

# ASYMPTOTIC ANALYSIS OF PERES' ALGORITHM FOR RANDOM NUMBER GENERATION

ZHAO GING LIM  AND CHEN-TUO LIAO

Division of Biometry, Institute of Agronomy, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd.,  
Taipei 106, Taiwan (ROC)

E-mail: [d02621201@ntu.edu.tw](mailto:d02621201@ntu.edu.tw); [ctliao@ntu.edu.tw](mailto:ctliao@ntu.edu.tw)

YI-CHING YAO

Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan (ROC)

E-mail: [yao@stat.sinica.edu.tw](mailto:yao@stat.sinica.edu.tw)

von Neumann [(1951). Various techniques used in connection with random digits. *National Bureau of Standards Applied Math Series* 12: 36–38] introduced a simple algorithm for generating independent unbiased random bits by tossing a (possibly) biased coin with unknown bias. While his algorithm fails to attain the entropy bound, Peres [(1992). Iterating von Neumann's procedure for extracting random bits. *The Annals of Statistics* 20(1): 590–597] showed that the entropy bound can be attained asymptotically by iterating von Neumann's algorithm. Let  $b(n, p)$  denote the expected number of unbiased bits generated when Peres' algorithm is applied to an input sequence consisting of the outcomes of  $n$  tosses of the coin with bias  $p$ . With  $p = 1/2$ , the coin is unbiased and the input sequence consists of  $n$  unbiased bits, so that  $n - b(n, 1/2)$  may be referred to as the cost incurred by Peres' algorithm when not knowing  $p = 1/2$ . We show that  $\lim_{n \rightarrow \infty} \log[n - b(n, 1/2)] / \log n = \theta = \log[(1 + \sqrt{5})/2]$  (where  $\log$  is the logarithm to base 2), which together with limited numerical results suggests that  $n - b(n, 1/2)$  may be a regularly varying sequence of index  $\theta$ . (A positive sequence  $\{L(n)\}$  is said to be regularly varying of index  $\theta$  if  $\lim_{n \rightarrow \infty} L(\lfloor \lambda n \rfloor) / L(n) = \lambda^\theta$  for all  $\lambda > 0$ , where  $\lfloor x \rfloor$  denotes the largest integer not exceeding  $x$ .) Some open problems on the asymptotic behavior of  $nh(p) - b(n, p)$  are briefly discussed where  $h(p) = -p \log p - (1 - p) \log(1 - p)$  denotes the Shannon entropy of a random bit with bias  $p$ .

**Keywords:** analysis of algorithms, elias' extractor, entropy, peres' extractor, regularly varying sequence, superadditivity, von neumann's extractor

## 1. INTRODUCTION AND RESULTS

In his seminal work [13], von Neumann introduced a simple algorithm  $\mathcal{A}_{\text{VN}}$  (also known as an extractor) for generating independent unbiased random bits by tossing a (possibly) biased coin with unknown bias. (A random bit is said to be unbiased if its value is 1 or 0 with equal probability.) Specifically, for  $i = 1, 2, \dots$ , let  $X_i \in \{0, 1\}$  denote the outcome of the  $i$ th toss of the coin, where 1 and 0 stand for heads and tails,

respectively. Assume that the input sequence  $(X_1, X_2, \dots)$  is independent and identically distributed (i.i.d.) with  $\mathbb{P}(X_i = 1) = p = 1 - \mathbb{P}(X_i = 0)$  where the bias  $p \in (0, 1)$  is unknown. (The coin is unbiased if  $p = 1/2$ .) The algorithm  $\mathcal{A}_N$  divides the  $X_i$ 's into pairs  $(X_1, X_2), (X_3, X_4), \dots$ , and discards those pairs of equal values. For each pair of unequal values,  $\mathcal{A}_N$  generates a bit equal to the first value of the pair, which is readily seen to be unbiased.

Let  $\mathcal{A}$  denote a generic algorithm that generates independent unbiased bits from the sequence  $(X_1, X_2, \dots)$ . Let  $\mathcal{A}(n)$  denote the set of unbiased bits generated by  $\mathcal{A}$  applied to  $(X_1, \dots, X_n)$ , the outcomes of the first  $n$  tosses. Denote by  $|\mathcal{A}(n)|$  the cardinality of  $\mathcal{A}(n)$ , which is an integer-valued random variable whose distribution depends on  $n$  and  $p$ . We say that  $\mathcal{A}$  is *nested* if  $\mathcal{A}(n_1) \subset \mathcal{A}(n_2)$  whenever  $n_1 < n_2$ , that is, the set of unbiased bits generated from  $(X_1, \dots, X_{n_1})$  is contained in the set of unbiased bits generated from  $(X_1, \dots, X_{n_1}, X_{n_1+1}, \dots, X_{n_2})$ . We write  $X \sim \text{binomial}(n, p)$  if a random variable  $X$  has the binomial distribution with parameters  $n$  and  $p$ . Then given  $|\mathcal{A}_N(n)| = k$ , the  $k$  bits generated by  $\mathcal{A}_N$  applied to  $(X_1, \dots, X_n)$  are (conditionally) independent unbiased. Moreover,  $|\mathcal{A}_N(n)| \sim \text{binomial}(\lfloor n/2 \rfloor, 2pq)$ , where  $q = 1 - p$  and  $\lfloor x \rfloor$  denotes the largest integer not exceeding  $x$ . When  $\mathcal{A}_N$  is applied to  $(X_1, \dots, X_n)$ , the expected number of unbiased bits generated per toss equals  $\mathbb{E}_p |\mathcal{A}_N(n)|/n = 2pq \lfloor n/2 \rfloor/n$ , which converges to  $pq$  as  $n \rightarrow \infty$ , where the subscript  $p$  in  $\mathbb{E}_p$  refers to the bias of each  $X_i$ . Note that  $pq$  is less than the entropy bound  $h(p) := -p \log p - q \log q$  (the Shannon entropy of each  $X_i$ ), where  $\log = \log_2$  (the logarithm to base 2). This indicates that  $\mathcal{A}_N$  does not make efficient use of information contained in  $X_1, X_2, \dots$ . It is also worth noting that  $\mathcal{A}_N$  is nested.

To improve the efficiency, Elias [2] presented a more sophisticated algorithm  $\mathcal{A}_E$  which generates unbiased bits from  $(X_1, \dots, X_n)$  by partitioning  $S_n = \{(x_1, \dots, x_n) : x_i \in \{0, 1\}\}$  (the set of all possible realizations of  $(X_1, \dots, X_n)$ ) into disjoint subsets  $S_{n,k} = \{(x_1, \dots, x_n) \in S_n : \sum_{i=1}^n x_i = k\}$ ,  $k = 0, 1, \dots, n$ . Write  $|S_{n,k}| = \binom{n}{k} = \sum_{\ell=0}^{\lfloor \log \binom{n}{k} \rfloor} c_\ell 2^\ell$  with  $c_\ell \in \{0, 1\}$  (binary representation of  $\binom{n}{k}$ ). Then, each  $S_{n,k}$  is further partitioned as  $S_{n,k} = \bigcup_{\{\ell: c_\ell=1\}} S_{n,k,\ell}$ , where  $|S_{n,k,\ell}| = 2^\ell$  for each  $\ell$  with  $c_\ell = 1$ . Specify an assignment of  $2^\ell$  distinct (output) sequences of  $\{0, 1\}^\ell$  to the  $2^\ell$  distinct sequences of  $S_{n,k,\ell}$ , so that if  $(X_1, \dots, X_n) \in S_{n,k,\ell}$ , then an output sequence of  $\ell$  bits is generated according to the assignment. While a naive implementation of Elias' algorithm requires an exponential memory size to make a table of assignment of output sequences, Ryabko and Matchikina [12] provided an efficient method to construct an assignment with much reduced memory size and running time. Note that  $\mathcal{A}_E$  is not nested. In fact, when  $\mathcal{A}_E$  is applied to  $(X_1, \dots, X_n)$ , all of  $X_1, \dots, X_n$  need to be observed before unbiased bits are generated. To show that  $\mathcal{A}_E$  attains the entropy bound asymptotically, Elias [2, Eq. (15)] proved that

$$\sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \log \binom{n}{k} - 3 \leq \mathbb{E}_p |\mathcal{A}_E(n)| \leq \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \log \binom{n}{k}. \tag{1.1}$$

