Source Codes as Random Number Generators

Karthik Visweswariah, *Student Member, IEEE*, Sanjeev R. Kulkarni, *Senior Member, IEEE*, and Sergio Verdú, *Fellow, IEEE*

***Abstract*—A random number generator generates fair coin flips by processing deterministically an arbitrary source of nonideal randomness. An optimal random number generator generates asymptotically fair coin flips from a stationary ergodic source at a rate of bits per source symbol equal to the entropy rate of the source. Since optimal noiseless data compression codes produce incompressible outputs, it is natural to investigate their capabilities as optimal random number generators. In this paper we show under general conditions that optimal variable-length source codes asymptotically achieve optimal variable-length random bit generation in a rather strong sense. In particular, we show in what sense the Lempel–Ziv algorithm can be considered an optimal universal random bit generator from arbitrary stationary ergodic random sources with unknown distributions.**

***Index Terms*— Data compression, entropy, Lempel–Ziv algorithm, random number generation, universal source coding.**

I. Introduction

In contrast to pseudorandom number generators which produce zero entropy rate sequences, a random number generator is a deterministic procedure to generate equiprobable independent bits from a random source Z. The problem was initially addressed by von Neumann in [12] where the source Z was a Bernoulli source with $p \neq 1/2$. More efficient algorithms for generating random bits from a biased coin were given by Hoeffding and Simons [6], Stout and Warren [9], and Peres [7]. Elias [4] showed that the entropy rate is an upper bound for the rate at which it is possible to generate random bits from stationary sources and found an optimal random number generator from stationary finite-state, finite-order Markov sources without making use of the Markov source distribution (other than its order). A simple algorithm to generate arbitrary distributions from a biased coin with a known distribution was given by Han and Hoshi [5]. The practically important problem of constructing a universal random number generator from arbitrary nonideal stationary sources has remained open. Vembu and Verdú [11] gave fundamental limits on the rate at which random bits can be generated from an arbitrary source. In particular, it was shown that for fixed-length random number generation the maximum achievable rate is the $\inf$-entropy rate of the source and for variable-rate random number generation the maximum achievable rate is the $\liminf$ of the entropy rate of the source. The proof of achievability in [11] was constructive but depended on knowing the source distribution.

Manuscript received July 3, 1996; revised August 25, 1997. This work was supported in part by the National Science Foundation under Grants NYI Award IRI-9457645 and NCR 9523805.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: verdu@princeton.edu).

Publisher Item Identifier S 0018-9448(98)00839-6.

Here we explore the possibility of using source codes as random number generators. The rationale is that an incompressible sequence must be random in some sense and thus an optimal source code (which eliminates all redundancy) must output a "random" sequence. The problem of generating random bits has been studied in two settings: variable-length and fixed-length, according to whether the number of random bits generated depends on the source realization or not. This paper focuses on the variable-length setting, as our main interest is in constructive methods for random number generation, and in particular, in universal methods. The Lempel–Ziv algorithm has been shown in [14] to be optimal in a certain sense to test whether or not a source generates independent equiprobable bits. However, the problem of investigating how far from being truly random is the output of a Lempel–Ziv algorithm (and other optimal source codes) appears to be new.

In the variable-length setting the "ideal" definition of randomness would be that, conditioned on the length of the output binary sequence being l, all the l-length binary sequences occur with probability 2^{-l}. The rate at which a variable-length generator produces random bits is the expected length of the binary output string per source symbol. However, we cannot go very far with such a strict requirement of randomness. Consider the simplest nontrivial setting: a memoryless source Z with alphabet $\{a, b, c\}$ such that $P(a) = 1/2, P(b) = 1/4$, and $P(c) = 1/4$. The Huffman code for this source assigns the strings 0, 10, 11 to a, b, c, respectively. Thus a string of n source symbols is mapped to a string of l bits, where l ranges from n to $2n$. The rate of bit generation is equal to the source entropy (1.5 b/symbol). All generated bit strings of length l are equiprobable; however, the Huffman code generates only

$$\binom{n}{l-n} 2^{l-n}$$

different strings of length l. Therefore, the Huffman encoder does not satisfy the "ideal" definition of randomness. The natural alternative advocated in a number of recent works is to adopt a distance measure between probability distributions and require that the distance between the output distributions (conditioned on the output length) and the ideal distribution vanishes as $n \to \infty$. The results of [11] show that for distance measures such as variational distance and normalized divergence the maximum rate of random number generation is equal to the source entropy rate for stationary ergodic sources. However, if we restrict attention to variable-length source codes as random number generators, variational distance reveals that the problem we consider in this paper is not as straightforward as it may have been surmised at first glance. Returning to the example of the memoryless ternary

