr[S_{u+1}[v] = \beta_v] = q^{u+1}$.

Now, Lemma 8 gives that $\Pr(S_N[u] = \beta_v | S_{u+1}[r] = \beta_v) = 0$, for all $0 \le r < u$. Therefore,

$$\begin{aligned}\Pr(S_N[u] = \beta_v) = \Pr(S_N[u] = \beta_v | S_{u+1}[u] = \beta_v) p_{u,\beta_v}^{(u+1)} \\ + \Pr(S_N[u] = \beta_v | S_{u+1}[v] = \beta_v) p_{v,\beta_v}^{(u+1)}.\end{aligned}$$

Lemma 7 gives, $\Pr[S_N[u] = \beta_v | S_{u+1}[u] = \beta_v] = q^{N-u-1}$. By the recursion relation for state array probabilities, $p_{u,\beta_v}^{(u+1)} = p$. Lemma 5 gives, $\Pr(S_N[u] = \beta_v | S_{u+1}[v] = \beta_v) = pq^{v-u-1}$. All these together give

$$\Pr[S_N[u] = \beta_v] = pq^{N-u-1} + pq^v.$$

□

Theorem 3.

$$p_{u,\beta_v}^{(N)} = pq^v(1 - q^{N-u-1}) + pq^{N-u-1}, \ \ \forall\, 0 \le v \le u \le N-1.$$

Proof. Similar to the previous analysis, we consider the cell where β_v is after $(u+1)$-th step. Lemma 8 guarantees that $\Pr(S_N[u] = \beta_v | S_{u+1}[r] = \beta_v) = 0$, for all $0 \le r < u$. Therefore,

$$\Pr[S_N[u] = \beta_v] = \sum_{k=u}^{N-1} \Pr(S_N[u] = \beta_v | S_{u+1}[k] = \beta_v) \Pr[S_{u+1}[k] = \beta_v]. \tag{2}$$

By Lemma 6,$\Pr(S_N[u] = \beta_v | S_{u+1}[k] = \beta_v) = pq^{k-u-1}$, for all $u+1 \le k \le N-1$. Lemma 7 gives, $\Pr[S_N[u] = \beta_v | S_{u+1}[u] = \beta_v] = q^{N-u-1}$. Lemma 5 gives $\Pr[S_{u+1}[k] = \beta_v] = pq^v$, for all $k \ge u+1$. And finally by the recursion relation $\Pr[S_{u+1}[u] = \beta_v] = p$. All these combined give,

$$\begin{aligned}\Pr[S_N[u] = \beta_v] &= pq^v \sum_{k=u+1}^{N-1} pq^{k-u-1} + pq^{N-u-1} \\ &= pq^v(1 - q^{N-u-1}) + pq^{N-u-1}.\end{aligned}$$

□

These two propositions complete our proof. Now, returning to the analysis presented in [18], their argument in the proof of Lemma 2 of that paper needs some correction. The proof argues that the only two ways for the event $(S_{v+1}[u] = v)$ to occur (where $v \ge u+1$) are

1. $(S_v[u] = v)$ has already occurred and the index u is not involved in the swap in $(v+1)$-th step.
2. $(S_v[u] \neq v)$ has occurred and the value v comes into the u-th cell from the v-th cell in the $(v+1)$-th step.

But when calculating the probability, the proof says

$$\Pr[S_v[u] \neq v, S_v[v] = v] = \Pr[S_v[u] \neq v] \Pr[S_v[v] = v],$$

whereas actually

$$\Pr[S_v[u] \neq v, S_v[v] = v] = \Pr[S_v[v] = v],$$

as the events on both sides are equal. This false independence assumption brings an extra negative term in their expression in Lemma 2, and gives false state probabilities.

2.3 Distribution of the state array after the r-th round of PRGA*

We have computed the distribution of the state array after KSA* for arbitrary initial distributions in Lemma 1 and Lemma 2. We now wish to carry out further calculations to obtain the distribution of the state array after the r-th round of the PRGA*. The scrambling process in both KSA* and PRGA* are the same in nature. Each loop goes by taking deterministic indices from 0 to $N-1$, one by one and then choosing a random index followed by interchanging their contents. We shall refer to this type of loop by *the usual loop*. The only difference between KSA* and PRGA* is that, while KSA* performs only one usual loop, PRGA* performs many of them consecutively but the first loop taking 1 to $N-1$ as deterministic indices. Our result in Theorem 1 enables us to calculate the distribution of the state array after one usual loop for any arbitrary initial distribution. So, if we can compute the distribution of the state array after one loop, where the deterministic indices run from 1 to $N-1$ (let us call such loops as *unusual loops*) for arbitrary initial distribution, we should be able to compute the distribution of the state array after any round of PRGA*. We may not be able to find a simple closed form expression as for the KSA*, but we shall try to make the expression as simple as possible.

To proceed in the above mentioned direction, let us first introduce some notations, which we will be handy in all of the coming sections.

1. $\mathbf{1}_N := N \times N$ identity matrix.
2. $\boldsymbol{x} := (p, pq, \ldots, pq^{N-1})^T$.
3. $\boldsymbol{x'} := (p, pq, \ldots, pq^{N-2}, 0)^T$.
4. $\boldsymbol{y}_u := (p, pq, \ldots, pq^u, 0, \ldots, 0)^T, \ \forall\, u \in \mathbb{Z}_N$.
5. $\boldsymbol{\psi} := (pq^{N-1}, \ldots, p)^T$.
6. $\boldsymbol{\psi'} := (pq^{N-2}, \ldots, p, 0)^T$.
7. $\boldsymbol{\zeta} := (0, \ldots, 0, q^{N-1})^T$.

Now, suppose KSA* is performed k-times on the state array, which means the state array has gone through k usual loops. We define

$$\boldsymbol{p}_v^{(k)} := (p_{0,v}^{(Nk)}, \ldots, p_{N-1,v}^{(Nk)})^T,\ k \geq 1;\ v \in \mathbb{Z}_N.$$

Here, $\boldsymbol{p}_v^{(k)}$ is actually the distribution of the element v in the cells of S at time Nk, i.e., after the completion of k-th usual loop. Now using Lemma 2 we have

$$p_{u,v}^{(Nk)} := \boldsymbol{x}^T \boldsymbol{p}_v^{(k-1)} - q^{N-u-1} \boldsymbol{y}_u^T \boldsymbol{p}_v^{(k-1)} + pq^{N-u-1};\ \ k \geq 1. \tag{3}$$

Now define

$$B := \begin{bmatrix} q^{N-1}\boldsymbol{y}_0^T \\ q^{N-2}\boldsymbol{y}_1^T \\ \cdots \\ q^0 \boldsymbol{y}_{N-1}^T \end{bmatrix}.$$

So, then using Equation (3) we can write

$$\boldsymbol{p}_v^{(k)} := \mathbf{1}_N \boldsymbol{x}^T \boldsymbol{p}_v^{(k-1)} - B\boldsymbol{p}_v^{(k-1)} + \boldsymbol{\psi}.$$

And if we denote

$$A := \mathbf{1}_N \boldsymbol{x}^T - B, \tag{4}$$

then our final formula for the evolution of the distribution of the element v becomes

$$\boldsymbol{p}_v^{(k)} := A\boldsymbol{p}_v^{(k-1)} + \boldsymbol{\psi},\ \ \forall\, v \in \mathbb{Z}_N. \tag{5}$$

Note that $A, \boldsymbol{\psi}$ don't depend upon v. The deviation of the distribution of the element v from the uniform distribution can be defined as

$$\boldsymbol{\delta}_v^{(k)} := \boldsymbol{p}_v^{(k)} - p\mathbf{1}_N,\ \ k \geq 1;\ v \in \mathbb{Z}_N.$$

Now, suppose $\boldsymbol{p}_v^{(0)} = p\mathbf{1}_N$, i.e., $p_{k,v}^{(0)} = p,\ \forall\, k \in \mathbb{Z}_N$. Then by Theorem 1 we have

$$\begin{aligned} p_{u,v}^{(N)} &= \sum_{j=0}^{N-1} pq^j p - \Big[\sum_{j=0}^{u} q^j p\Big] pq^{N-u-1} + pq^{N-u-1} \\ &= p^2 \frac{1-q^N}{1-q} - p^2 q^{N-u-1} \frac{1-q^{u+1}}{1-q} + pq^{N-u-1} \\ &= p(1-q^N) - pq^{N-u-1} + pq^N + pq^{N-u-1} = p \end{aligned}$$

Hence,

$$\boldsymbol{p}_v^{(0)} = p\mathbf{1}_N \Rightarrow \boldsymbol{p}_v^{(1)} = p\mathbf{1}_N.$$

And hence by Equation 5, we have

$$p\mathbf{1}_N = pA\mathbf{1}_N + \boldsymbol{\psi}.$$

Combining the above equation with Equation (5) we get the following result.