Letting  $H(Z)$  denote the Shannon entropy of a random variable  $Z$  and noting that  $\sum_{i=1}^n X_i \sim \text{binomial}(n, p)$ , we have

$$\begin{aligned} nh(p) - H\left(\sum_{i=1}^n X_i\right) &= -np \log p - nq \log q + \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \log \left[\binom{n}{k} p^k q^{n-k}\right] \\ &= \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \log \binom{n}{k}, \end{aligned}$$

from which it follows that (1.1) is equivalent to

$$H\left(\sum_{i=1}^n X_i\right) \leq nh(p) - \mathbb{E}_p |\mathcal{A}_E(n)| \leq H\left(\sum_{i=1}^n X_i\right) + 3. \quad (1.2)$$

Since  $H(\sum_{i=1}^n X_i) = \frac{1}{2} \log n + \frac{1}{2} \log e + \log \sqrt{2\pi pq} + O(1/n)$  (cf. [3]), we have

$$nh(p) - \mathbb{E}_p |\mathcal{A}_E(n)| = \frac{1}{2} \log n + O(1). \quad (1.3)$$

Consequently,  $\lim_{n \rightarrow \infty} \mathbb{E}_p |\mathcal{A}_E(n)|/n = h(p)$ . Later, Pae and Loui [8] established the exact optimality of  $\mathcal{A}_E$  that for any algorithm  $\mathcal{A}$ ,  $\mathbb{E}_p |\mathcal{A}_E(n)| \geq \mathbb{E}_p |\mathcal{A}(n)|$  for all  $p \in (0, 1)$  and  $n \geq 1$ .

While  $\mathcal{A}_{\mathcal{N}}$  fails to attain the entropy bound asymptotically, Peres [9] pointed out that the entropy bound can be attained asymptotically by iterating  $\mathcal{A}_{\mathcal{N}}$ . To describe Peres' ingenious idea, we consider an infinite i.i.d. Bernoulli sequence  $\mathbf{X} = (X_1, X_2, \dots)$  with bias  $p$ , which is decomposed into three infinite Bernoulli sequences  $\psi_i(\mathbf{X})$ ,  $i = 1, 2, 3$ , as follows. First divide  $X_1, X_2, \dots$  into pairs  $(X_1, X_2), (X_3, X_4), \dots$ . The  $i$ th bit of  $\psi_1(\mathbf{X})$  is 1 or 0 according as the  $i$ th pair  $(X_{2i-1}, X_{2i})$  is of equal values or of unequal values. Then, separate those pairs of equal values from the other pairs of unequal values. The  $i$ th bit of  $\psi_2(\mathbf{X})$  is the common value of the  $i$ th pair of equal values. The  $i$ th bit of  $\psi_3(\mathbf{X})$  is the first value of the  $i$ th pair of unequal values. As an example, let  $\mathbf{X} = (0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, \dots)$ . Then,  $\psi_1(\mathbf{X}) = (0, 0, 1, 1, 0, 1, \dots)$ ,  $\psi_2(\mathbf{X}) = (1, 0, 0, \dots)$  and  $\psi_3(\mathbf{X}) = (0, 1, 1, \dots)$ . It is readily seen that (i)  $\psi_1(\mathbf{X})$ ,  $\psi_2(\mathbf{X})$  and  $\psi_3(\mathbf{X})$  are mutually independent, (ii)  $\psi_1(\mathbf{X})$ ,  $\psi_2(\mathbf{X})$  and  $\psi_3(\mathbf{X})$  are each an i.i.d. Bernoulli sequence with respective biases  $f_1(p) := p^2 + q^2$ ,  $f_2(p) := p^2/(p^2 + q^2)$  and  $1/2$ , (iii)  $\mathbf{X}$  can be recovered from  $\psi_1(\mathbf{X})$ ,  $\psi_2(\mathbf{X})$  and  $\psi_3(\mathbf{X})$ , implying that they together contain all information in  $\mathbf{X}$ . The first iteration of Peres' algorithm  $\mathcal{A}_P$  yields  $\psi_i(\mathbf{X})$ ,  $i = 1, 2, 3$ , where  $\psi_3(\mathbf{X})$  is precisely the output sequence generated by  $\mathcal{A}_{\mathcal{N}}$  applied to  $\mathbf{X}$ . On the second iteration of  $\mathcal{A}_P$ ,  $\psi_i(\mathbf{X})$ ,  $i = 1, 2$  are each decomposed into three i.i.d. Bernoulli sequences  $\psi_1(\psi_i(\mathbf{X}))$ ,  $\psi_2(\psi_i(\mathbf{X}))$  and  $\psi_3(\psi_i(\mathbf{X}))$  with respective biases  $f_1(f_i(p))$ ,  $f_2(f_i(p))$  and  $1/2$ . Thus, after 2 iterations, there are  $7 (= 2^3 - 1)$  Bernoulli sequences,  $3 (= 2^2 - 1)$  of which have bias  $1/2$ . More generally, after  $\nu$  iterations ( $\nu = 1, 2, \dots$ ), there are  $2^{\nu+1} - 1$  Bernoulli sequences,  $2^\nu - 1$  of which have bias  $1/2$ . We refer to the  $2^\nu - 1$  Bernoulli sequences having bias  $1/2$  as *unbiased* Bernoulli sequences, and refer to the other  $2^\nu$  Bernoulli sequences as *biased* Bernoulli sequences. Note that the  $2^{\nu+1} - 1$  Bernoulli sequences are all mutually independent, from which  $\mathbf{X}$  can be recovered.

We now consider the finite setting where only the first  $n$  terms of the infinite input sequence  $\mathbf{X}$  are available. Let  $(\mathbf{X})_n = (X_1, \dots, X_n)$ , the subsequence of  $\mathbf{X}$  consisting of the first  $n$  terms. Then,  $(\mathbf{X})_n$  induces the first  $n_i$  terms of  $\psi_i(\mathbf{X})$ ,  $i = 1, 2, 3$ , where  $n_1 = n_2 + n_3 = \lfloor n/2 \rfloor$ ,  $n_2 \sim \text{binomial}(\lfloor n/2 \rfloor, p^2 + q^2)$  and  $n_3 \sim \text{binomial}(\lfloor n/2 \rfloor, 2pq)$ . In shorthand notation,  $(\mathbf{X})_n$  induces  $(\psi_i(\mathbf{X}))_{n_i}$ ,  $i = 1, 2, 3$ . While the infinite sequences  $\psi_i(\mathbf{X})$ ,  $i = 1, 2, 3$  are mutually independent, the subsequences  $(\psi_i(\mathbf{X}))_{n_i}$ ,  $i = 1, 2, 3$  are no longer independent. Indeed, the numbers of 1's and 0's in  $(\psi_1(\mathbf{X}))_{n_1}$  are equal to  $n_2$  and  $n_3$ , respectively. It is readily seen that, given the value of  $n_3$ , the bits in  $(\psi_3(\mathbf{X}))_{n_3}$  are (conditionally) independent unbiased. In fact, given the values of the bits in  $(\psi_1(\mathbf{X}))_{n_1}$  and  $(\psi_2(\mathbf{X}))_{n_2}$ , the bits in  $(\psi_3(\mathbf{X}))_{n_3}$  remain (conditionally) independent unbiased. Furthermore, for even  $n$ ,  $(\mathbf{X})_n$  can be recovered from  $(\psi_i(\mathbf{X}))_{n_i}$ ,  $i = 1, 2, 3$ , but for odd  $n$ , the last term of  $(\mathbf{X})_n$  cannot be recovered, resulting in a loss of information. After  $\nu$  iterations ( $\nu = 1, 2, \dots$ ),  $(\mathbf{X})_n$  induces a (possibly empty) subsequence of each of the  $2^{\nu+1} - 1$  infinite Bernoulli sequences as decomposed from  $\mathbf{X}$ . Let  $\mathcal{A}_{P,\nu}(n)$  denote the set of all unbiased bits contained in the subsequences of those  $2^\nu - 1$  infinite unbiased Bernoulli sequences. Since