source, we can check that the variational distance (sum of the absolute difference between all probability masses) between the generated and ideal distributions of sequences of length l is

$$2\left(1-\binom{n}{l-n}2^{-n}\right)$$

for $l=n,n+1,\cdots,2n$, which approaches 2 for every l. If, instead, we consider the normalized divergence between the generated and ideal distributions we obtain

$$\log 2-\frac{1}{n}\log\binom{n}{l-n}. \tag{1}$$

This means that for generated lengths other than those in the neighborhood of $3n/2$, the normalized divergence does not vanish as $n\to\infty$. Fortunately, these "unfavorable" lengths have vanishing probability. Indeed, normalized divergence along with the elimination from consideration of unfavorable generated lengths will serve as the basis for our definition of random rate generation. Unlike variational distance such a definition will lead to the demonstration that optimal source codes are optimal random bit generators, while at the same time being a *bona fide* definition of randomness in the sense that the Kolmogorov complexity of the output is maximal.

Section II formalizes our definition of a rate-R random bit generator (RBG). Although this definition is different from those introduced in [11], we show that the fundamental limits proved in [11] for stationary ergodic sources also hold for the new definition, namely, for every stationary ergodic source there exist generators of random bits at the entropy rate, but not at any higher rate. Unlike the definitions of [11], the definition we give in Section II can be used to prove positive results on optimal universal random bit generation.

In Section III, we show (under conditions more general than stationarity and ergodicity of the source) that optimal variable-length source codes (in the sense of probability or expected length) are generators of random bits at the maximum possible rate. Consequently, we show that Shannon, Huffman, and Lempel–Ziv codes[1] are optimal random bit generators for stationary ergodic sources. The latter result establishes yet another use of the celebrated Lempel–Ziv algorithm: optimal universal random bit generation from a stationary ergodic source.

Section IV shows that optimal random bit generators in the sense of Section II generate strings with maximal Kolmogorov complexity when driven by stationary ergodic sources.

II. Preliminaries

We deal with discrete random sources, characterized by their sequence of finite-dimensional distributions

$$\boldsymbol{Z}=\{P_{Z^n}\}_{n=1}^{\infty}$$

where Z^n takes values in A^n, and A is a finite set. Note that the sources we allow include but are not restricted to random processes since at this point we need not place consistency requirements on the finite-dimensional distributions.

We denote the set of finite-length binary sequences by $\{0,1\}^*$ and $l\colon\{0,1\}^*\mapsto N$ is the function that maps a finite-length bit string to its length.

Definition 1: A sequence of deterministic mappings

$$\{\phi_n\colon A^n\mapsto\{0,1\}^*\}$$

is a rate-R RBG for a source $\boldsymbol{Z}$ if there exists a sequence of sets G_n of positive integers such that the following conditions are met:

[C1]

$$\lim_{n\to\infty}P(l(\phi_n(Z^n))\in G_n)=1 \tag{2}$$

[C2]

$$\liminf_{n\to\infty}\frac{1}{n}\sum_{r\in G_n}rP(l(\phi_n(Z^n))=r)=R \tag{3}$$

and

[C3]

$$\limsup_{n\to\infty}\max_{r\in G_n}\frac{1}{r}D(\phi_n(Z_r^n)\|B^r)=0 \tag{4}$$

where B^r has the equiprobable distribution on $\{0,1\}^r$ and Z_r^n is Z^n restricted to $\{z^n\colon l(\phi_n(z^n))=r\}$.

The above definition is a modified version of [11, Definition 4]; it requires that the average length of the output bit string be sufficiently high [C2] and that conditioned on the length of the output bit string, we get "almost" equiprobable bits [C3]. The notion of "almost" equiprobable is made precise by using normalized divergence as a measure of distance from pure random bits. Note that [C3] is equivalent to the condition on the entropy of generated strings

$$\liminf_{n\to\infty}\min_{r\in G_n}\frac{H(\phi_n(Z_r^n))}{r}=1.$$

Note that [C3] is a condition on "good" lengths (i.e., lengths in G_n) and not on all output lengths. Also note that the the summation in [C2] is only over the set of "good" lengths. [C1] ensures that the probability that the generated length is "good" tends to 1.