Theorem 4. *If $\boldsymbol{\delta}_v^{(0)}$ and $\boldsymbol{\delta}_v^{(k)}$ denote the deviation of the distribution of the element v, before and after $k \geq 1$ KSA* loops (i.e., before and after k usual loops) respectively, from the uniform distribution, then*

$$\boldsymbol{\delta}_v^{(k)} = A^k \boldsymbol{\delta}_v^{(0)}, \ \forall \ v \in \mathbb{Z}_N,$$

where A is as defined in (4).

Thus the matrix A acts like a transition matrix for an usual loop.

Let $\left\{p'^{(0)}_{u,v}\right\}_{0\leq u,v\leq N-1}$ be the initial distribution entering an unusual loop, where $p'^{(0)}_{u,v} = \Pr[S_0[u] = v]$. And suppose $\left\{p'^{(N-1)}_{u,v}\right\}_{0\leq u,v\leq N-1}$ be the distribution of the state array after the unusual loop is completed. Now we shall again use the trick of rearranging the state array. Consider the 0-th cell of the state array S, and we put this cell at the end of the array and relabel the cells of the new array from 0 to $N-1$ in usual way. Call the new state array S^*. Therefore, the i-th cell of S' was actually $(i+1)(\mathrm{mod}N)$-th cell of S. Consider the first $(N-1)$ steps of the usual loops performed on the array S^*. Then the array obtained after placing the last cell of S^* at the first, is the same in distribution with the array S after an unusual loop is performed. If $\left\{p^{*(0)}_{u,v}\right\}_{0\leq u,v\leq N-1}$ is the initial distribution for S^*, then by construction we have the following relation

$$p^{*(0)}_{u,v} = p'^{(0)}_{(u+1) \bmod N,v} \ , \ \forall \ u, v \in \mathbb{Z}_N.$$

If $\left\{p^{*(N-1)}_{u,v}\right\}_{0\leq u,v\leq N-1}$ is the distribution of S^* after $(N-1)$ steps of the usual loop, then

$$p^{*(N-1)}_{u,v} = p'^{(N-1)}_{(u+1) \bmod N,v} \ , \ \forall \ u, v \in \mathbb{Z}_N.$$

We again define,

$$\boldsymbol{p}'^{(k)}_v := (p'^{((N-1)k)}_{0,v}, \ldots, p'^{((N-1)k)}_{N-1,v})^T, \ k = 0, 1; \ v \in \mathbb{Z}_N,$$

and

$$\boldsymbol{p}^{*(k)}_v := (p^{*((N-1)k)}_{0,v}, \ldots, p^{*((N-1)k)}_{N-1,v})^T, \ k = 0, 1; \ v \in \mathbb{Z}_N.$$

As $\boldsymbol{p}^{*(1)}_v$ is obtained by performing first $(N-1)$ operations on the initial distribution $\boldsymbol{p}^{*(0)}_v$, using Lemma 1 and Lemma 2 we can write,

$$\boldsymbol{p}^{*(1)}_v = \mathbf{1}_N \boldsymbol{x}'^T \boldsymbol{p}^{*(0)}_v - B' \boldsymbol{p}^{*(0)}_v + \boldsymbol{\psi}', \ \forall \ v \in \mathbb{Z}_N,$$

where,

$$B' := \begin{bmatrix} q^{N-2}\boldsymbol{y}_0^T \\ q^{N-3}\boldsymbol{y}_1^T \\ \cdots \\ q^0 \boldsymbol{y}_{N-2}^T \\ -\zeta^T \end{bmatrix}.$$

Then define,

$$A' = \mathbf{1}_N \boldsymbol{x}'^T - B',$$

which gives

$$\boldsymbol{p}_v^{*(1)} = A'\boldsymbol{p}_v^{*(0)} + \boldsymbol{\psi}',\ \forall\ v \in \mathbb{Z}_N.$$

Now we want to write everything in terms of $\boldsymbol{p}_v'^{(k)}$s. Denote by P, the permutation matrix defined as

$$P := \begin{bmatrix} \mathbf{0}_{N-1} & I_{N-1} \\ 1 & \mathbf{0}_{N-1}^T \end{bmatrix}.$$

Then, $P\boldsymbol{p}_v'^{(k)} = \boldsymbol{p}_v^{*(k)}$, which gives,

$$P\boldsymbol{p}_v'^{(1)} = A'P\boldsymbol{p}_v'^{(0)} + \boldsymbol{\psi}',\ \forall\ v \in \mathbb{Z}_N,$$

i.e.,

$$\boldsymbol{p}_v'^{(1)} = P^{-1}A'P\boldsymbol{p}_v'^{(0)} + P^{-1}\boldsymbol{\psi}',\ \forall\ v \in \mathbb{Z}_N.$$

It is also easy to see from Lemma 1 and Lemma 2 that

$$\boldsymbol{p}_v'^{(0)} = p\mathbf{1}_N \Rightarrow \boldsymbol{p}_v^{*(0)} = p\mathbf{1}_N \Rightarrow \boldsymbol{p}_v^{*(1)} = p\mathbf{1}_N \Rightarrow \boldsymbol{p}_v'^{(1)} = p\mathbf{1}_N,$$

which gives

$$p\mathbf{1}_N = pP^{-1}A'P\mathbf{1}_N + P^{-1}\boldsymbol{\psi}'.$$

So, if we define $\boldsymbol{\delta}_v'^{(k)} := \boldsymbol{p}_v'^{(1)} - p\mathbf{1}_N$, and $A_0 := P^{-1}A'P$, then we have the following result

Theorem 5. *If $\boldsymbol{\delta}_v'^{(0)}$ and $\boldsymbol{\delta}_v'^{(1)}$ denote the deviation of the distribution of the element v in the state array, before and after an unusual loop respectively, from the uniform distribution, then*

$$\boldsymbol{\delta}_v'^{(1)} = A_0\boldsymbol{\delta}_v'^{(0)},\ \forall\ \ v \in \mathbb{Z}_N.$$

As a corollary we can also write the following.

Corollary 1. *If $\boldsymbol{\delta}_v'^{(0)}$ and $\boldsymbol{\delta}_v'^{(k)}$ denote the deviation of the distribution of the index v on the state array, before and after $k \geq 1$ unusual loops respectively, from the uniform distribution, then*

$$\boldsymbol{\delta}_v'^{(k)} = A_0^k\boldsymbol{\delta}_v'^{(0)},\ \forall\ \ v \in \mathbb{Z}_N.$$

Therefore, after every usual and unusual loop, the deviation vector for each v is pre-multiplied by the matrix A and A_0 respectively. We therefore refer these matrices by the *associated matrix* of the two loops respectively.

Now using Theorem 4 and Theorem 5 we have the following result.

Theorem 6. *consider KSA*,PRGA* together performing k loops, where $k = k_1 + k_2$, and k_1 loops are of KSA* and k_2 loops are of PRGA*, and they have been performed sequentially, i.e., the KSA* loops are performed first. If $\boldsymbol{\delta}_v^{(0,0)}$ and $\boldsymbol{\delta}_v^{(k_1,k_2)}$ denote the deviation of the distribution of the index v on the state array, before and after these k loops respectively, from the uniform distribution, then if $k_2 \geq 1$, we have*

$$\boldsymbol{\delta}_v^{(k_1,k_2)} = A^{k_2-1} A_0 A^{k_1} \boldsymbol{\delta}_v^{(0,0)}, \ \forall \ \ v \in \mathbb{Z}_N.$$

one of the two special cases of Theorem 6 are when KSA* is performed k times, which means $k = k_1, k_2 = 0$ and for this the result is already specified in Theorem 4. The other special case happens when we consider the usual RC4 scenario where KSA* is performed once and the rest are PRGA*. This corresponds to the case $k_1 = 1, k_2 = k - 1$. Then the theorem gives us following corollary.

Corollary 2. *If $\boldsymbol{\delta}_v^{(0,0)}$ and $\boldsymbol{\delta}_v^{(1,k-1)}$ denote the deviation of the distribution of the element v in the state array, before and after k loops of RC4 KSA* and PRGA* respectively, from the uniform distribution, then if $k \geq 2$, we have*

$$\boldsymbol{\delta}_v^{(1,k-1)} = A^{k-2} A_0 A \boldsymbol{\delta}_v^{(0,0)}, \ \forall \ \ v \in \mathbb{Z}_N.$$

Note that from $\boldsymbol{\delta}_v$s we can easily find out the exact distributions. Therefore Corollary 2 gives us the expression for the distribution of the state array after KSA* and $(k-1)$ loops in PRGA*.

3 Analysis of Burn-in Period

After the derivation of the state probabilities in previous section we shall now turn our attention to the rate at which the state probabilities converge to the uniform distribution.