after  $\lfloor \log n \rfloor$  iterations, the longest biased subsequence has length 1, no more unbiased bits can be produced by further iteration. We have  $\mathcal{A}_{\mathcal{P},\nu}(n) = \mathcal{A}_{\mathcal{P},\lfloor \log n \rfloor}(n)$  for  $\nu \geq \lfloor \log n \rfloor$ . Let  $\mathcal{A}_{\mathcal{P}}(n) = \mathcal{A}_{\mathcal{P},\lfloor \log n \rfloor}(n)$ , the set of all unbiased bits generated by  $\mathcal{A}_{\mathcal{P}}$  applied to  $(\mathbf{X})_n$ . Consider again the example where  $\mathbf{X} = (0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, \dots)$ . For  $n = 12$ , we have  $\mathcal{A}_{\mathcal{P},1}(12) = \{0, 1, 1\}$ ,  $\mathcal{A}_{\mathcal{P},2}(12) = \{0, 1, 1, 0, 1\}$ , and  $\mathcal{A}_{\mathcal{P}}(12) = \mathcal{A}_{\mathcal{P},3}(12) = \{0, 1, 1, 0, 1, 0\}$ . It is shown in Peres [9] that (i) for each  $\nu$ , given  $|\mathcal{A}_{\mathcal{P},\nu}(n)| = k$ , the  $k$  bits in  $\mathcal{A}_{\mathcal{P},\nu}(n)$  are independent unbiased, (ii) the rates  $r_{\nu}(p) := \lim_{n \rightarrow \infty} \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)|/n$  satisfy  $r_1(p) = pq$  and the recursion

$$r_{\nu}(p) = pq + \frac{1}{2}r_{\nu-1}(p^2 + q^2) + \frac{1}{2}(p^2 + q^2)r_{\nu-1} \left( \frac{p^2}{p^2 + q^2} \right) \quad \text{for } \nu \geq 2, \tag{1.4}$$

and (iii)  $r_{\nu}(p)$  increases as  $\nu \rightarrow \infty$  to  $h(p)$  uniformly in  $p \in (0, 1)$ . As a consequence,  $\mathbb{E}_p |\mathcal{A}_{\mathcal{P}}(n)|/n \rightarrow h(p)$  as  $n \rightarrow \infty$ , showing that  $\mathcal{A}_{\mathcal{P}}$  attains the entropy bound asymptotically. Moreover,  $\mathcal{A}_{\mathcal{P}}$  is nested.

While  $\mathcal{A}_{\mathcal{E}}$  and  $\mathcal{A}_{\mathcal{P}}$  both attain the entropy bound asymptotically, (1.2) and (1.3) provide a precise (second-order) behavior of  $nh(p) - \mathbb{E}_p |\mathcal{A}_{\mathcal{E}}(n)|$ . In contrast, there is not much known about the behavior of  $nh(p) - \mathbb{E}_p |\mathcal{A}_{\mathcal{P}}(n)|$  for large  $n$ . In this regard, Pae [7] gave a formula to compute  $\mathbb{E}_p |\mathcal{A}_{\mathcal{P}}(n)|$ , which is not convenient for deriving the asymptotic behavior of  $nh(p) - \mathbb{E}_p |\mathcal{A}_{\mathcal{P}}(n)|$ . Recently, Prasitsupparote *et al.* [10] showed, based on some heuristics, that for  $p = 1/2$ ,

$$nh(p) - \mathbb{E}_p |\mathcal{A}_{\mathcal{P}}(n)| = n - \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P}}(n)| \geq n^{\log 3 - 1}. \tag{1.5}$$

To derive (1.5), they assumed, without rigorous justification, that

$$\frac{1}{n} \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)| \leq r_{\nu}(p) \quad \text{for } p \in (0, 1), n \geq 1, \nu \geq 1. \tag{1.6}$$

In this paper, we establish the following results.

PROPOSITION 1: Let  $a(n, p, \nu) = \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)|$ . Then, for  $p \in (0, 1)$  and  $\nu = 1, 2, \dots$ , the sequence  $(a(1, p, \nu), a(2, p, \nu), \dots)$  is superadditive, that is,  $a(n, p, \nu) + a(m, p, \nu) \leq a(n + m, p, \nu)$  for  $n, m \geq 1$ . Consequently,  $\lim_{n \rightarrow \infty} a(n, p, \nu)/n$  exists and is equal to  $\sup_{n \geq 1} a(n, p, \nu)/n$ . That is,

$$r_{\nu}(p) := \lim_{n \rightarrow \infty} \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)|/n = \sup_{n \geq 1} \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)|/n,$$

which implies (1.6).

PROPOSITION 2: For  $p = 1/2$ , let  $b(n) = \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P}}(n)|$ . Then,

(i) the  $b(n)$  satisfy  $b(0) = b(1) = 0$  and the recursion

$$b(n) = \left\lfloor \frac{n}{2} \right\rfloor / 2 + b \left( \left\lfloor \frac{n}{2} \right\rfloor \right) + \mathbb{E} b(B_{\lfloor n/2 \rfloor, 1/2}) \quad \text{for } n = 2, 3, \dots, \tag{1.7}$$

where  $B_{n,p}$  denotes a binomial( $n, p$ ) random variable;

(ii)

$$\lim_{n \rightarrow \infty} \frac{\log(n - b(n))}{\log n} = \log \left( \frac{1 + \sqrt{5}}{2} \right).$$

The next section contains the proofs of Propositions 1 and 2. In addition, for completeness, a rigorous proof of (1.5) is also given, which is needed for the proof of Proposition 2(ii). Section 3 presents the numerical results, and Section 4 concludes the paper with some open problems.

We close this section with a few remarks.

REMARK 1: *von Neumann's algorithm has been used to remove bias in the output of a true random number generator from a variety of physical devices (see, e.g., [14]). It has applications in cryptography to generate random cryptographic keys for secure data transmission.*

REMARK 2: *The so-called Bernoulli factory refers to the problem of using a  $p$ -coin (coin with probability of heads  $p$ ) to simulate an  $f(p)$ -coin (coin with probability of heads  $f(p)$ ) where (unknown)  $p$  is known only to belong to a given subset  $S$  of  $(0, 1)$  and the function  $f : S \rightarrow [0, 1]$  is known. The problem considered in the present paper deals with the special case that  $S = (0, 1)$  and  $f(p) = 1/2$  for all  $p$ . Keane and O'Brien [5] obtained necessary and sufficient conditions on  $f$  under which there exists an algorithm to simulate an  $f(p)$ -coin using a  $p$ -coin. See also Nacu and Peres [6] for related results.*

REMARK 3: *While  $\mathcal{A}_E$  generates more unbiased bits than  $\mathcal{A}_P$ ,  $\mathcal{A}_P$  is much simpler to implement. To implement Elias' algorithm efficiently, Ryabko and Matchikina [12] made use of the fast method of coding (proposed by Ryabko [11]) with a much reduced encoding time of  $O(\log^3 n \log \log n)$  bit-operations per letter. Prasitsupparote et al. [10] conducted an extensive numerical study of  $\mathcal{A}_P$  and  $\mathcal{A}_E$  (with the method of Ryabko and Matchikina [12]) by comparing their memory and running time requirements. Their empirical results suggest that  $\mathcal{A}_P$  would be superior to  $\mathcal{A}_E$  in practical applications.*

## 2. PROOFS OF PROPOSITIONS 1 AND 2 AND (1.5)

PROOF OF PROPOSITION 1: Recall that when  $\mathcal{A}_P$  is applied to  $(X_1, \dots, X_n)$ , 3 subsequences  $(\psi_i(\mathbf{X}))_{n_i}$ ,  $i = 1, 2, 3$ , are induced where  $n_1 = n_2 + n_3 = \lfloor n/2 \rfloor$  and  $n_3 \sim \text{binomial}(\lfloor n/2 \rfloor, 2pq)$ . It follows that  $\mathcal{A}_{P,\nu}(n)$ , the set of unbiased bits generated after  $\nu$  iterations, is the union of 3 disjoint subsets  $\mathcal{S}_i$ ,  $i = 1, 2, 3$ , where  $\mathcal{S}_3 = (\psi_3(\mathbf{X}))_{n_3}$ , and  $\mathcal{S}_i$ ,  $i = 1, 2$ , is the set of unbiased bits generated when  $\mathcal{A}_P$  is applied to  $(\psi_i(\mathbf{X}))_{n_i}$  with  $\nu - 1$  iterations. We have

$$a(n, p, \nu) = \mathbb{E}_p |\mathcal{A}_{P,\nu}(n)| = \mathbb{E}_p |\mathcal{S}_1| + \mathbb{E}_p |\mathcal{S}_2| + 2pq \lfloor n/2 \rfloor. \quad (2.1)$$

Noting that  $(\psi_1(\mathbf{X}))_{n_1}$  is a sequence of  $n_1 = \lfloor n/2 \rfloor$  i.i.d. Bernoulli random variables with bias  $f_1(p)$ , we have

$$\mathbb{E}_p |\mathcal{S}_1| = \mathbb{E}_{f_1(p)} |\mathcal{A}_{P,\nu-1}(\lfloor n/2 \rfloor)| = a(\lfloor n/2 \rfloor, f_1(p), \nu - 1). \quad (2.2)$$