Definition 2: The maximum randomness rate $V(\boldsymbol{Z})$ of a source $\boldsymbol{Z}$ is the supremum over R for which there exists a rate-R random bit generator for the source $\boldsymbol{Z}$.

Definition 2 differs from [11, Definition 5]; it allows that the set of "good" lengths G_n be unknown to the random bit generator. This turns out to be important in order to show that the Lempel–Ziv algorithm is an optimal universal random bit generator.

As in [11], the maximum randomness rate of a stationary ergodic source is equal to its entropy rate. First we prove the following general converse. Note that this converse strengthens the converse in [11] which itself strengthens the converse in [4].

[1]Throughout this paper "Lempel–Ziv code" refers to the incremental parsing scheme (LZ '78) as described in [2].

Theorem 1: For any source $\boldsymbol{Z}$ we have

$$V(\boldsymbol{Z}) \le \liminf_{n\to\infty} \frac{H(Z^n)}{n} \triangleq H. \tag{5}$$

Proof: Suppose we have a rate-R RBG $\{\phi_n\colon A^n \mapsto \{0,1\}^*\}$ for the source $\boldsymbol{Z}$. Let $L_n = l(\phi_n(Z^n))$. Fix $\epsilon > 0$ and $\delta > 0$ and select a sequence of sets G_n of positive integers such that for sufficiently large n

$$\min_{r\in G_n} \frac{H(\phi_n(Z_r^n))}{r} > 1-\epsilon \tag{6}$$

and

$$\frac{1}{n}\sum_{r\in G_n} rP(l(\phi_n(Z^n)) = r) \ge R - \delta. \tag{7}$$

Then we have

$$\begin{aligned}
\frac{H(Z^n)}{n} &\ge \frac{1}{n}\sum_{r\in G_n} H(Z_r^n) P_{L_n}(r) && (8)\\
&\ge \frac{1}{n}\sum_{r\in G_n} r\frac{H(\phi_r(Z_r^n))}{r} P_{L_n}(r) && (9)\\
&\ge \frac{1}{n}\sum_{r\in G_n} r(1-\epsilon) P_{L_n}(r) && (10)\\
&\ge (1-\epsilon)(R-\delta) && (11)
\end{aligned}$$

where (8) holds because conditioning reduces entropy and we are not conditioning over all lengths; (9) holds since the ϕ_n are deterministic mappings; and (10) as well as (11) hold for sufficiently large n due to (6) and (7). Since $\epsilon > 0$ and $\delta > 0$ are arbitrary, we must have

$$\liminf_{n\to\infty} \frac{H(Z^n)}{n} \ge R. \qquad \square$$

Unlike [11, Definition 5], (5) does not hold with equality for all sources because of condition [C1] in Definition 1. This is shown by an example in Appendix I.

We will show in Section III that Shannon codes, Lempel–Ziv codes, and Huffman codes can be used as random bit generators to generate random bits (in the sense defined above) from a stationary, ergodic source at the maximum possible rate. Thus we have the following theorem.

Theorem 2: The maximum randomness rate of a stationary ergodic source $\boldsymbol{Z}$ is equal to its entropy rate.

III. Optimal Source Codes as Random Number Generators

In this section we derive two sufficient conditions on the source so that optimal source codes generate random bits at the maximum possible rate. To do so we must first formalize the notion of optimal source codes.

Definition 3: A variable-length lossless source code ϕ for a random source $\boldsymbol{Z}$, is a sequence of one-to-one fixed-length to variable-length mappings $\{\phi_n\}$ where ϕ_n maps A^n to $\{0,1\}^*$.

Definition 4: A variable-length lossless source code is optimal for the source $\boldsymbol{Z}$ if for any $\delta > 0$

$$\lim_{n\to\infty} P_{Z^n}(|l(\phi_n(Z^n)) - H(Z^n)| \le H(Z^n)\delta) = 1.$$

Definition 5: A variable-length lossless source code for the source $\boldsymbol{Z}$ is mean-optimal if for any $\delta > 0$

$$E[l(\phi_n(Z^n))] < H(Z^n) + n\delta$$

for all sufficiently large n.