Let us first recall Theorem 4, 5, 6. Recall that $\boldsymbol{\delta}_v^{(k_1,k_2)}$ gives us the deviation of the distribution of index v from the uniform distribution after k_1 KSA* loops and $k_2 \geq 1$ PRGA* loops. After each usual loop this deviation is multiplied by A whereas after each unusual loop it gets multiplied by A_0. But these deviations being vectors, we want to summarize them by a single quantity, i.e., we would like to take some kind of norm of this deviation vectors. Therefore, we define,

$$d_v^{(k_1,k_2)} := ||\boldsymbol{\delta}_v^{(k_1,k_2)}||,$$

$$\Delta_{(k_1,k_2)} := \max_{0 \leq v \leq N-1} d_v^{(k_1,k_2)} = \max_{0 \leq v \leq N-1} ||\boldsymbol{\delta}_v^{(k_1,k_2)}||;$$

Where $||.||$ denotes some kind of vector norm. Then by Theorem 6, we have

$$d_v^{(k_1,k_2)} = ||\boldsymbol{\delta}_v^{(k_1,k_2)}|| = ||A^{k_2-1} A_0 A^{k_1-1} \boldsymbol{\delta}_v^{(0,0)}|| \leq ||A||_M^{k_1+k_2-1} ||A_0||_M ||\boldsymbol{\delta}_v^{(0,0)}|| \ ;$$

where $||.||_M$ denotes some sort of matrix norm for which the relation $||A\boldsymbol{v}|| \leq ||A||_M ||\boldsymbol{v}||$ holds true. Our next task is to search for such suitable norm.

Our first choice should be the most popular *operator norm* for the matrix and L_2 norm for the vector. But the vectors, $\delta_v^{(k_1,k_2)}$s which we are dealing with in our case have some special properties, i.e. their coordinates add up to 0 and the absolute values of their each coordinate are less than 1. The L_2 norm ignores these special properties and therefore does not give good bound for the rate of convergence in our case (empirically confirmed). The vector norm which we shall use in our analysis is the L_∞ norm which is defined as

$$||\boldsymbol{\delta}_v^{(k_1,k_2)}||_\infty := \max_{1\le i\le N} |\boldsymbol{e}_i^T \boldsymbol{\delta}_v^{(k_1,k_2)}|;$$

where $\boldsymbol{e}_i$ is the i-th coordinate vector. The matrix norm which we shall use is very particular to fit our purpose. It is obtained in the following manner. Take any $C = ((c_{ij}))_{1\le i,j\le N}$. Suppose $c_{i(1)}, \ldots, c_{i(N)}$ are the elements of the i-th row of C in non-increasing order. We consider two cases. First consider the case where N is even. Then for any $\boldsymbol{v}$ with $\mathbf{1}_N^T\boldsymbol{v} = 0$, we have,

$$\begin{aligned}
\boldsymbol{e}_i^T C\boldsymbol{v} &= \boldsymbol{e}_i^T \sum_{j=1}^N \sum_{k=1}^N c_{kj}\boldsymbol{e}_k\boldsymbol{e}_j^T\boldsymbol{v} \\
&= \sum_{j=1}^N c_{ij}\boldsymbol{e}_j^T\boldsymbol{v} \\
&= \sum_{j=1}^N c_{ij}\Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij}.
\end{aligned}$$

Note that, $0 \le \boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty \le 2||\boldsymbol{v}||_\infty$, for all j and $\sum_{j=1}^N(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty) = N||\boldsymbol{v}||_\infty$ as $\mathbf{1}_N^T\boldsymbol{v} = 0$. Hence,

$$\begin{aligned}
\boldsymbol{e}_i^T C\boldsymbol{v} &= \sum_{j=1}^N c_{ij}\Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij} \\
&\le \sum_{j=1}^{\frac{N}{2}} c_{i(j)}\Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) + c_{i(\frac{N}{2}+1)} \sum_{j=\frac{N}{2}+1}^N \Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij} \\
&= \sum_{j=1}^{\frac{N}{2}} c_{i(j)}\Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) + \sum_{j=1}^{\frac{N}{2}} c_{i(\frac{N}{2}+1)}\Big(-\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij} \\
&\le \sum_{j=1}^{\frac{N}{2}} c_{i(j)}\Big(\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) + \sum_{j=1}^{\frac{N}{2}} c_{i(j)}\Big(-\boldsymbol{e}_j^T\boldsymbol{v} + ||\boldsymbol{v}||_\infty\Big) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij} \\
&= \sum_{j=1}^{\frac{N}{2}} c_{i(j)}(2||\boldsymbol{v}||_\infty) - ||\boldsymbol{v}||_\infty \sum_{j=1}^N c_{ij} \\
&= \left[\sum_{j=1}^{\frac{N}{2}} c_{i(j)} - \sum_{j=\frac{N}{2}+1}^N c_{i(j)}\right] ||\boldsymbol{v}||_\infty; \ \ \forall\, i.
\end{aligned}$$

Therefore,

$$||C\boldsymbol{v}||_\infty \leq \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N}{2}} c_{i(j)} - \sum_{j=\frac{N}{2}+1}^{N} c_{i(j)}\right]||\boldsymbol{v}||_\infty.$$

Similarly if N is odd, then

$$||C\boldsymbol{v}||_\infty \leq \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N+1}{2}} c_{i(j)} - \sum_{j=\frac{N+1}{2}}^{N} c_{i(j)}\right]||\boldsymbol{v}||_\infty.$$

Therefore, if we take,

$$||C||_M := \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N}{2}} c_{i(j)} - \sum_{j=\frac{N}{2}+1}^{N} c_{i(j)}\right], \text{ if } N \text{ is even;}$$

and

$$||C||_M := \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N+1}{2}} c_{i(j)} - \sum_{j=\frac{N+1}{2}}^{N} c_{i(j)}\right], \text{ if } N \text{ is odd,}$$

then it serves our purpose. Note that, this norm is invariant under row and column permutations. From now on, by *measure* of a row, we shall refer the quantity associated with the row, which we maximize over the rows to get the matrix norm.

Now our next target is to find the expressions for $||A||_M$ and $||A_0||_M$. We now have the following result.

Result 1 *Let $Np = 1$ and $q = 1 - p$. $A = ((a_{ij}))_{1\leq i,j\leq N}$. Then if N is even, we have*

$$||A||_M := \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N}{2}} a_{i(j)} - \sum_{j=\frac{N}{2}+1}^{N} a_{i(j)}\right] = \left[\sum_{j=1}^{\frac{N}{2}} a_{1(j)} - \sum_{j=\frac{N}{2}+1}^{N} a_{1(j)}\right];$$

and this expression simplifies to

$$||A||_M = 1 - 2q^{\frac{N}{2}} - q^{N-1} + 2q^N.$$

On the other hand if N is odd and $N \geq 15$, then

$$||A||_M := \max_{1\leq i\leq N}\left[\sum_{j=1}^{\frac{N+1}{2}} a_{i(j)} - \sum_{j=\frac{N+1}{2}}^{N} a_{i(j)}\right] = \left[\sum_{j=1}^{\frac{N+1}{2}} a_{1(j)} - \sum_{j=\frac{N+1}{2}}^{N} a_{1(j)}\right],$$

which simplifies to

$$||A||_M = 1 - q^{\frac{N-1}{2}} - q^{\frac{N+1}{2}} - q^{N-1} + 2q^N.$$

The proof of this result is pretty long and full of calculations, hence it is given in Appendix A.

Let us denote the term $||A||_M$ by η_N to point out that it is a function of N. Using the fact that $q^N \longrightarrow e^{-1}$ and $q \longrightarrow 1$ if $N \longrightarrow \infty$, we have

$$\eta_N \approx 1 - \frac{2}{\sqrt{e}} - \frac{1}{e} + \frac{2}{e} = (1 - \frac{1}{\sqrt{e}})^2, \quad \text{for large } N.$$

Let us call this term as $\alpha := (1 - \frac{1}{\sqrt{e}})^2$. The approximation $\eta_N \approx \alpha$ is quite good for $N \geq 60$, with the approximation error being less than 0.005, as seen from Fig. 1. Also note that $\eta_N \leq \alpha$, for all N.