Similarly, conditioning on  $n_2$ ,  $(\psi_2(\mathbf{X}))_{n_2}$  is a sequence of  $n_2$  i.i.d. Bernoulli random variables with bias  $f_2(p)$ , so that the conditional expectation of  $|\mathcal{S}_2|$  given  $n_2$  equals

$\mathbb{E}_{f_2(p)} |\mathcal{A}_{P, \nu-1}(n_2)| = a(n_2, f_2(p), \nu - 1)$ . Since  $n_2 \sim \text{binomial}(\lfloor n/2 \rfloor, 1 - 2pq)$ , we have

$$\mathbb{E}_p |\mathcal{S}_2| = \mathbb{E} a(B_{\lfloor n/2 \rfloor, 1-2pq}, f_2(p), \nu - 1), \tag{2.3}$$

where the expectation operator  $\mathbb{E}$  on the right-hand side is on  $B_{\lfloor n/2 \rfloor, 1-2pq}$  (a  $\text{binomial}(\lfloor n/2 \rfloor, 1 - 2pq)$  random variable). By (2.1), (2.2) and (2.3),

$$a(n, p, \nu) = a(\lfloor n/2 \rfloor, f_1(p), \nu - 1) + \mathbb{E} a(B_{\lfloor n/2 \rfloor, 1-2pq}, f_2(p), \nu - 1) + 2pq \lfloor n/2 \rfloor. \tag{2.4}$$

We now prove by induction on  $\nu$  that

$$a(n, p, \nu) + a(m, p, \nu) \leq a(n + m, p, \nu). \tag{2.5}$$

For  $\nu = 1$ ,  $a(n, p, 1) = 2pq \lfloor n/2 \rfloor$ . Since  $\lfloor n/2 \rfloor + \lfloor m/2 \rfloor \leq \lfloor (n + m)/2 \rfloor$  for  $n, m \geq 1$ , we have  $a(n, p, 1) + a(m, p, 1) \leq a(n + m, p, 1)$ , implying that (2.5) holds for  $\nu = 1$ . Suppose that for an integer  $k > 0$ , (2.5) holds for all  $n, m \geq 1$ , all  $p \in (0, 1)$ , and  $\nu = k$ . We need to show that (2.5) holds for  $n, m \geq 1$ ,  $p \in (0, 1)$  and  $\nu = k + 1$ . By the induction hypothesis,

$$\begin{aligned} a(\lfloor n/2 \rfloor, f_1(p), k) + a(\lfloor m/2 \rfloor, f_1(p), k) &\leq a(\lfloor n/2 \rfloor + \lfloor m/2 \rfloor, f_1(p), k) \\ &\leq a(\lfloor (n + m)/2 \rfloor, f_1(p), k), \end{aligned} \tag{2.6}$$

where the second inequality follows from the fact that  $a(n, p, \nu)$  is non-decreasing in  $n$ . Let  $U$  and  $V$  be independent random variables with  $U \sim \text{binomial}(\lfloor n/2 \rfloor, 1 - 2pq)$  and  $V \sim \text{binomial}(\lfloor m/2 \rfloor, 1 - 2pq)$ . Then,  $U + V \sim \text{binomial}(\lfloor n/2 \rfloor + \lfloor m/2 \rfloor, 1 - 2pq)$ . If  $n$  and  $m$  are both odd, let  $W$  be independent of  $U$  and  $V$  with  $W \sim \text{binomial}(1, 1 - 2pq)$ . If at least one of  $n$  and  $m$  is even, let  $W$  be identically 0. Then,  $U + V + W \sim \text{binomial}(\lfloor (n + m)/2 \rfloor, 1 - 2pq)$ . We have by the induction hypothesis that

$$\begin{aligned} &\mathbb{E} a(B_{\lfloor n/2 \rfloor, 1-2pq}, f_2(p), k) + \mathbb{E} a(B_{\lfloor m/2 \rfloor, 1-2pq}, f_2(p), k) \\ &= \mathbb{E} \{ a(U, f_2(p), k) + a(V, f_2(p), k) \} \\ &\leq \mathbb{E} a(U + V, f_2(p), k) \\ &\leq \mathbb{E} a(U + V + W, f_2(p), k) \\ &= \mathbb{E} a(B_{\lfloor (n+m)/2 \rfloor}, f_2(p), k). \end{aligned} \tag{2.7}$$

Moreover,

$$2pq \lfloor n/2 \rfloor + 2pq \lfloor m/2 \rfloor \leq 2pq \lfloor (n + m)/2 \rfloor. \tag{2.8}$$

By (2.4) and (2.6)–(2.8),

$$a(n, p, k + 1) + a(m, p, k + 1) \leq a(n + m, p, k + 1),$$

showing that (2.5) holds for  $n, m \geq 1$ ,  $p \in (0, 1)$  and  $\nu = k + 1$ . The proof is complete. ■

PROOF OF (1.5): The following argument is taken from the proof of Theorem 1 in Prasitsupparote [10]. With  $p = 1/2$ , we have  $r_1(1/2) = pq = \frac{1}{4}$  and, by (1.4)

$$r_\nu(1/2) = \frac{1}{4} + \frac{3}{4}r_{\nu-1}(1/2) \quad \text{for } \nu \geq 2,$$

from which it follows that  $r_\nu(1/2) = 1 - (\frac{3}{4})^\nu$ ,  $\nu \geq 1$ . By Proposition 1,

$$1 - (\frac{3}{4})^\nu = r_\nu(1/2) \geq \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P},\nu}(n)|/n,$$

so that with  $\nu = \lfloor \log n \rfloor$  and  $b(n) = \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P}}(n)|$ , we have

$$1 - (\frac{3}{4})^{\lfloor \log n \rfloor} \geq \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P},\lfloor \log n \rfloor}(n)|/n = \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P}}(n)|/n = b(n)/n,$$

implying that

$$n - b(n) \geq n(\frac{3}{4})^{\lfloor \log n \rfloor} \geq n(\frac{3}{4})^{\log n} = n^{\log 3 - 1},$$

proving (1.5). ■

PROOF OF PROPOSITION 2(i): For  $p = 1/2$ ,  $f_1(1/2) = f_2(1/2) = 1/2$  and  $B_{\lfloor n/2 \rfloor, 1-2pq} = B_{\lfloor n/2 \rfloor, 1/2}$ . Letting  $a(n, p, \nu) = \mathbb{E}_p |\mathcal{A}_{\mathcal{P},\nu}(n)|$ , we have by (2.4) that

$$a(n, 1/2, \nu) = a(\lfloor n/2 \rfloor, 1/2, \nu - 1) + \mathbb{E} a(B_{\lfloor n/2 \rfloor, 1/2}, 1/2, \nu - 1) + \lfloor n/2 \rfloor / 2. \tag{2.9}$$

Recall that  $\mathcal{A}_{\mathcal{P}}(n) = \mathcal{A}_{\mathcal{P},\nu}(n)$  for  $\nu \geq \lfloor \log n \rfloor$ . By (2.9),

$$\begin{aligned} b(n) &= \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P}}(n)| = \mathbb{E}_{1/2} |\mathcal{A}_{\mathcal{P},\lfloor \log n \rfloor}(n)| \\ &= a(n, 1/2, \lfloor \log n \rfloor) \\ &= a(\lfloor n/2 \rfloor, 1/2, \lfloor \log n \rfloor - 1) + \mathbb{E} a(B_{\lfloor n/2 \rfloor, 1/2}, 1/2, \lfloor \log n \rfloor - 1) + \lfloor n/2 \rfloor / 2 \\ &= b(\lfloor n/2 \rfloor) + \mathbb{E} b(B_{\lfloor n/2 \rfloor, 1/2}) + \lfloor n/2 \rfloor / 2, \end{aligned}$$

proving (1.7). ■

To prove Proposition 2(ii), we need the following lemmas. Proposition 2(ii) follows immediately from Lemmas 4 and 5 below. For the rest of this section, to simplify notation, we write  $B_n = B_{n,1/2}$  for a binomial( $n, 1/2$ ) random variable. Let  $g(n) = n - b(n) \geq 0$ , for  $n = 0, 1, \dots$ . We have  $g(0) = 0$ ,  $g(1) = 1$ , and by Proposition 2(i), for even  $n \geq 0$ ,

$$\begin{aligned} g(n) &= n - b(n) = n - \left[ \frac{n}{4} + b\left(\frac{n}{2}\right) + \mathbb{E} b(B_{n/2}) \right] \\ &= \left[ \frac{n}{2} - b\left(\frac{n}{2}\right) \right] + \mathbb{E} [B_{n/2} - b(B_{n/2})] \\ &= g\left(\frac{n}{2}\right) + \mathbb{E} g(B_{n/2}), \end{aligned}$$

and for odd  $n \geq 1$ ,

$$\begin{aligned} g(n) &= n - b(n) = n - \left[ \frac{n-1}{4} + b\left(\frac{n-1}{2}\right) + \mathbb{E} b(B_{(n-1)/2}) \right] \\ &= 1 + \left[ \frac{n-1}{2} - b\left(\frac{n-1}{2}\right) \right] + \mathbb{E} [B_{(n-1)/2} - b(B_{(n-1)/2})] \\ &= 1 + g\left(\frac{n-1}{2}\right) + \mathbb{E} g(B_{(n-1)/2}). \end{aligned}$$

So, for  $n \geq 0$ ,

$$g(n) = g\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + \mathbb{E}g(B_{\lfloor \frac{n}{2} \rfloor}) + \mathbf{1}_{\{n \text{ is odd}\}}, \tag{2.10}$$

where  $\mathbf{1}$  denotes the indicator function.