We now show that an optimal lossless source code generates random bits at the optimum rate.[2]

Theorem 3: Consider a source $\boldsymbol{Z}$ such that

$$H = \liminf_{n\to\infty} \frac{H(Z^n)}{n} > 0. \tag{12}$$

If ϕ is an optimal lossless source code for $\boldsymbol{Z}$ then ϕ is a rate-H random bit generator.

Proof: Let

$$\begin{aligned}
I_n &= \{r\colon l(\phi_n(z^n)) = r, \text{ for some } z^n \in A^n\} && (13)\\
J_{n,r} &= \{z^n \in A^n\colon l(\phi_n(z^n)) = r\}
\end{aligned}$$

and

$$L_n = l(\phi_n(Z^n)).$$

I_n is the set of possible output lengths at stage n, $J_{n,r}$ is the subset of A^n that is mapped to an output of length r.

Note that

$$P_{L_n}(r) = P_{Z^n}(J_{n,r}). \tag{14}$$

First we will show that $E[H(\phi_n(Z_{L_n}^n))/L_n] \to 1$ as $n \to \infty$. Since the mapping

$$\phi_n\colon A^n \mapsto \{0,1\}^*$$

is one-to-one, we have

$$H(\phi_n(Z_{L_n}^n)) = H(Z_{L_n}^n).$$

Fix $\delta > 0$. Let

$$I_n' = I_n \cap [H(Z^n)(1-\delta), H(Z^n)(1+\delta)].$$

Then by the definition of an optimal code we have

$$P(L_n \in I_n \cap I_n'^c) \to 0 \tag{15}$$

as $n \to \infty$.

Now (see (16)–(21) on the following page).

To bound the second and third terms in the numerator of (21), we use the following elementary result.

[2]Notice that a source may satisfy condition (12) and yet it may be such that no optimal lossless source code exists.

$$E\left[\frac{H(Z^n_{L_n})}{L_n}\right] = \sum_{r\in I_n} P_{L_n}(r)\frac{H(Z^n_r)}{r} \tag{16}$$

$$\geq \sum_{r\in I'_n} P_{L_n}(r)\frac{H(Z^n_r)}{r} \tag{17}$$

$$\geq \frac{1}{H(Z^n)(1+\delta)} \sum_{r\in I'_n} P_{L_n}(r) H(Z^n_r) \tag{18}$$

$$= \frac{1}{H(Z^n)(1+\delta)} \sum_{r\in I'_n} \sum_{z^n\in J_{n,r}} P_{Z^n}(z^n) \log \frac{P_{L_n}(r)}{P_{Z^n}(z^n)} \tag{19}$$

$$= \frac{\displaystyle\sum_{r\in I'_n} \sum_{z^n\in J_{n,r}} P_{Z^n}(z^n) \log \frac{1}{P_{Z^n}(z^n)} - \sum_{r\in I'_n} P_{L_n}(r) \log \frac{1}{P_{L_n}(r)}}{H(Z^n)(1+\delta)} \tag{20}$$

$$= \frac{H(Z^n) - \displaystyle\sum_{r\in I_n\cap I'^c_n} \sum_{z^n\in J_{n,r}} P_{Z^n}(z^n) \log \frac{1}{P_{Z^n}(z^n)} - \sum_{r\in I'_n} P_{L_n}(r) \log \frac{1}{P_{L_n}(r)}}{H(Z^n)(1+\delta)}. \tag{21}$$

Lemma 1: Let X be a random variable taking values on a set F. Let G be a finite subset of F. Then we have

$$\sum_{x\in G} P_X(x) \log \frac{1}{P_X(x)} \leq P_X(G)\left(\log|G| + \log \frac{1}{P_X(G)}\right).$$

Using Lemma 1 twice along with the fact that $|I_n| \leq |A|^n$ and $|J_{n,r}| \leq |A|^n$ for every $r \in I_n$ we have

$$\sum_{r\in I_n\cap I'^c_n} \sum_{z^n\in J_{n,r}} P_{Z^n}(z^n) \log\left(\frac{1}{P_{Z^n}(z^n)}\right) \leq P_{L_n}(I_n \cap I'^c_n)\left(2n\log|A| + \log \frac{1}{P_{L_n}(I_n \cap I'^c_n)}\right).$$

Thus we have

$$\limsup_{n\to\infty} \frac{\displaystyle\sum_{r\in I_n\cap I'^c_n} \sum_{z^n\in J_{n,r}} P_{Z^n}(z^n) \log\left(\frac{1}{P_{Z^n}(z^n)}\right)}{H(Z^n)(1+\delta)} \tag{22}$$

$$\leq \limsup_{n\to\infty} \frac{P_{L_n}(I_n \cap I'^c_n)\left(2n\log|A| + \log \dfrac{1}{P_{L_n}(I_n \cap I'^c_n)}\right)}{H(Z^n)(1+\delta)} \tag{23}$$

$$= 0 \tag{24}$$

where (24) follows from (12) and (15).