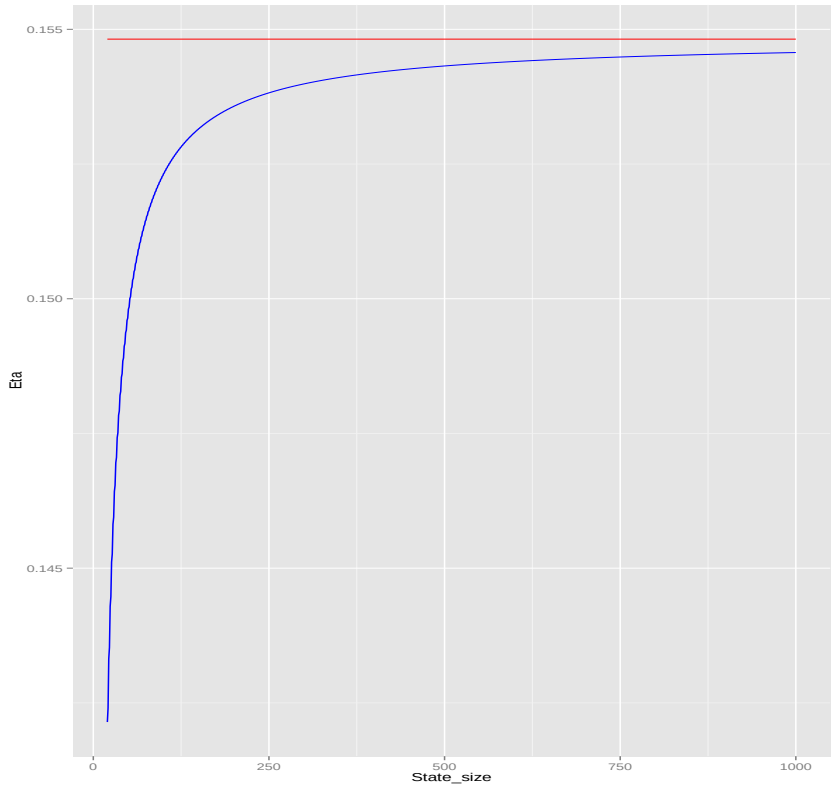

Fig. 1. Plot of η_N versus N

It now remains to compute the value of $||A_0||_M$. As already observed, the norm is invariant under column and row operations, we get $||A_0||_M = ||A'||_M$. Look at the k-th row of the matrix A'. It is

$$(p - pq^{N-k-1}, \ldots, pq^{k-1} - pq^{N-2}, pq^k, \ldots, pq^{N-2}, 0),$$

for $k = 1, \ldots, N-1$. Compare it with the $(k+1)$-th row of A, which looks like

$$(p - pq^{N-k-1}, \ldots, pq^{k-1} - pq^{N-2}, pq^k - pq^{N-1}, pq^{k+1}, \ldots, pq^{N-2}, pq^{N-1}).$$

The interesting thing is to note that, only two terms differ in the above which gives us that the measure for these rows differ at most by $2pq^{N-1}$. The last row of the matrix A' is

$$(p, pq, \ldots, pq^{N-2}, q^{N-1}),$$

and the measure for this row is therefore $q^{N-1} + p(1 + \cdots + q^{\frac{N}{2}-2}) - p(q^{\frac{N}{2}-1} + \cdots + q^{N-2}) = q^{N-1} + (1 - q^{\frac{N}{2}-1}) - q^{\frac{N}{2}-1}(1 - q^{\frac{N}{2}}) = 1 - 2q^{\frac{N}{2}-1} + 2q^{N-1}$. Therefore,

$$||A_0||_M \leq max(||A||_M + 2pq^{N-1}, 1 - 2q^{\frac{N}{2}-1} + 2q^{N-1}).$$

Using the approximation which we have used earlier we conclude that

$$||A_0||_M \leq 1 - \frac{2}{\sqrt{e}} + \frac{2}{e} = (1 - \frac{1}{\sqrt{e}})^2 + \frac{1}{e} =: \beta.$$

Now using Theorem 6 we can conclude the following theorem.

Theorem 7. *Suppose $\boldsymbol{\delta}_v^{(0,0)}$ and $\boldsymbol{\delta}_v^{(k_1,k_2)}$ denote the deviation of the distribution of the index v on the state array, before and after k_1 loops of KSA* and k_2 loops of PRGA* respectively, from the uniform distribution. The first loop of PRGA* is always an unusual loop, and we allow the different loops of KSA* to PRGA* to be performed in any pre-fixed order. Then for $N \geq 15$, if $k_2 \geq 1$, we have,*

$$||\boldsymbol{\delta}_v^{(k_1,k_2)}||_\infty \leq \alpha^{(k_1+k_2-1)}\beta||\boldsymbol{\delta}_v^{(0,0)}||_\infty, \ \forall \ v \in \mathbb{Z}_N,$$

and if $k_2 = 0$ then we have

$$||\boldsymbol{\delta}_v^{(k_1,0)}||_\infty \leq \alpha^{k_1}||\boldsymbol{\delta}_v^{(0,0)}||_\infty, \ \forall \ v \in \mathbb{Z}_N,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$, and $\beta = \alpha + \frac{1}{e}$

In the above theorem we have allowed a slight flexibility than Theorem 6 that the different loops can be performed in any order. This is possible because after taking the norm all the norms corresponding to A will come together.

We shall do now two special cases of Theorem 7. Considering only KSA* is performed k times, We can write a corollary to Theorem 7 by taking $k_1 = k, k_2 = 0$.

Corollary 3. *Consider $k \geq 1$. Suppose $\boldsymbol{\delta}_v^{(0)}$ and $\boldsymbol{\delta}_v^{(k)}$ denote the deviation of the distribution of the index v on the state array, before and after k loops of KSA*, from the uniform distribution. Then for $N \geq 15$,*

$$||\boldsymbol{\delta}_v^{(k)}||_\infty \leq \alpha^k||\boldsymbol{\delta}_v^{(0)}||_\infty, \ \forall \ v \in \mathbb{Z}_N,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$.

Another corollary can be written considering the usual RC4 scenario, i.e., $k_1 = 1, k_2 = k - 1$.

Corollary 4. *Consider $k \geq 2$. If $\boldsymbol{\delta}_v^{(0,0)}$ and $\boldsymbol{\delta}_v^{(1,k-1)}$ denote the deviation of the distribution of the index v on the state array, before and after k loops of RC4 KSA*-PRGA* respectively, from the uniform distribution, then for $N \geq 15$,*

$$||\boldsymbol{\delta}_v^{(1,k-1)}||_\infty \leq \alpha^{(k-1)}\beta||\boldsymbol{\delta}_v^{(0,0)}||_\infty, \ \forall \ v \in \mathbb{Z}_N,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$, and $\beta = \alpha + \frac{1}{e}$.

Recall the definition of Δ_k. Similarly we can define

$$\Delta_k^* := \max_{0 \leq v \leq N-1} ||\boldsymbol{\delta}_v^{(1,k-1)}||;$$

Thus Δ_k^* is a measure of the deviation of the distribution of the state array from the uniform distribution after k loops of RC4 KSA*-PRGA*. Taking maximum w.r.t. v on both sides of Corollary 4, we get the following result.

Lemma 9. *Consider $k \geq 2$. Then for all $N \geq 15$,*

$$\Delta_k^* \leq \alpha^{(k-1)}\beta\Delta_0^*,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$, and $\beta = \alpha + \frac{1}{e}$.

We now have to answer the question, *how small Δ_k^* would be convenient to say that the distribution of the state array after k loops of KSA* and PRGA* is almost uniform ?*. This value should depend on N, as even small departure from the uniform distribution for large N would be significant since the uniform probability in each cell is very small in magnitude. We therefore set the threshold to be of $O(N^{-\delta})$ for some convenient δ chosen. This δ should be greater than 1 though, in order to make the bound sensible.

We shall here work with the threshold $N^{-\delta}$. Let R_N is the minimum number of loops after which the departure from uniform, Δ_k becomes less than $N^{-\delta}$. Using Lemma 9 we can write

$$\alpha^{(k-1)}\beta\Delta_0^* \leq N^{-\delta}. \tag{6}$$

In other words,

$$k \geq \frac{-\delta \log N - \log \Delta_0^* - \log \beta}{\log \alpha} + 1, \tag{7}$$

as $\alpha < 1$. Combining (6) and (7), we can write the following theorem,

Theorem 8. *For $N \geq 15$,*

$$R_N \leq \lceil \frac{-\delta \log N - \log \Delta_0^* - \log \beta}{\log \alpha} + 1 \rceil.$$

4 Application to RC4 and RC4$^+$: A Comparative Study

4.1 Application to RC4:

Let us apply the theorem to RC4 scenario. Here, $N = 256$, and we take $\delta = 3$, which shall ensure that the maximum deviation from the uniform distribution will be less than 10^{-7}. To calculate Δ_0^*, we observe that

$$\boldsymbol{p}_v^{(0)} = \boldsymbol{e}_v, \ \ \forall v \in \mathbb{Z}_N,$$

where $\boldsymbol{e}_v$ is the $(v+1)$-th co-ordinate vector in $\mathbb{R}^N$. Hence,

$$||\boldsymbol{\delta}_v^{(0,0)}||_\infty = ||\boldsymbol{p}_v^{(0)} - p\mathbf{1}_N||_\infty = q, \forall v \in \mathbb{Z}_N,$$

which implies

$$\Delta_0^* = q.$$

Therefore, putting the numerical values in Theorem 8, we get that that, $R_N \leq 10$. This result implies that total 10 loops is sufficient in RC4 to get a state-space distribution which deviates from the uniform distribution by at most 10^{-7}.

4.2 Application to RC4$^+$:

RC4$^+$ is a refinement of the RC4 algorithm, with some more state array scrambling steps. Like RC4, we also here work with an idealized model of RC4$^+$ consistent with the intention of the cipher designer. We shall concentrate here on only on key scheduling part of it, as the idealized model for its PRGA part does not exactly fall in th framework we have developed. The KSA$^+$ algorithm is given below.