LEMMA 1: For  $\delta \in (0, 1)$ , we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(B_n > \frac{n}{2}(1 + \delta)\right) = -\frac{1}{2}[(1 - \delta) \log(1 - \delta) + (1 + \delta) \log(1 + \delta)] < 0.$$

LEMMA 2: If  $f(0) \leq f(1) \leq \dots \leq f(n + 1)$ , then  $\mathbb{E}f(B_{n+1}) \geq \mathbb{E}f(B_n)$ .

Lemma 1 is a standard result in large deviation theory; see, for example, [4, pp. 539–540]. Lemma 2 follows from the fact that  $B_n$  is stochastically smaller than  $B_{n+1}$ . By (2.10),

$$g(0) = 0, \quad g(1) = 1, \quad g(2) = \frac{3}{2}, \quad g(3) = \frac{5}{2}, \quad g(4) = \frac{19}{8} < g(3), \tag{2.11}$$

so that  $g(n)$  is not non-decreasing. Lemma 3 below constructs two non-decreasing sequences  $G$  and  $H$  that are closely related to  $g$  and satisfy  $0 \leq H(n) \leq g(n) \leq G(n) \leq n$ .

LEMMA 3: Let  $G(n)$  and  $H(n)$ ,  $n = 0, 1, \dots$  be defined by

$$\begin{aligned} G(n) &= g(n) && \text{for } n = 0, 1, 2, 3, \\ H(n) &= g(n) && \text{for } n = 0, 1, \end{aligned}$$

and recursively

$$G(n) = G\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + \mathbb{E}G(B_{\lfloor \frac{n}{2} \rfloor}) + 1 \quad \text{for } n \geq 4, \tag{2.12}$$

$$H(n) = H\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + \mathbb{E}H(B_{\lfloor \frac{n}{2} \rfloor}) \quad \text{for } n \geq 2. \tag{2.13}$$

Then, (i)  $G$  is non-decreasing and  $g(n) \leq G(n) \leq n$  for all  $n$ , and (ii)  $H$  is non-decreasing and  $g(n) \geq H(n) \geq 0$  for all  $n$ .

PROOF: In view of (2.10) and (2.12), it is easily shown by induction that  $g(n) \leq G(n)$  for all  $n$ . By (2.11),  $G(n) = g(n) \leq n$  and  $G(n)$  is non-decreasing for  $n \leq 3$ . For  $n \geq 4$ , if  $G(\ell) \leq \ell$  for all  $\ell < n$ , then

$$\begin{aligned} G(n) &= G\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + \mathbb{E}G(B_{\lfloor \frac{n}{2} \rfloor}) + 1 \\ &\leq \left\lfloor \frac{n}{2} \right\rfloor + \mathbb{E}B_{\lfloor \frac{n}{2} \rfloor} + 1 \\ &\leq \frac{3}{4}n + 1 \leq n. \end{aligned}$$

It follows by induction that  $G(n) \leq n$  for all  $n$ . That  $G(n)$  is non-decreasing in  $n$  also follows by induction and Lemma 2. This proves part (i). In view of (2.10) and (2.13), part (ii) can be proved similarly. ■



LEMMA 4: For each  $\delta \in (0, 1)$ , there exists an  $N \geq 4$  and a non-decreasing sequence  $(G'(0), G'(1), \dots)$  such that  $G'(n) \geq g(n)$  for all  $n$  and

$$G'(n) = G' \left( \left\lfloor \frac{n}{c} \right\rfloor \right) + \frac{4}{c^2} G' \left( \left\lfloor \frac{n}{c^2} \right\rfloor \right), \quad \text{for all } n \geq N,$$

where  $c = c(\delta) = 2/\sqrt{1 + \delta}$ . Moreover,

$$\limsup_{n \rightarrow \infty} \frac{\log g(n)}{\log n} \leq \limsup_{n \rightarrow \infty} \frac{\log G'(n)}{\log n} \leq \frac{1}{\log c} \log \left( \frac{1}{2} + \sqrt{\frac{4}{c^2} + \frac{1}{4}} \right).$$

Consequently, letting  $\delta \rightarrow 0$  so that  $c = c(\delta) \rightarrow 2$ , we have

$$\limsup_{n \rightarrow \infty} \frac{\log g(n)}{\log n} \leq \log \left( \frac{1 + \sqrt{5}}{2} \right).$$

PROOF: Let  $\delta \in (0, 1)$  be fixed. Let  $G$  be defined as in Lemma 3, so that  $G$  is non-decreasing and  $0 \leq g(n) \leq G(n) \leq n$  for all  $n$ . We have

$$\begin{aligned} \mathbb{E} G(B_{\lfloor n/2 \rfloor}) + 1 &\leq 1 + G \left( \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) \mathbb{P} \left( B_{\lfloor n/2 \rfloor} \leq \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) \\ &\quad + G \left( \left\lfloor \frac{n}{2} \right\rfloor \right) \mathbb{P} \left( B_{\lfloor n/2 \rfloor} > \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) \\ &\leq 1 + G \left( \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) + \left\lfloor \frac{n}{2} \right\rfloor \mathbb{P} \left( B_{\lfloor n/2 \rfloor} > \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right). \end{aligned} \tag{2.14}$$

By (1.5),  $G(\lfloor (\lfloor n/2 \rfloor / 2)(1 + \delta) \rfloor) \geq g(\lfloor (\lfloor n/2 \rfloor / 2)(1 + \delta) \rfloor) \rightarrow \infty$  as  $n \rightarrow \infty$ , and by Lemma 1,

$$\lim_{n \rightarrow \infty} \left\lfloor \frac{n}{2} \right\rfloor \mathbb{P} \left( B_{\lfloor n/2 \rfloor} > \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) = 0,$$

so that by (2.14), there is a (large)  $N \geq 4$  such that

$$\mathbb{E} G(B_{\lfloor n/2 \rfloor}) + 1 \leq (1 + \delta) G \left( \left\lfloor \frac{n}{2} \right\rfloor \left( \frac{1 + \delta}{2} \right) \right) \quad \text{for all } n \geq N. \tag{2.15}$$

Letting  $c = 2/\sqrt{1 + \delta}$ , we have by (2.12) and (2.15) that for all  $n \geq N (\geq 4)$ ,

$$\begin{aligned} G(n) &= G \left( \left\lfloor \frac{n}{2} \right\rfloor \right) + \mathbb{E} G(B_{\lfloor n/2 \rfloor}) + 1 \\ &\leq G \left( \left\lfloor \frac{n}{c} \right\rfloor \right) + \frac{4}{c^2} G \left( \left\lfloor \frac{n}{c^2} \right\rfloor \right). \end{aligned} \tag{2.16}$$

Define  $G'(n)$ ,  $n = 0, 1, \dots$  by  $G'(n) = G(n)$  for  $n < N$  and recursively

$$G'(n) = G' \left( \left\lfloor \frac{n}{c} \right\rfloor \right) + \frac{4}{c^2} G' \left( \left\lfloor \frac{n}{c^2} \right\rfloor \right) \quad \text{for } n \geq N. \tag{2.17}$$

(Note that for  $c > \sqrt{2}$  and  $n \geq N \geq 4$ ,  $\lfloor n/c^2 \rfloor \leq \lfloor n/c \rfloor \leq n - 1$ , so  $G'$  is well-defined.) Since  $G(n)$  is non-decreasing and  $G(n) = G'(n)$  for all  $n < N$ , we have by (2.16), (2.17)

and induction that  $G'(n) \geq G(n) (\geq g(n))$  for all  $n$ . To show that  $G'(n)$  is non-decreasing, note that  $G'(N) \geq G(N) \geq G(N - 1) = G'(N - 1)$ . Since  $G'(0) \leq G'(1) \leq \dots \leq G'(N)$ , it follows by (2.17) and induction that  $G'(n) \leq G'(n + 1)$  for all  $n \geq N$ .