Similarly we upper-bound the third term in the numerator of (21) using Lemma 1 to get

$$\limsup_{n\to\infty} \frac{\displaystyle\sum_{r\in I'_n} P_{L_n}(r) \log\left(\frac{1}{P_{L_n}(r)}\right)}{H(Z^n)(1+\delta)} \tag{25}$$

$$\leq \limsup_{n\to\infty} \frac{P_{L_n}(I'_n)\left(\log(2H(Z^n)\delta + 1) + \log \dfrac{1}{P_{L_n}(I'_n)}\right)}{H(Z^n)(1+\delta)} \tag{26}$$

$$= 0 \tag{27}$$

where the equality follows because of (12).

Thus using (21), (24), and (27) we have

$$\liminf_{n\to\infty} E\left[\frac{H(\phi_n(Z_{L_n}^n))}{L_n}\right] \geq \frac{1}{1+\delta}.$$

Since

$$E\left[\frac{H(\phi_n(Z_{L_n}^n))}{L_n}\right] \leq 1$$

for each n and since $\delta > 0$ was arbitrary we have

$$\lim_{n\to\infty} E\left[\frac{H(\phi_n(Z_{L_n}^n))}{L_n}\right] = 1. \tag{28}$$

Fix $\epsilon > 0$. Define

$$K_n(\epsilon) = \left\{r \in I_n \colon \frac{H(\phi_n(Z_r^n))}{r} > 1-\epsilon\right\}. \tag{29}$$

We will show by contradiction that $P_{L_n}(K_n(\epsilon)) \to 1$ as $n \to \infty$. Suppose there exists $\alpha > 0$ such that $P_{L_n}(K_n(\epsilon)) \leq 1-\alpha$ infinitely often. Since $H(\phi_n(Z_{L_n}^n)) \leq L_n$ we have

$$E\left[\frac{H(\phi_n(Z_{L_n}^n))}{L_n}\right] \leq (1-\epsilon)\alpha + (1-\alpha) = 1-\alpha\epsilon$$

infinitely often. But this contradicts (28), and so there is no $\alpha > 0$ such that $P_{L_n}(K_n(\epsilon)) \leq 1-\alpha$ infinitely often. Thus

$$\lim_{n\to\infty} P_{L_n}(K_n(\epsilon)) = 1.$$

Let

$$H' = \limsup_{n\to\infty} \frac{H(Z^n)}{n}.$$

Now fix $\beta > 0$ and $\epsilon > 0$. Let $\gamma = \beta/H'$. Note that $\gamma > 0$ since $H' < \infty$. Let

$$G_n(\epsilon) = K_n(\epsilon) \cap \left[H(Z^n)\left(1-\frac{\gamma}{2}\right), H(Z^n)\left(1+\frac{\gamma}{2}\right)\right].$$

By the definition of $K_n(\epsilon)$ in (29) we have

$$\min_{r\in K_n(\epsilon)} \frac{H(\phi_n(Z_r^n))}{r} > 1-\epsilon$$

for each n and thus

$$\max_{r\in G_n(\epsilon)} \frac{D(\phi_n(Z_r^n)\|B^r)}{r} < \epsilon \tag{30}$$

for each n. Also since $P_{L_n}(K_n(\epsilon)) \to 1$ as $n \to \infty$ and

$$P_{L_n}([H(Z^n)(1-(\gamma/2)), H(Z^n)(1+(\gamma/2))]) \to 1$$

as $n \to \infty$ we have

$$\lim_{n\to\infty} P_{L_n}(G_n(\epsilon)) = 1. \tag{31}$$

Thus

$$\frac{1}{n}\sum_{r\in G_n(\epsilon)} r P_{L_n}(r) \geq \frac{H(Z^n)}{n}\left(1-\frac{\gamma}{2}\right)\sum_{r\in G_n(\epsilon)} P_{L_n}(r) \tag{32}$$