Algorithm KSA$^+$

Initialization:
For $i \leftarrow 0, \ldots, N-1$
$S[i] \leftarrow i$;

Basic Scrambling:
For $i \leftarrow 0, \ldots, N-1$
$j \leftarrow Uniform\{0, \ldots, N-1\}$;
Swap$(S[i], S[j])$;

IV Scrambling:
For $i \leftarrow \frac{N}{2} - 1, \ldots, 0$
$j \leftarrow Uniform\{0, \ldots, N-1\}$;
Swap$(S[i], S[j])$;
For $i \leftarrow \frac{N}{2}, \ldots, N-1$
$j \leftarrow Uniform\{0, \ldots, N-1\}$;
Swap$(S[i], S[j])$;

Zigzag scrambling:
For $y \leftarrow 0, \ldots, N-1$
if $y \cong 0 \bmod 2$ then
$i \leftarrow \frac{y}{2}$
else
$i \leftarrow N - \frac{y+1}{2}$
$j \leftarrow Uniform\{0, \ldots, N-1\}$;
Swap$(S[i], S[j])$;

So, KSA$^+$ contains three types of loops, usual loop, IV scrambling loop and the zigzag loop. The operations in each loop are exactly the same except for the fact that the index i is varied over 0 to $N-1$ in different permutations. Let us develop a theory for this in general. Consider a permutation π of $\{0,\ldots,N-1\}$, say $(\pi(0),\ldots,\pi(N-1))$. Consider the loop described below.

π *scrambling*:
For $i \leftarrow 0,\ldots,N-1$
$j \leftarrow Uniform\,\{0,\ldots,N-1\}$;
Swap$(S[\pi(i)],S[j])$;

Suppose $\{p_{u,v}\}_{0\le u,v\le N-1}$ is the distribution entering the loop with the usual notations carried out, and $\left\{p^*_{u,v}\right\}_{0\le u,v\le N-1}$ be the distribution after the loop is carried out. Consider a transformed state array S' obtained from S by the relation $S'[i]=S[\pi(i)]$. Then the initial distribution on S' is

$$\left\{\Pr[S'_0[u]=v] := q_{u,v} = p_{\pi(u),v}\right\}_{0\le u,v\le N-1},$$

and the final distribution is

$$\left\{\Pr[S'_N[u]=v] := q^*_{u,v} = p^*_{\pi(u),v}\right\}_{0\le u,v\le N-1}.$$

On the other hand, the scrambling loop on S is actually an usual loop on S', as $S[j] \stackrel{d}{=} S[\pi(j)]$ if $j \leftarrow Uniform\,\{0,\ldots,N-1\}$. Now let us define $\boldsymbol{p}_v$,$\boldsymbol{p}^*_v$,$\boldsymbol{q}_v$,$\boldsymbol{q}^*_v$ in the usual way as defined earlier, for each index v. And then define the deviations from the uniform as $\boldsymbol{\delta}_v = \boldsymbol{p}_v - p\mathbf{1}_N$,$\boldsymbol{\delta}^*_v = \boldsymbol{p}^*_v - p\mathbf{1}_N$,$\boldsymbol{\delta}'_v = \boldsymbol{q}_v - p\mathbf{1}_N$,$\boldsymbol{\delta}'^*_v = \boldsymbol{q}^*_v - p\mathbf{1}_N$. Let us now write some relations which are obvious from the context.

$$P_\pi \boldsymbol{\delta}_v = \boldsymbol{\delta}'_v;\ P_\pi \boldsymbol{\delta}^*_v = \boldsymbol{\delta}'^*_v,$$

where P_π is the permutation matrix corresponding to the permutation π. And Theorem 6 gives

$$\boldsymbol{\delta}'^*_v = A\boldsymbol{\delta}'_v,$$

and therefore,

$$\boldsymbol{\delta}^*_v = P_\pi^{-1} A P_\pi \boldsymbol{\delta}'_v = A_\pi \boldsymbol{\delta}'_v,$$

where $A_\pi = P_\pi^{-1} A P_\pi$. Note that A_π is a matrix obtained by just row and column permutation of A, and therefore,

$$||A_\pi||_M = ||A||_M.$$

As the norm of the associated matrix denotes the reduction in the deviation, we confirm that, all the permutation loops are same w.r.t. that criterion. Hence, we can write down the following two results.

Lemma 10. *Consider $k \geq 2$. If $\boldsymbol{\delta}_v^{(0)}$ and $\boldsymbol{\delta}_v^{(k)}$ denote the deviation of the distribution of the index v on the state array, before and after k loops of KSA$^+$ respectively, (with the terminology that each KSA$^+$ consists of three loops), from the uniform distribution, then for $N \geq 15$,*

$$||\boldsymbol{\delta}_v^{(k)}||_\infty \leq \alpha^k ||\boldsymbol{\delta}_v^{(0)}||_\infty, \ \forall \ v \in \mathbb{Z}_N,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$.

Lemma 11. *Consider $k \geq 2$. If $\boldsymbol{\delta}_v^{(0)}$ and $\boldsymbol{\delta}_v^{(k)}$ denote the deviation of the distribution of the index v on the state array, before and after k loops of KSA$^+$ respectively, (with the terminology that each KSA$^+$ consists of three loops), from the uniform distribution, and*

$$\Delta_k^+ := \max_{0 \leq v \leq N-1} ||\boldsymbol{\delta}_v^{(k)}||_\infty.$$

Then for $N \geq 15$,

$$\Delta_k^+ \leq \alpha^k \Delta_0^+,$$

where $\alpha = (1 - \frac{1}{\sqrt{e}})^2$.

Theorem 9. *If R_N^+ is the minimum number of loops in KSA$^+$ after which the departure from the uniform distribution becomes less than $N^{-\delta}$, then*

$$R_N^+ \leq \lceil \frac{-\delta \log N - \log \Delta_0^+}{\log \alpha} \rceil,$$

for $N \geq 15$, and Δ_0^+ is as defined in Lemma 11.

As a remark we mention that, if N is even, then the bound $N \geq 15$ is not needed.

5 Conclusion

RC4 stream cipher has a nice combinatorial structure. The evolution of its internal state, which is a permutation over $\mathbb{Z}_N$, is drived by a random key with a hope to reach an almost uniformly random permutation as quickly as possible. In this paper, we perform rigorous analysis of this evolution and determine how the distance between the RC4 internal state and an uniformly random permutation decreases as a function of the number of rounds. This yields interesting results on the "burn-in" period of the RC4 structure, i.e., the number of KSA-PRGA rounds before the cipher can be used in practice. One caveat of our analysis is that we assume the pseudo-random index j of RC4 permutation to be uniformly random over $\mathbb{Z}_N$. This is a standard assumption in all the prior works as well. In reality, however, this assumption is not true. It remains an interesting open question if similar analysis can be performed by first deriving the exact distribution of j values in different rounds and then using this distribution in the computation of burn-in period.

References

1. Carlisle M. Adams, Ali Miri, and Michael J. Wiener, editors. *Selected Areas in Cryptography, 14th International Workshop, SAC 2007, Ottawa, Canada, August 16-17, 2007, Revised Selected Papers*, volume 4876 of *Lecture Notes in Computer Science*. Springer, 2007.
2. Nadhem J. AlFardan, Daniel J. Bernstein, Kenneth G. Paterson, Bertram Poettering, and Jacob C. N. Schuldt. On the security of RC4 in TLS. In Samuel T. King, editor, *Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013*, pages 305–320. USENIX Association, 2013.
3. Eli Biham and Yaniv Carmeli. Efficient reconstruction of RC4 keys from internal states. In Nyberg [16], pages 270–288.
4. Scott R. Fluhrer, Itsik Mantin, and Adi Shamir. Weaknesses in the key scheduling algorithm of RC4. In Serge Vaudenay and Amr M. Youssef, editors, *Selected Areas in Cryptography, 8th Annual International Workshop, SAC 2001 Toronto, Ontario, Canada, August 16-17, 2001, Revised Papers*, volume 2259 of *Lecture Notes in Computer Science*, pages 1–24. Springer, 2001.
5. Scott R. Fluhrer and David A. McGrew. Statistical analysis of the alleged RC4 keystream generator. In Bruce Schneier, editor, *Fast Software Encryption, 7th International Workshop, FSE 2000, New York, NY, USA, April 10-12, 2000, Proceedings*, volume 1978 of *Lecture Notes in Computer Science*, pages 19–30. Springer, 2000.
6. Christina Garman, Kenneth G. Paterson, and Thyla Van der Merwe. Attacks only get better: Password recovery attacks against RC4 in TLS. In Jaeyeon Jung and Thorsten Holz, editors, *24th USENIX Security Symposium, USENIX Security 15, Washington, D.C., USA, August 12-14, 2015.*, pages 113–128. USENIX Association, 2015.
7. Jovan Dj. Golic. Linear statistical weakness of alleged RC4 keystream generator. In Walter Fumy, editor, *Advances in Cryptology - EUROCRYPT '97, International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, May 11-15, 1997, Proceeding*, volume 1233 of *Lecture Notes in Computer Science*, pages 226–238. Springer, 1997.
8. Sourav Sen Gupta, Subhamoy Maitra, Goutam Paul, and Santanu Sarkar. (non-)random sequences from (non-)random permutations - analysis of RC4 stream cipher. *J. Cryptology*, 27(1):67–108, 2014.
9. R. J. Jenkins. Isaac and rc4, 1996.
10. Subhamoy Maitra and Goutam Paul. New form of permutation bias and secret key leakage in keystream bytes of RC4. In Nyberg [16], pages 253–269.
11. Istik Mantin. The security of the stream cipher rc4. Master Thesis, The Weizmann Institue of Science, 2001.
12. Itsik Mantin. A practical attack on the fixed RC4 in the WEP mode. In Bimal K. Roy, editor, *Advances in Cryptology - ASIACRYPT 2005, 11th International Conference on the Theory and Application of Cryptology and Information Security, Chennai, India, December 4-8, 2005, Proceedings*, volume 3788 of *Lecture Notes in Computer Science*, pages 395–411. Springer, 2005.
13. Itsik Mantin. Predicting and distinguishing attacks on RC4 keystream generator. In Ronald Cramer, editor, *Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings*, volume 3494 of *Lecture Notes in Computer Science*, pages 491–506. Springer, 2005.