It remains to prove that

$$\limsup_{n \rightarrow \infty} \frac{\log G'(n)}{\log n} \leq \frac{1}{\log c} \log \left( \frac{1}{2} + \sqrt{\frac{4}{c^2} + \frac{1}{4}} \right). \tag{2.18}$$

Let  $\ell_k = \lfloor c^k N \rfloor$ ,  $k = 0, 1, \dots$ . Let  $x_0 = G'(\ell_0)$ ,  $x_1 = G'(\ell_1)$ , and

$$x_k = x_{k-1} + \frac{4}{c^2} x_{k-2}, \quad k = 2, 3, \dots \tag{2.19}$$

By (2.17) and monotonicity of  $G'$ , we have for  $k \geq 2$

$$\begin{aligned} G'(\ell_k) &= G'(\lfloor c^k N \rfloor) = G' \left( \left\lfloor \frac{\lfloor c^k N \rfloor}{c} \right\rfloor \right) + \frac{4}{c^2} G' \left( \left\lfloor \frac{\lfloor c^k N \rfloor}{c^2} \right\rfloor \right) \\ &\leq G'(\lfloor c^{k-1} N \rfloor) + \frac{4}{c^2} G'(\lfloor c^{k-2} N \rfloor) \\ &= G'(\ell_{k-1}) + \frac{4}{c^2} G'(\ell_{k-2}). \end{aligned} \tag{2.20}$$

Since  $x_k = G'(\ell_k)$  for  $k = 0, 1$ , it follows by (2.19), (2.20) and induction that

$$G'(\ell_k) \leq x_k \quad \text{for all } k \geq 0. \tag{2.21}$$

Since  $x_k$  satisfies the difference equation (2.19), we have

$$x_k = \alpha_1 \lambda_1^k + \alpha_2 \lambda_2^k, \quad k = 0, 1, \dots$$

where

$$\begin{aligned} \lambda_1 &= \frac{1}{2}(1 + \gamma), & \lambda_2 &= \frac{1}{2}(1 - \gamma) \\ \alpha_1 &= \frac{1}{\gamma} \left( \frac{1}{2}(\gamma - 1)x_0 + x_1 \right), & \alpha_2 &= \frac{1}{\gamma} \left( \frac{1}{2}(\gamma + 1)x_0 - x_1 \right) \end{aligned}$$

and  $\gamma = \sqrt{1 + 16/c^2}$ . Noting that  $-1 < \lambda_2 < 0 < 1 < \lambda_1$  (since  $\sqrt{2} < c < 2$ ) and  $\alpha_1 > 0$ , it follows that

$$\lim_{k \rightarrow \infty} \frac{\log x_k}{k} = \log \lambda_1 = \log \left( \frac{1}{2} + \sqrt{\frac{4}{c^2} + \frac{1}{4}} \right). \tag{2.22}$$

By (2.21) and (2.22),

$$\limsup_{k \rightarrow \infty} \frac{\log G'(\ell_k)}{\log \ell_k} \leq \limsup_{k \rightarrow \infty} \frac{\log x_k}{\log \lfloor c^k N \rfloor} = \frac{\log \lambda_1}{\log c}.$$

Since  $G'$  is non-decreasing, for each  $n \geq 1$ , let  $k = k(n)$  be such that  $\ell_k \leq n < \ell_{k+1}$ , so that

$$\frac{\log G'(n)}{\log n} \leq \frac{\log G'(\ell_{k+1})}{\log \ell_k},$$

implying that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{\log G'(n)}{\log n} &\leq \limsup_{k \rightarrow \infty} \frac{\log G'(\ell_{k+1})}{\log \ell_k} \\ &= \limsup_{k \rightarrow \infty} \frac{\log G'(\ell_{k+1}) \log \ell_{k+1}}{\log \ell_{k+1} \log \ell_k} \\ &\leq \frac{\log \lambda_1}{\log c} = \frac{1}{\log c} \log \left( \frac{1}{2} + \sqrt{\frac{4}{c^2} + \frac{1}{4}} \right), \end{aligned}$$

proving (2.18). The proof is complete. ■

LEMMA 5: For each  $\delta \in (0, 1)$ , there exists an  $N \geq 4$  and a non-decreasing sequence  $(H'(0), H'(1), \dots)$  such that  $0 \leq H'(n) \leq g(n)$  for all  $n$  and

$$H'(n) = H' \left( \left\lceil \frac{n}{d} \right\rceil \right) + \frac{4}{d^2} H' \left( \left\lceil \frac{n}{d^2} \right\rceil \right), \quad \text{for all } n \geq N,$$

where  $d = d(\delta) = 2 + \delta$  and  $\lceil x \rceil$  denotes the smallest integer not less than  $x$ . Moreover,

$$\liminf_{n \rightarrow \infty} \frac{\log g(n)}{\log n} \geq \liminf_{n \rightarrow \infty} \frac{\log H'(n)}{\log n} \geq \frac{1}{\log d} \log \left( \frac{1}{2} + \sqrt{\frac{4}{d^2} + \frac{1}{4}} \right).$$

Consequently, letting  $\delta \rightarrow 0$  so that  $d = d(\delta) \rightarrow 2$ , we have

$$\liminf_{n \rightarrow \infty} \frac{\log g(n)}{\log n} \geq \log \left( \frac{1 + \sqrt{5}}{2} \right).$$

PROOF: The following proof is similar to that of Lemma 4. Let  $\delta \in (0, 1)$  be fixed. Let  $H$  be defined as in Lemma 3, so that  $H$  is non-decreasing and  $0 \leq H(n) \leq g(n)$  for all  $n$ . Also  $H(0) = g(0) = 0$ ,  $H(1) = g(1) = 1$ .

For  $d = 2 + \delta > 2$ , by the law of large numbers,  $\mathbb{P}(B_{\lfloor n/2 \rfloor} < \lceil \frac{n}{d^2} \rceil) \rightarrow 0$  as  $n \rightarrow \infty$ . So there exists an  $N \geq 4$  such that for all  $n \geq N$ ,

$$H \left( \left\lfloor \frac{n}{2} \right\rfloor \right) \geq H \left( \left\lceil \frac{n}{d} \right\rceil \right) \quad \text{and} \quad \mathbb{P} \left( B_{\lfloor n/2 \rfloor} \geq \left\lceil \frac{n}{d^2} \right\rceil \right) \geq \frac{4}{d^2}.$$

By (2.13), for  $n \geq N \geq 4$ ,

$$\begin{aligned} H(n) &= H \left( \left\lfloor \frac{n}{2} \right\rfloor \right) + \mathbb{E} H(B_{\lfloor n/2 \rfloor}) \\ &\geq H \left( \left\lceil \frac{n}{d} \right\rceil \right) + \mathbb{P} \left( B_{\lfloor n/2 \rfloor} \geq \left\lceil \frac{n}{d^2} \right\rceil \right) H \left( \left\lceil \frac{n}{d^2} \right\rceil \right) \\ &\geq H \left( \left\lceil \frac{n}{d} \right\rceil \right) + \frac{4}{d^2} H \left( \left\lceil \frac{n}{d^2} \right\rceil \right). \end{aligned} \tag{2.23}$$

Define  $H'(n)$ ,  $n = 0, 1, \dots$  by  $H'(0) = 0$ ,  $H'(1) = \dots = H'(N - 1) = 1$  and recursively

$$H'(n) = H' \left( \left\lceil \frac{n}{d} \right\rceil \right) + \frac{4}{d^2} H' \left( \left\lceil \frac{n}{d^2} \right\rceil \right) \quad \text{for } n \geq N. \tag{2.24}$$

(Note that  $\lceil n/d^2 \rceil \leq \lceil n/d \rceil \leq n - 1$  for  $n \geq N \geq 4$ , so that the recursion is well-defined.) Since by (2.24),  $H'(N) = H'(\lceil N/d \rceil) + (4/d^2)H'(\lceil N/d^2 \rceil) = 1 + 4/d^2 > 1$ , we have  $H'(0) <$

$H'(1) = \dots = H'(N - 1) < H'(N)$ . It follows by (2.24) and induction that  $H'$  is a non-decreasing sequence. Since  $H(n) \geq H'(n)$  for all  $n < N$ , we have by (2.23), (2.24) and induction that  $H'(n) \leq H(n)$  for all  $n$ .