$$= \frac{H(Z^n)}{n}\left(1-\frac{\gamma}{2}\right)P_{L_n}(G_n(\epsilon)) \tag{33}$$

$$\geq \frac{H(Z^n)}{n}P_{L_n}(G_n(\epsilon)) - \frac{\gamma H(Z^n)}{2n} \tag{34}$$

$$\geq \liminf_{n\to\infty}\frac{H(Z^n)}{n} - \beta \tag{35}$$

$$= H - \beta \tag{36}$$

where (35) holds for sufficiently large n. Since $\epsilon > 0$ and $\beta > 0$ were arbitrary, (30), (31), and (35) imply that the source code ϕ is a rate-H RBG. Since we have seen that the maximum randomness rate of a source satisfies

$$V(\boldsymbol{Z}) \leq \liminf_{n\to\infty}\frac{H(Z^n)}{n}$$

the source code ϕ generates bits at the maximum possible rate. □

We now show that a mean-optimal source code generates random bits at the optimal rate. Recall that the inf-entropy rate of a source $\boldsymbol{Z}$, $\underline{H}(\boldsymbol{Z})$ is the largest extended real number α that satisfies for all $\delta > 0$

$$\lim_{n\to\infty} P_{Z^n}\left(\frac{1}{n}\log\frac{1}{P_{Z^n}(Z^n)} \leq \alpha - \delta\right) = 0.$$

Theorem 4: Consider a source $\boldsymbol{Z}$ such that Z^n takes values in A^n and

$$\underline{H}(\boldsymbol{Z}) > 0. \tag{37}$$

If ϕ is a mean-optimal source code for $\boldsymbol{Z}$ then ϕ is a rate-H RBG.

The proof of this theorem will make use of the following result.

Lemma 2: Consider a source $\boldsymbol{Z}$ with Z^n taking values in A^n such that $\underline{H}(\boldsymbol{Z}) = \alpha > 0$. Suppose that ϕ is a variable-length mean-optimal source code for the source. If there exists a set of source n-strings C_n such that for any $\delta > 0$

$$\frac{1}{n}\sum_{z^n\in C_n} l(\phi(z^n))P_{Z^n}(z^n) \geq \frac{H(Z^n)}{n} - \delta \tag{38}$$

for all sufficiently large n then $P_{Z^n}(C_n) \to 1$ as $n \to \infty$.

Proof: We will proceed by contradiction. Suppose that there exists $\beta > 0$, such that $P_{Z^n}(C_n^c) > \beta$ for all $n \in I$ where I is an infinite set of integers. Let

$$S_n = \{z^n \in A^n \colon P_{Z^n}(z^n) \leq 2^{-n(3\alpha/4)}\}.$$

By the definition of $\underline{H}(\boldsymbol{Z})$ we have $P_{Z^n}(S_n) \to 1$ as $n \to \infty$. So we must have

$$P_{Z^n}(C_n^c \cap S_n) > \frac{\beta}{2}$$

for sufficiently large $n \in I$. Now

$$\frac{1}{n}\sum_{z^n\in C_n^c\cap S_n} P_{Z^n}(z^n)l(\phi(z^n)) = P_{Z^n}(C_n^c \cap S_n)\frac{1}{n}\sum_{z^n\in C_n^c\cap S_n} \cdot \frac{P_{Z^n}(z^n)}{P_{Z^n}(C_n^c \cap S_n)} l(\phi(z^n)) \tag{39}$$

$$\geq \frac{H(Z^n_{C_n^c\cap S_n})}{n} P_{Z^n}(C_n^c \cap S_n) \tag{40}$$

$$\geq \frac{\log\left(\frac{\beta}{2}2^{n(3\alpha/4)}\right)}{n}\,\frac{\beta}{2} \tag{41}$$

$$\geq \frac{3\alpha\beta}{16}. \tag{42}$$

where (41) follows from [3, Problem 1.1.10].

Equation (42) along with (38) with $\delta = 3\alpha\beta/32$ contradicts the fact that ϕ is a mean-optimal code for $\boldsymbol{Z}$. Thus we must have $P_{Z^n}(C_n) \to 1$ as $n \to \infty$. □

Proof of Theorem 4: We define $I_n, J_{n,r}$, and L_n as in the proof of Theorem 3.