14. Itsik Mantin and Adi Shamir. A practical attack on broadcast RC4. In Mitsuru Matsui, editor, *Fast Software Encryption, 8th International Workshop, FSE 2001 Yokohama, Japan, April 2-4, 2001, Revised Papers*, volume 2355 of *Lecture Notes in Computer Science*, pages 152–164. Springer, 2001.
15. Ilya Mironov. (not so) random shuffles of RC4. In Moti Yung, editor, *Advances in Cryptology - CRYPTO 2002, 22nd Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 2002, Proceedings*, volume 2442 of *Lecture Notes in Computer Science*, pages 304–319. Springer, 2002.
16. Kaisa Nyberg, editor. *Fast Software Encryption, 15th International Workshop, FSE 2008, Lausanne, Switzerland, February 10-13, 2008, Revised Selected Papers*, volume 5086 of *Lecture Notes in Computer Science*. Springer, 2008.
17. Goutam Paul and Subhamoy Maitra. Permutation after RC4 key scheduling reveals the secret key. In Adams et al. [1], pages 360–377.
18. Goutam Paul, Subhamoy Maitra, and Rohit Srivastava. On non-randomness of the permutation after RC4 key scheduling. In Serdar Boztas and Hsiao-feng Lu, editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, 17th International Symposium, AAECC-17, Bangalore, India, December 16-20, 2007, Proceedings*, volume 4851 of *Lecture Notes in Computer Science*, pages 100–109. Springer, 2007.
19. Goutam Paul, Siddheshwar Rathi, and Subhamoy Maitra. On non-negligible bias of the first output byte of RC4 towards the first three bytes of the secret key. *Des. Codes Cryptography*, 49(1-3):123–134, 2008.
20. Ronald L Rivest and Jacob CN Schuldt. Spritz–A spongy RC4-like stream cipher and hash function. CRYPTO 2014 Rump Session, 2014.
21. Andrew Roos. A class of weak keys in the rc4 stream cipher, 1995.
22. Pouyan Sepehrdad, Petr Susil, Serge Vaudenay, and Martin Vuagnoux. Smashing WEP in a passive attack. In Shiho Moriai, editor, *Fast Software Encryption - 20th International Workshop, FSE 2013, Singapore, March 11-13, 2013. Revised Selected Papers*, volume 8424 of *Lecture Notes in Computer Science*, pages 155–178. Springer, 2013.
23. Pouyan Sepehrdad, Serge Vaudenay, and Martin Vuagnoux. Discovery and exploitation of new biases in RC4. In Alex Biryukov, Guang Gong, and Douglas R. Stinson, editors, *Selected Areas in Cryptography - 17th International Workshop, SAC 2010, Waterloo, Ontario, Canada, August 12-13, 2010, Revised Selected Papers*, volume 6544 of *Lecture Notes in Computer Science*, pages 74–91. Springer, 2010.
24. Pouyan Sepehrdad, Serge Vaudenay, and Martin Vuagnoux. Statistical attack on RC4 - distinguishing WPA. In Kenneth G. Paterson, editor, *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings*, volume 6632 of *Lecture Notes in Computer Science*, pages 343–363. Springer, 2011.
25. Erik Tews, Ralf-Philipp Weinmann, and Andrei Pyshkin. Breaking 104 bit WEP in less than 60 seconds. In Sehun Kim, Moti Yung, and Hyung-Woo Lee, editors, *Information Security Applications, 8th International Workshop, WISA 2007, Jeju Island, Korea, August 27-29, 2007, Revised Selected Papers*, volume 4867 of *Lecture Notes in Computer Science*, pages 188–202. Springer, 2007.
26. Serge Vaudenay and Martin Vuagnoux. Passive-only key recovery attacks on RC4. In Adams et al. [1], pages 344–359.

27. D Wagner. My rc4 weak keys. *Post in sci. crypt, message-id 447o1l $ cbj@ cnn. princeton. EDU*, 26:1–1, 1995.

A Proof of Result 1

Recall our previous notation that $p = \frac{1}{N}$ and $q = 1 - p$. Recall the definition of A. Note that k-th row of A looks like

$$(p(1 - q^{N-k}), \ldots, p(q^{k-1} - q^{N-1}), pq^k, \ldots, pq^{N-1}),\ \forall\ 1 \leq k \leq N.$$

One observation which we shall use repeatedly in the proof is that $(1 - \frac{1}{N})^N$ i.e., q^N is increasing in N, while q^{N-1} is decreasing in N. This can be proved simply by differentiating the functions $(1 - \frac{1}{x})^x$ and $(1 - \frac{1}{x})^{x-1}$ respectively.

Case 1: N is even

Let J denotes a subset of size $\dfrac{N}{2}$ of the index set $\{1, \ldots, N\}$. Then

$$\begin{aligned}
||A||_M &= \max_{1 \leq i \leq N} \left[\sum_{j=1}^{\frac{N}{2}} a_{i(j)} - \sum_{j=\frac{N}{2}+1}^{N} a_{i(j)} \right] \\
&= \max_{1 \leq i \leq N} \max_J \left[\sum_{j \in J} a_{ij} - \sum_{j \notin J} a_{ij} \right] \\
&= \max_{1 \leq i \leq N} \left[\max_J \left[2 \sum_{j \in J} a_{ij} \right] - \sum_{j=1}^{N} a_{ij} \right].
\end{aligned}$$

The first row of A is the following,

$$p(1 - q^{N-1}, q, \ldots, q^{N-1});$$

and let us define,

$$2p(1 + q + \cdots + q^{\frac{N}{2}-1} - q^{N-1}) - p(1 + \cdots + q^{N-2}) =: I.$$

Let us take a closure look of the structure of row k of A. Note that,

$$a_{k1} > \cdots > a_{kk};\ \ a_{k,k+1} > \cdots > a_{k,N}.$$

Therefore,

$$\max_J \left[\sum_{j \in J} a_{kj} \right] = \max_{(l,m) \in \mathcal{Q}_k} \left[\sum_{i=1}^{l} a_{k,k+i} + \sum_{j=1}^{m} a_{k,j} \right],$$

where $\mathcal{Q}_k := \left\{ (l,m) |\ l,m \geq 0;\ l+m = \frac{N}{2}, k+l \leq N, m \leq k \right\}$. So, let us define,

$$\begin{aligned} D_{k,l,m} &= 2\left[\sum_{i=1}^{l} a_{k,k+i} + \sum_{j=1}^{m} a_{k,j} \right] - \sum_{j=1}^{N} a_{kj} \\ &= 2p(q^k + \cdots + q^{k+l-1} + (1-q^{N-k}) + \cdots + (q^{m-1} - q^{N+m-k-1})) - p(1 + \cdots + q^{N-k-1}). \end{aligned}$$

Our target is to show

$$D_{k,l,m} \leq I,\ \forall (l,m) \in \mathcal{Q}_k, k \in \{1, \ldots, N\}. \tag{8}$$

We shall only show for $N \geq 10$. For smaller values of N, i.e., $N = 2, 4, 6, 8$, the correctness of the result can be checked directly by calculating the matrix A.

Let us consider first the case $k = 1$. Then $m = 0, 1$, and $D_{1,\frac{N}{2}-1,1} = I$. So, it is enough to show $1 - q^{N-1} \geq q^{\frac{N}{2}}$, as this will imply that $D_{1,\frac{N}{2},0} \leq D_{1,\frac{N}{2}-1,1}$. Now we have, q^{N-1} decreasing in N and hence, $q^{N-1} + q^{\frac{N-1}{2}}$ is decreasing in N. Therefore,

$$q^{N-1} + q^{\frac{N}{2}} \leq q^{N-1} + q^{\frac{N-1}{2}} \leq (\frac{15}{16})^{15} + (\frac{15}{16})^{7.5} < 1,\ \ \forall\ N \geq 16.$$

For smaller values of N we have to check directly from the expression.