It remains to prove that

$$\liminf_{n \rightarrow \infty} \frac{\log H'(n)}{\log n} \geq \frac{1}{\log d} \log \left( \frac{1}{2} + \sqrt{\frac{4}{d^2} + \frac{1}{4}} \right). \tag{2.25}$$

Let  $\ell_k = \lceil d^k N \rceil$ ,  $k = 0, 1, \dots$ . Let  $x_0 = H'(\ell_0)$ ,  $x_1 = H'(\ell_1)$ , and

$$x_k = x_{k-1} + \frac{4}{d^2} x_{k-2}, \quad k = 2, 3, \dots \tag{2.26}$$

By (2.24) and monotonicity of  $H'$ , we have for  $k \geq 2$

$$\begin{aligned} H'(\ell_k) &= H'(\lceil d^k N \rceil) = H' \left( \left\lceil \frac{\lceil d^k N \rceil}{d} \right\rceil \right) + \frac{4}{d^2} H' \left( \left\lceil \frac{\lceil d^k N \rceil}{d^2} \right\rceil \right) \\ &\geq H'(\lceil d^{k-1} N \rceil) + \frac{4}{d^2} H'(\lceil d^{k-2} N \rceil) \\ &= H'(\ell_{k-1}) + \frac{4}{d^2} H'(\ell_{k-2}). \end{aligned} \tag{2.27}$$

Since  $x_k = H'(\ell_k)$  for  $k = 0, 1$ , it follows by (2.26), (2.27) and induction that

$$H'(\ell_k) \geq x_k \quad \text{for all } k \geq 0. \tag{2.28}$$

Note that the difference equation (2.26) is the same as (2.19) with  $c$  replaced by  $d$ . Solving (2.26) yields (cf. (2.22))

$$\lim_{k \rightarrow \infty} \frac{\log x_k}{k} = \log \left( \frac{1}{2} + \sqrt{\frac{4}{d^2} + \frac{1}{4}} \right).$$

By (2.28),

$$\liminf_{k \rightarrow \infty} \frac{\log H'(\ell_k)}{\log \ell_k} \geq \liminf_{k \rightarrow \infty} \frac{\log x_k}{\log \lceil d^k N \rceil} = \frac{1}{\log d} \log \left( \frac{1}{2} + \sqrt{\frac{4}{d^2} + \frac{1}{4}} \right).$$

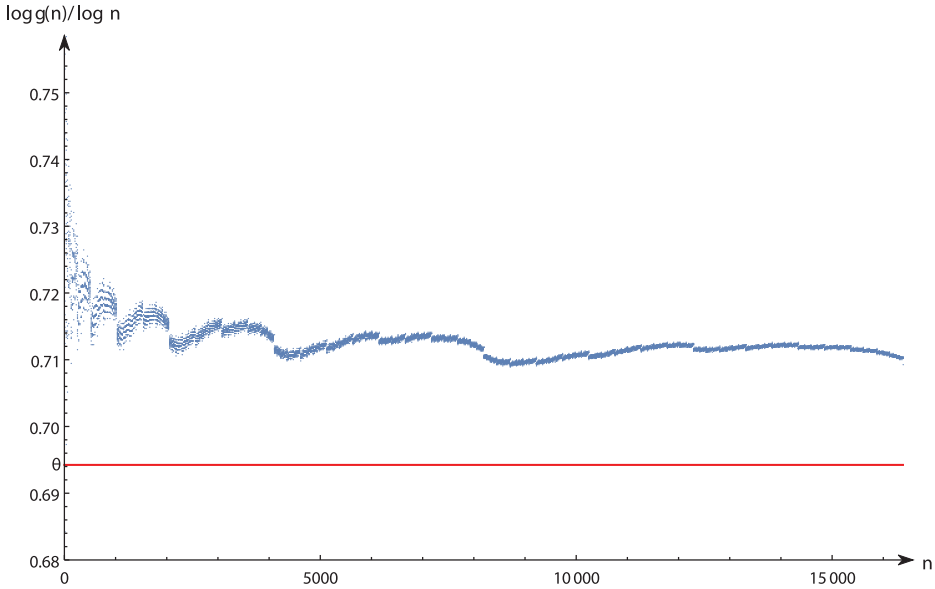
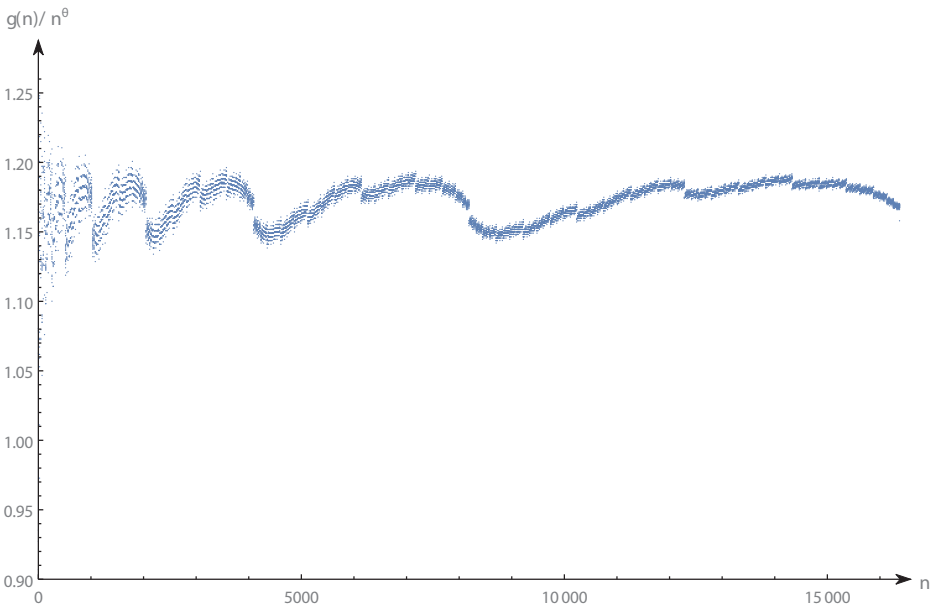
Since  $H'$  is non-decreasing, for each  $n \geq 1$ , let  $k = k(n)$  be such that  $\ell_k \leq n < \ell_{k+1}$ , so that

$$\frac{\log H'(n)}{\log n} \geq \frac{\log H'(\ell_k)}{\log \ell_{k+1}},$$

implying that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{\log H'(n)}{\log n} &\geq \liminf_{k \rightarrow \infty} \frac{\log H'(\ell_k)}{\log \ell_{k+1}} \\ &= \liminf_{k \rightarrow \infty} \frac{\log H'(\ell_k)}{\log \ell_k} \frac{\log \ell_k}{\log \ell_{k+1}} \\ &\geq \frac{1}{\log d} \log \left( \frac{1}{2} + \sqrt{\frac{4}{d^2} + \frac{1}{4}} \right), \end{aligned}$$

proving (2.25). The proof is complete. ■

FIGURE 1. Plot of  $\log g(n)/\log n$  versus  $n$ .FIGURE 2. Plot of  $g(n)/n^\theta$  versus  $n$  with  $\theta = \log[(1 + \sqrt{5})/2]$ .

### 3. NUMERICAL RESULTS

Recall that  $g(n) = n - b(n) = n - \mathbb{E}_{1/2} |\mathcal{A}_P(n)|$ . By (2.10), we computed  $g(n)$  for all  $n \leq 16384$ . Figure 1 plots  $\log g(n)/\log n$  versus  $n$  for  $n \leq 16384$ , where  $\theta = \log[(1 + \sqrt{5})/2] \approx 0.694$ . It shows that  $\log g(n)/\log n$  is slightly greater than  $\theta$  and appears to converge to  $\theta$

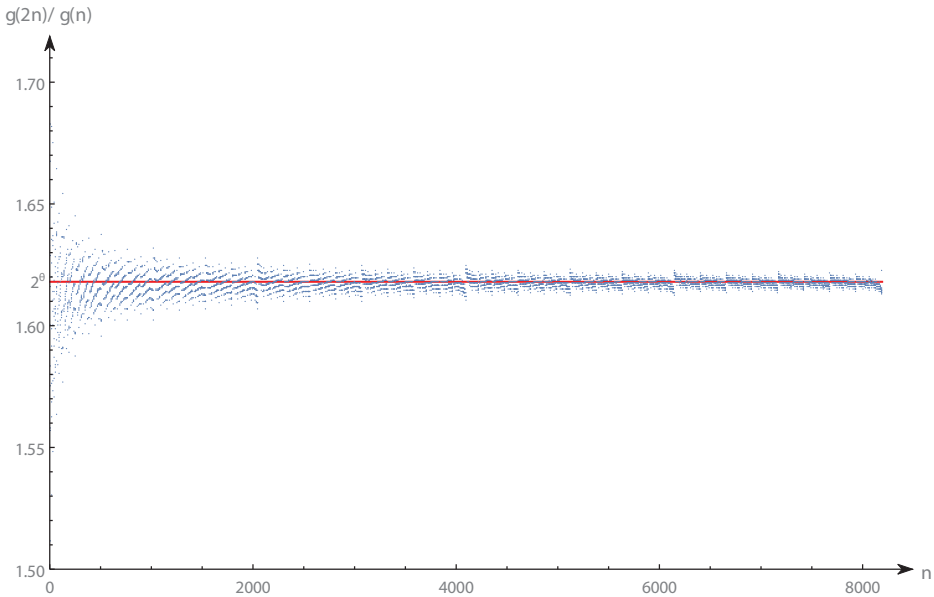


FIGURE 3. Plot of  $g(2n)/g(n)$  versus  $n$ .