Fix $\epsilon > 0$. Define the set of "bad" output lengths

$$B_n(\epsilon) = \left\{ r \in I_n \colon \frac{H(Z_r^n)}{r} \le 1 - \epsilon \right\}.$$

Our objective is to show that

$$\lim_{n\to\infty} P_{L_n}(B_n(\epsilon)) = 0 \tag{43}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{r\in B_n(\epsilon)} r P_{L_n}(r) = 0. \tag{44}$$

Since ϕ is a mean-optimal source code, (44) will follow if we can show that for any $\delta > 0$

$$\frac{1}{n} \sum_{r\in B_n(\epsilon)^c} r P_{L_n}(r) \ge \frac{H(Z^n)}{n} - \delta \tag{45}$$

for sufficiently large n.

Let us proceed by contradiction: assume there exists $\delta > 0$ such that (45) is not satisfied for $n \in I$ where I is an infinite set of integers.

Now consider a new source code ϕ' for the source $\boldsymbol{Z}$. To code $z^n \in J_{n,r}$ we code Z_r^n using a Huffman code and then prefix that code with a code for the length r. We code r by repeating every bit in the binary expansion of r twice and then ending the description with a 01. Thus r is coded using at most $2\log r + 4$ bits. For example, 5 would be coded as 11001101. Thus we have a prefix code for the source $\boldsymbol{Z}$. Now the average length of the new code is

$$\sum_{z^n\in A^n} l(\phi'(z^n)) P_{Z^n}(z^n) \tag{46}$$

$$= \sum_{r\in I_n} \sum_{z^n\in J_{n,r}} l(\phi'(z^n)) P_{Z^n}(z^n) \tag{47}$$

$$= \sum_{r\in B_n(\epsilon)} \sum_{z^n\in J_{n,r}} l(\phi'(z^n)) P_{Z^n}(z^n) + \sum_{r\in B_n(\epsilon)^c} \sum_{z^n\in J_{n,r}} l(\phi'(z^n)) P_{Z^n}(z^n) \tag{48}$$

$$= \sum_{r\in B_n(\epsilon)} P_{L_n}(r) \sum_{z^n\in J_{n,r}} l(\phi'(z^n)) \frac{P_{Z^n}(z^n)}{P_{L_n}(r)} + \sum_{r\in B_n(\epsilon)^c} P_{L_n}(r) \sum_{z^n\in J_{n,r}} l(\phi'(z^n)) \frac{P_{Z^n}(z^n)}{P_{L_n}(r)} \tag{49}$$

$$\le \sum_{r\in B_n(\epsilon)} P_{L_n}(r)(r(1-\epsilon) + 1 + 2\log r + 4) + \sum_{r\in B_n(\epsilon)^c} P_{L_n}(r)(r + 2\log r + 4) \tag{50}$$

$$\le \sum_{r\in I_n} r P_{L_n}(r) + 2 \sum_{r\in I_n} P_{L_n}(r) \log r + 5 - \epsilon \sum_{r\in B_n(\epsilon)} r P_{L_n}(r) \tag{51}$$

$$\le \sum_{r\in I_n} r P_{L_n}(r) + 2\log\left(\sum_{r\in I_n} r P_{L_n}(r) \right) + 5 - \epsilon \sum_{r\in B_n(\epsilon)} r P_{L_n}(r) \tag{52}$$

$$\le H(Z^n) + \frac{n\epsilon\delta}{4} + 2\log\left(H(Z^n) + \frac{n\epsilon\delta}{4} \right) + 5 - \epsilon \sum_{r\in B_n(\epsilon)} r P_{L_n}(r) \tag{53}$$

$$\le H(Z^n) + \frac{n\epsilon\delta}{4} + 2\log\left(H(Z^n) + \frac{n\epsilon\delta}{4} \right) + 5 - n\epsilon\delta \tag{54}$$

$$\le H(Z^n) - \frac{n\epsilon\delta}{2} \tag{55}$$

where (50) follows because the average code length when Z_r^n is coded by a Huffman code is upper-bounded by $H(Z_r^n) + 1$, (52) follows because of the concavity of the logarithm, and (53)–(55) hold for $n \in I$.

Since ϕ' is a prefix code for $\boldsymbol{Z}$ the average length in (46) is lower-bounded by $H(Z^n)$. This is contradicted by (55). Thus (45) is established; (45) implies (44) and due to Lemma 2, (45) also implies (43) and the theorem is proved. □

IV. Shannon, Huffman, and Lempel–Ziv Codes

We first introduce some notation. Let us denote the sequences of mappings corresponding to the Shannon, Huffman, and the Lempel–Ziv codes by ϕ_n^S, ϕ_n^H, and ϕ_n^L, respectively. By "Lempel–Ziv code" we mean the LZ'78 incremental parsing scheme as described in [2].