Now, we have to consider the case where $k \geq 2$. First consider $l, m > 0$.Then, $0 < l, m < \frac{N}{2}$, and

$$\begin{aligned} I - D_{k,l,m} &= 2p(\sum_{i=0}^{\frac{N}{2}-1} q^i - \sum_{i=0}^{m-1} q^i + \sum_{j=N-k}^{N-k+m-1} q^j - \sum_{j=k}^{k+l-1} q^j - q^{N-1}) \\ &\quad -p\sum_{j=0}^{N-2} q^j + p\sum_{j=0}^{N-k-1} q^j \\ &= 2p(\sum_{i=m}^{\frac{N}{2}-1} q^i + \sum_{j=N-k}^{N-k+m-1} q^j - \sum_{j=k}^{k+l-1} q^j - q^{N-1}) - p\sum_{j=N-k}^{N-2} q^j \end{aligned}$$

Therefore, enough to show

$$2(\sum_{i=m}^{\frac{N}{2}-1} q^i + \sum_{j=N-k}^{N-k+m-1} q^j - \sum_{j=k}^{k+l-1} q^j - q^{N-1}) \geq \sum_{j=N-k}^{N-2} q^j \tag{9}$$

Note that, if $m > k-2$, i.e. $m = k-1, k$, then

$$\sum_{i=m}^{\frac{N}{2}-1} q^i - \sum_{j=k}^{k+l-1} q^j = \sum_{i=m}^{m+l-1} q^i - \sum_{i=k}^{k+l-1} q^i \geq 0,$$

and

$$\sum_{j=N-k}^{N-k+m-1} q^j \geq \sum_{j=N-k}^{N-2} q^j;\ \ \sum_{j=N-k}^{N-k+m-1} q^j \geq q^{N-1},$$

which ensures that (9) holds true. So, now we should consider only the case $m \leq k-2$. Then $m+2 \leq k \leq \frac{N}{2}+m$ as $k+l \leq N$. In this case we can simplify (9) and conclude that it is enough to show

$$2(1-q^{k-m})(\sum_{i=m}^{\frac{N}{2}-1} q^i) + \sum_{j=k}^{N-k+m-1} q^j - \sum_{j=N-k+m}^{N-2} q^j - 2q^{N-1} \geq 0, \tag{10}$$

i.e.

$$2(1-q^{k-m})q^m \frac{1-q^l}{1-q} + q^{N-k}\frac{1-q^m}{1-q} - q^{N-k+m}\frac{1-q^{k-m}}{1-q} - q^{N-1} \geq 0,$$

i.e., dividing both sides by $\frac{q^{N-k+m}}{1-q}$ we conclude that it is enough to show

$$2(1-q^l)q^{k-N}(1-q^{k-m}) + (2q-1)q^{k-m-1} + q^{-m} \geq 2.$$

Let us define,

$$u_{N,m} = 2(1-q^{\frac{N}{2}-m})q^{-N};\ v_{N,m} = (2q-1)q^{-m-1},$$

and

$$E_{k,N,m} = u_{N,m}(q^k - q^{2k-m}) + v_{N,m}q^k + q^{-m},$$

and therefore we have to show

$$E_{k,N,m} \geq 2.$$

Note that

$$E_{k,N,m} - E_{k+1,N,m} = pq^k\left[u_{N,m}(1-(1+q)q^{k-m}) + v_{N,m}\right],$$

and $u_{N,m}(1-(1+q)q^{k-m})+v_{N,m}$ is an non-decreasing function of k when n, m are held at constant. Therefore,

$$E_{k,N,m} \geq E_{k+1,N,m} \Rightarrow E_{k+1,N,m} \geq E_{k+2,N,m},$$

i.e., when $E_{k,N,m}$ starts increasing it goes on increasing. Therefore, to find the minimum it is enough to search at the extremes i.e. $k = m+2$ and $k = m+\frac{N}{2}$. We shall instead search for the minimum at $k = m, m+\frac{N}{2}$ as it will suffice.

$$E_{m,N,m} = 2 - q^{-1} + q^{-m} \geq 2, \text{ as } m > 0.$$

$$\begin{aligned}
E_{m+\frac{N}{2},N,m} &= 2(1-q^{\frac{N}{2}-m})q^{m-\frac{N}{2}}(1-q^{\frac{N}{2}}) + (2q-1)q^{\frac{N}{2}-1} + q^{-m} \\
&= 2q^{m-\frac{N}{2}} - 2q^m + 4q^{\frac{N}{2}} - q^{\frac{N}{2}-1} + q^{-m} - 2 \\
&= q^{m-\frac{N}{2}} - 2q^m + 4q^{\frac{N}{2}} - q^{\frac{N}{2}-1} + (q^{m-\frac{N}{2}} + q^{-m}) - 2 \\
&\geq q^{m-\frac{N}{2}} - 2q^m + 4q^{\frac{N}{2}} - q^{\frac{N}{2}-1} + 2q^{-\frac{N}{4}} - 2,
\end{aligned}$$

where the last expression is increasing in m because $q^N \geq \frac{1}{4}$ which implies $q^{-\frac{N}{2}} \leq 2$. Therefore it is enough to check at $m = 0$ which gives us

$$\begin{aligned} E_{m+\frac{N}{2},N,m} &\geq q^{-\frac{N}{2}} + 2q^{-\frac{N}{4}} + 4q^{\frac{N}{2}} - q^{\frac{N}{2}-1} - 4 \\ &= q^{-\frac{N}{2}} + 2q^{-\frac{N}{4}} + (3 - \frac{p}{q})q^{\frac{N}{2}} - 4 \\ &\geq q^{-\frac{N}{2}} + 2q^{-\frac{N}{4}} + (3 - \frac{1}{17})q^{\frac{N}{2}} - 4, \ \forall \ N \geq 18; \end{aligned}$$

Now consider the function $x \longrightarrow x + 2\sqrt{x} + (3 - \frac{1}{17})\frac{1}{x}$, and it can easily be seen by differentiating that this function is increasing when $x^2 + x^{\frac{3}{2}} \geq (3 - \frac{1}{17})$. Now, we know, $q^{-\frac{N}{2}}$ is decreasing in N, and hence, $q^{-\frac{N}{2}} \geq \lim_{N\to\infty} q^{-\frac{N}{2}} = \sqrt{e}$ and $e + e^{\frac{3}{4}} \geq (3 - \frac{1}{17})$. Therefore we have that the function $q^{-\frac{N}{2}} + 2q^{-\frac{N}{4}} + (3 - \frac{1}{17})q^{\frac{N}{2}} - 4$ is decreasing in N and therefore the minimum value is the limiting value when N goes to ∞, i.e., $\sqrt{e} + 2e^{\frac{1}{4}} + (3 - \frac{1}{17})\frac{1}{\sqrt{e}} - 4 > 2$. Hence, $E_{m+\frac{N}{2},N,m} \geq 2, \ \forall \ N \geq 18$. And for smaller values of N i.e., for $N = 10, 12, 14, 16$, we can directly check by calculating $q^{-\frac{N}{2}} + 2q^{-\frac{N}{4}} + 4q^{\frac{N}{2}} - q^{\frac{N}{2}-1} - 4$. So, we are done with the first case of the proof except for the case that $l = 0$ or $m = 0$.

Note that, $q^k + q^{N-k} \geq 2q^{\frac{N}{2}} \geq 2(1/2) = 1$, as q^N is increasing in N. This implies that, $a_{k,k+1} \geq a_{k1}$, which in turn implies that $I \geq D_{k,1,\frac{N}{2}-1} \geq D_{k,0,\frac{N}{2}}$. Therefore the case for $l = 0$ is solved.

Now, $k > \frac{N}{2}$ implies $l \leq N - k < \frac{N}{2}$, and hence, $m > 0$. So, to consider $1 \leq k \leq \frac{N}{2}$. Note that it is enough to prove that

$$q^{k+\frac{N}{2}-1} + q^{N-k} < 1,$$

as it guarantees $a_{k1} < a_{k,k+\frac{N}{2}}$ which implies $D_{k,\frac{N}{2},0} < D_{k,\frac{N}{2}-1,1} \leq I$.

$$(q^{k+\frac{N}{2}-1}+q^{N-k})-(q^{k+1+\frac{N}{2}-1}+q^{N-k-1}) = q^{k+\frac{N}{2}-1}(1-q)-q^{N-k-1}(1-q) = (1-q)q^{N-k-1}(q^{2k-\frac{N}{2}}-1).$$

Therefore, $q^{k+\frac{N}{2}-1} + q^{N-k}$ is at first decreasing and then increasing as k varies to 1 to $\frac{N}{2}$. Hence for maximum value it is enough to check at $k = 1, \frac{N}{2}$, and at both of these points the value is $q^{\frac{N}{2}} + q^{\frac{N}{2}}$ which is less than 1, already proven. Therefore,

$$\begin{aligned} ||A||_M = I &= 2p(1 + q + \cdots + q^{\frac{N}{2}-1} - q^{N-1}) - p(1 + \cdots + q^{N-2}) \\ &= 1 - 2q^{\frac{N}{2}} - q^{N-1} + 2q^N. \end{aligned}$$