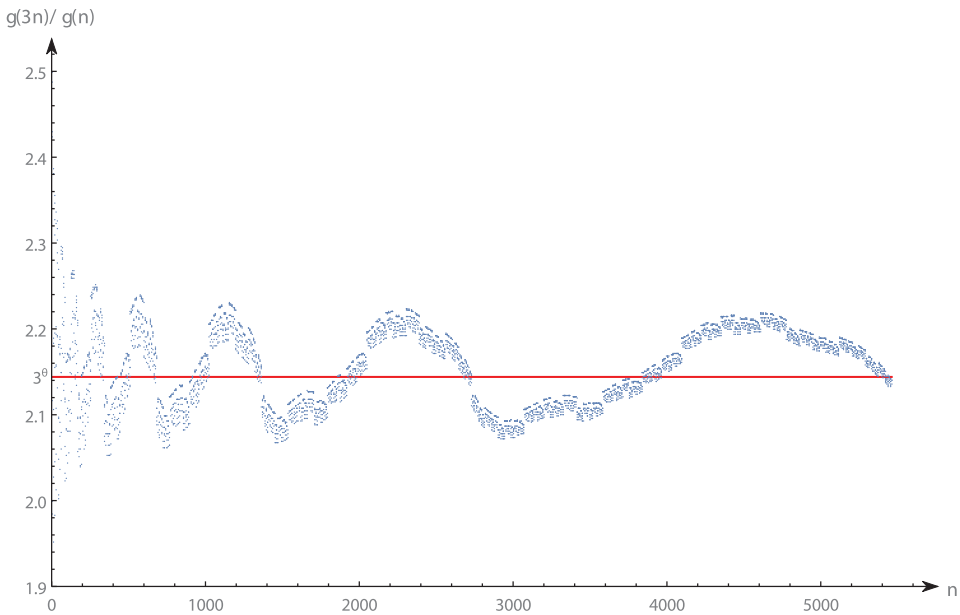


FIGURE 4. Plot of  $g(3n)/g(n)$  versus  $n$ .

slowly. Figure 2 plots  $g(n)/n^\theta$  versus  $n$  for  $n \leq 16384$ . By Proposition 2(ii),

$$\lim_{n \rightarrow \infty} \log [g(n)/n^\theta] / \log n = 0.$$

While it is unclear whether  $g(n)/n^\theta$  converges to some constant eventually, it appears that  $g(n)/n^\theta$  fluctuates less when  $n$  becomes larger. Figure 3 plots  $g(2n)/g(n)$  versus  $n$  for  $n \leq 8192$ . It appears that  $g(2n)/g(n)$  is close to  $2^\theta$  for large  $n$ . Figure 4 plots  $g(3n)/g(n)$

versus  $n$  for  $n \leq 5461$ , where  $g(3n)/g(n)$  oscillates around  $3^\theta$ . Our limited numerical results provide weak evidence that  $g(3n)/g(n)$  converges to  $3^\theta$  eventually.

#### 4. CONCLUDING REMARKS

Recall that  $\mathbb{E}_p |\mathcal{A}_P(n)|$  denotes the expected number of unbiased bits generated when Peres' algorithm is applied to the input sequence  $(X_1, \dots, X_n)$  where  $X_1, \dots, X_n$  are i.i.d. with  $\mathbb{P}(X_i = 1) = p = 1 - \mathbb{P}(X_i = 0)$ . When  $p = 1/2$ ,  $X_1, \dots, X_n$  are unbiased, so that  $g(n) = n - \mathbb{E}_{1/2} |\mathcal{A}_P(n)|$  may be referred to as the cost incurred by Peres' algorithm when not knowing  $p = \frac{1}{2}$ . We derived

$$\lim_{n \rightarrow \infty} \log [n - \mathbb{E}_{1/2} |\mathcal{A}_P(n)|] / \log n = \theta = \log \left( \frac{1 + \sqrt{5}}{2} \right) \quad (4.1)$$

by exploiting the recursion in (1.7). It is a challenging task to obtain more refined results. A positive sequence  $\{L(n)\}$  is said to be regularly varying of index  $\theta$  if  $\lim_{n \rightarrow \infty} L(\lfloor \lambda n \rfloor) / L(n) = \lambda^\theta$  for all  $\lambda > 0$ . (See Ref. [1] for a unified theory of regularly varying sequences.) Figures 3 and 4 suggest that  $g(n)$  may be regularly varying of index  $\theta$ . Furthermore, it is of interest to see if  $g(n)/n^\theta$  converges to a constant (which would imply that  $g(n)$  is regularly varying of index  $\theta$ ). If so, how can this constant be characterized?

For  $p \neq 1/2$ , no recursion like (1.7) is available. It seems difficult to obtain an asymptotic result on  $nh(p) - \mathbb{E}_p |\mathcal{A}_P(n)|$  similar to (4.1). Furthermore,  $\text{Var}_p(|\mathcal{A}_P(n)|)$ , the variance of  $|\mathcal{A}_P(n)|$ , is also of interest and importance. Even for  $p = 1/2$ , it seems challenging to derive the asymptotic behavior of  $\text{Var}_{1/2}(|\mathcal{A}_P(n)|)$  as  $n \rightarrow \infty$ .

#### Acknowledgments

The authors gratefully acknowledge support from the Ministry of Science and Technology, Taiwan, ROC.

#### References

1. Bojanic, R. & Seneta, E. (1973). A unified theory of regularly varying sequences. *Mathematische Zeitschrift* 134(2): 91–106.
2. Elias, P. (1972). The efficient construction of an unbiased random sequence. *Annals of Mathematical Statistics* 43(3): 865–870.
3. Jacquet, P. & Szpankowski, W. (1999). Entropy computations via analytic depoissonization. *IEEE Transactions on Information Theory* 45(4): 1072–1081.
4. Kallenberg, O. (2002). *Foundations of modern probability*. New York: Springer.
5. Keane, M.S. & O'Brien, G.L. (1994). A Bernoulli factory. *ACM Transactions on Modeling and Computer Simulation* 4(2): 213–219.
6. Nacu, Ş. & Peres, Y. (2005). Fast simulation of new coins from old. *Annals of Applied Probability* 15(1A): 93–115.
7. Pae, S.-i. (2013). Exact output rate of Peres's algorithm for random number generation. *Information Processing Letters* 113(5): 160–164.
8. Pae, S.-i. & Loui, M.C. (2006). Randomizing functions: simulation of a discrete probability distribution using a source of unknown distribution. *IEEE Transactions on Information Theory* 52(11): 4965–4976.
9. Peres, Y. (1992). Iterating von Neumann's procedure for extracting random bits. *The Annals of Statistics* 20(1): 590–597.
10. Prasitsupparote, A., Konno, N., & Shikata, J. (2018). Numerical and non-asymptotic analysis of Elias's and Peres's extractors with finite input sequences. *Entropy* 20(10): Article ID 729, 19 pages.
11. Ryabko, B. (1998). Fast enumeration of combinatorial objects. *Discrete Mathematics and Applications* 8(2): 163–182.

12. Ryabko, B.Y. & Matchikina, E. (2000). Fast and efficient construction of an unbiased random sequence. *IEEE Transactions on Information Theory* 46(3): 1090–1093.
13. von Neumann, J. (1951). Various techniques used in connection with random digits. *National Bureau of Standards Applied Math Series* 12: 36–38.
14. Wei, W. & Guo, H. (2009). Bias-free true random-number generator. *Optics Letters* 34(12): 1876–1878.