Case 2: N is odd

We shall follow similar technique as used in the even case. Let J denotes a subset of size $\frac{N-1}{2}$ of the index set $\{1,\ldots,N\}$, and u denotes a single index from the same set.. Then

$$\begin{aligned}||A||_M &= \max_{1\le i\le N}\left[\sum_{j=1}^{\frac{N+1}{2}} a_{i(j)} - \sum_{j=\frac{N+1}{2}}^{N} a_{i(j)}\right]\\ &= \max_{1\le i\le N}\max_{J,u:u\notin J}\left[\sum_{j\in J} a_{ij} - \sum_{j\notin J, j\neq u} a_{ij}\right]\\ &= \max_{1\le i\le N}\left[\max_{J,u:u\notin J}\left[2\sum_{j\in J} a_{ij} + a_{iu}\right] - \sum_{j=1}^{N} a_{ij}\right].\end{aligned}$$

The first row of A is the following,

$$p(1-q^{N-1}, q, \ldots, q^{N-1});$$

and let us define,

$$2p(1+q+\cdots+q^{\frac{N-1}{2}-1}-q^{N-1}) + pq^{\frac{N-1}{2}} - p(1+\cdots+q^{N-2}) =: I.$$

As expected from previous experience, here also we have,

$$a_{k1} > \cdots > a_{kk};\ \ a_{k,k+1} > \cdots > a_{k,N}.$$

Therefore,

$$\max_{J,u:u\notin J}\left[2\sum_{j\in J} a_{kj} + a_{ku}\right] = \max_{(l,m)\in\mathcal{Q}_k}\left[2\sum_{i=1}^{l} a_{k,k+i} + 2\sum_{j=1}^{m} a_{kj} + a_{k,k+l+1}, 2\sum_{i=1}^{l} a_{k,k+i} + 2\sum_{j=1}^{m} a_{kj} + a_{k,m+1}\right].$$

where $\mathcal{Q}_k := \left\{(l,m)|\ l,m\ge 0;\ l+m=\frac{N-1}{2}, k+l\le N, m\le k\right\}$.

So, let us define, for $l < N-k$,

$$\begin{aligned}D'_{k,l,m} &= 2\left[\sum_{i=1}^{l} a_{k,k+i} + \sum_{j=1}^{m} a_{kj} + a_{k,k+l+1}\right] - \sum_{j=1}^{N} a_{kj}\\ &= 2p(q^k+\cdots+q^{k+l-1} + (1-q^{N-k}) + \cdots + (q^{m-1}-q^{N+m-k-1})) + pq^{k+l} - p(1+\cdots+q^{N-k-1}),\end{aligned}$$

and for $m < k$,

$$\begin{aligned}D''_{k,l,m} &= 2\left[\sum_{i=1}^{l} a_{k,k+i} + \sum_{j=1}^{m} a_{kj} + a_{k,m+1}\right] - \sum_{j=1}^{N} a_{kj}\\ &= 2p(q^k+\cdots+q^{k+l-1} + (1-q^{N-k}) + \cdots + (q^{m-1}-q^{N+m-k-1})) + p(q^m - q^{m+N-k}) - p(1+\cdots+q^{N-k-1})\end{aligned}$$

Our target is to show

$$D'_{k,l,m}, D''_{k,l,m} \leq I,\ \forall (l,m) \in \mathcal{Q}_k, k \in \{1,\ldots,N\}\,, \tag{11}$$

and where they are defined. After this part the proof is completely similar to the proof given in even part. So, to avoid the repetition of the calculations, we just sketch the next part of the proof.

Let us first consider the case for $D'_{k,l,m}$. Let us consider first the case $k=1$. Then $m=0,1$, and $D'_{1,\frac{N-1}{2}-1,1} = I$. So, it is enough to show $1-q^{N-1} \geq q^{\frac{N-1}{2}}$, as this will imply that $D'_{1,\frac{N-1}{2},0} \leq D'_{1,\frac{N-1}{2}-1,1}$. Now we have, q^{N-1} decreasing in N and hence, $q^{N-1}+q^{\frac{N-1}{2}}$ is decreasing in N. Therefore,

$$q^{N-1}+q^{\frac{N-1}{2}} \leq (\frac{14}{15})^{14} + (\frac{14}{15})^{7} < 1,\ \ \forall\ N \geq 15.$$

The situations with $l=0$,$m=0$ and $m \geq k-1$ are similarly solved as for the even case. For, $l,m>0$ and $m \leq k-2$ case, simplifying the expression we see that it is enough to show

$$E'_{k,N,m} := 2(1-q^l)q^{k-N}(1-q^{k-m}) + (2q-1)q^{k-m-1} + q^{-m} + pq^{k-m-\frac{N+1}{2}} - pq^{2k+l-N-m} \geq 2.$$

Considering it as a function of k only, keeping N,m fixed, we see that

$$\frac{E'_{k,N,m} - E'_{k+1,N,m}}{pq^k}$$

is increasing in k. Therefore, $E'_{k,N,m}$ can have minimum only at two ends, i.e. $k=m, m+\frac{N+1}{2}$.

$$E'_{m,N,m} = (2q-1)q^{-1} + q^{-m} \geq 2.$$

$$\begin{aligned} E'_{\frac{N+1}{2}+m,N,m} &= 2q^{m-\frac{N-1}{2}} - 2q^{m+1} + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} + q^{-m} - 2 + p - pq^{\frac{N+1}{2}} \\ &\geq q^{m-\frac{N-1}{2}} + 2q^{-\frac{N-1}{4}} - 2q^{m+1} + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} - 2 + p - pq^{\frac{N+1}{2}}, \end{aligned}$$

and the final RHS term is increasing in m as $q^{-\frac{N+1}{2}} \leq 2$. So, enough to check at $m=0$. So, enough to show,

$$q^{-\frac{N-1}{2}} + 2q^{-\frac{N-1}{4}} - 2q + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} - 2 + p - pq^{\frac{N+1}{2}} \geq 2.$$

The LHS converges to $\sqrt{e} + 2e^{\frac{1}{4}} + 3\frac{1}{\sqrt{e}} - 4 > 2$. So, after some N LHS will be greater than 2. For some initial terms we have to check directly. Now for the last case we have to show that $D''_{k,l,m} \leq I$. The cases for $l=0, m=0$ and $m \geq k-1$ are easy to handle. For $l,m>0$ and $m \leq k-2$, simplifying the expression we see that it is enough to show,

$$E''_{k,N,m} = 2(1-q^l)q^{k-N}(1-q^{k-m}) + (2q-1)q^{k-m-1} + q^{-m} + pq^{k-m-\frac{N+1}{2}} - pq^{k-N} + p \geq 2.$$

Easy to observe that

$$\frac{E''_{k,N,m} - E''_{k+1,N,m}}{pq^k}$$

is increasing in k considering N, m fixed. So, $E''_{k,N,m}$ can have minimum only at the two ends, $k = m+1, m+\frac{N+1}{2}$. Let us define,

$$E''_{m+1,N,m} = F_{m,N}.$$

Observe that, $F_{m,N} - F_{m+1,N}$ is decreasing in m and

$$F_{0,N} - F_{1,N} = q^{-N}p(p(2q-1) - q^{N-1}) \leq 0,$$

as $2q + q^{N-2} \geq 2$. Therefore, $F_{m,N} - F_{m+1,N} \leq 0$, and hence, $F_{m,N}$ will be minimum at $m = 1$(as $m = 0$ case is done separately) and observe that,

$$\frac{F_{1,N} - 2}{p} \longrightarrow e - \sqrt{e} + 1 > 0.$$

Therefore, $F_{1,N} \geq 2$ after some terms and those cases are easy to check. This finishes checking at $k = m+1$. For other end point,

$$\begin{aligned} E''_{\frac{N+1}{2}+m,N,m} &= 2q^{m-\frac{N-1}{2}} - 2q^{m+1} + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} + q^{-m} - 2 + p - pq^{m-\frac{N-1}{2}} + p \\ &\geq q^{m-\frac{N-1}{2}} + 2q^{-\frac{N-1}{4}} - 2q^{m+1} + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} - 2 + 2p - pq^{m-\frac{N-1}{2}}, \end{aligned}$$

and the final RHS term is increasing in m as $q^{-\frac{N+1}{2}} \leq 2$. So, enough to check at $m = 0$. So, enough to show,

$$q^{-\frac{N-1}{2}} + 2q^{-\frac{N-1}{4}} - 2q + 4q^{\frac{N+1}{2}} - q^{\frac{N-1}{2}} - 2 + 2p - pq^{-\frac{N-1}{2}} \geq 2.$$

The LHS converges to $\sqrt{e} + 2e^{\frac{1}{4}} + 3\frac{1}{\sqrt{e}} - 4 > 2$. So, after some N LHS will be greater than 2. For some initial terms we have to check directly. Then simplifying for the expression of I, we get

$$||A||_M = 1 - q^{\frac{N-1}{2}} - q^{\frac{N+1}{2}} - q^{N-1} + 2q^N.$$

This completes the proof. □