We can now state and prove our results establishing that Shannon, Huffman, and Lempel–Ziv codes are optimal random bit generators.

Theorem 5: Let $\boldsymbol{Z}$ be a source with $\underline{H}(\boldsymbol{Z}) > 0$. The Shannon code is a random bit generator generating bits at the maximum possible rate, $V(\boldsymbol{Z})$.

Proof: We use Theorem 4 to prove the result. Since the source $\boldsymbol{Z}$ satisfies the required conditions, we only have to check that the Shannon code is mean-optimal. This is true since the Shannon code is lossless and from [2, Theorem 5.4.3] we have

$$E[l(\phi_n^S(Z^n))] \le H(Z^n) + 1.$$

Thus the Shannon code satisfies all the assumptions required in Theorem 4. □

Theorem 6: Let $\boldsymbol{Z}$ be a source $\boldsymbol{Z}$ with $\underline{H}(\boldsymbol{Z}) > 0$. The Huffman code is a random bit generator that generates bits at the maximum possible rate.

Proof: We use Theorem 4 to prove the result. Since the source $\boldsymbol{Z}$ satisfies the required conditions, we only have to check that the Huffman code is mean-optimal. This is true since the Huffman code is lossless and from [2, Theorems 5.4.1 and 5.8.1] we have

$$E[l(\phi_n^H(Z^n))] \le H(Z^n) + 1.$$

Thus the Huffman code satisfies all the assumptions required in Theorem 4. □

Theorem 7: Let $\boldsymbol{Z}$ be a stationary, ergodic source with a finite entropy rate $H>0$. The Lempel–Ziv code is a rate-H RBG.

Proof: First we note that since $\boldsymbol{Z}$ is a stationary, ergodic source with finite entropy rate $H>0$, (12) is satisfied. We now verify that the Lempel–Ziv code is an optimal lossless source code. The Lempel–Ziv code is lossless and so the mapping ϕ_n^L is a one-to-one mapping for each n. Thus we only need to check that the code is optimal in the sense of Definition 4. Fix $\delta>0$.

Consider the set

$$S_n=\{z^n\in A^n\colon l(\phi_n^L(z^n))\le H(Z^n)-n\delta\}.$$

The number of binary sequences with length smaller than L is $2^{(L+1)}-1$. Thus we have

$$|S_n|\le 2^{H(Z^n)-n\delta+1}. \tag{56}$$

Stationary ergodic sources satisfy the asymptotoc equipartition property (AEP) and for sources that satisfy the AEP it can be shown that if a sequence of sets S_n satisfies (56) then $P_{Z^n}(S_n)\to 0$ as $n\to\infty$. Also from [2, Theorem 12.10.2], we have

$$P_{Z^n}(\{l(\phi_n^L(Z^n))\le n(H+\delta)\})\to 1$$

as $n\to\infty$. Since $(H(Z^n)/n)\to H$ as $n\to\infty$, we have

$$P_{Z^n}\left(\left\{z^n\in A^n\colon \frac{1}{n}|l(\phi_n^L(z^n))-H(Z^n)|\le\frac{H\delta}{2}\right\}\right)\to 1$$

as $n\to\infty$. For sufficiently large n, $(H(Z^n)/n)>(H/2)$ and so we have

$$P_{Z^n}(\{z^n\in A^n\colon |l(\phi_n^L(z^n))-H(Z^n)|\le H(Z^n)\delta\})\to 1$$

as $n\to\infty$.

Thus the Lempel–Ziv code satisfies all the assumptions required in Theorem 3. □

Remark: For our proof of the optimality of the Lempel–Ziv code we require [2, Theorem 12.10.2]. In the proof of that theorem it is assumed that if the number of phrases in the parsing of the input is c_n, then $\log c_n$ bits will be needed to code a pointer to a phrase. Since the number of phrases is not known *a priori*, we can use the Elias code or any other universal coding of the integers to either i) code the pointers directly, or ii) first code the number of pointers and then use $\log c_n$ bits to code the pointer to a phrase. Using either of these schemes does not affect the result in [2, Theorem 12.10.2] as a result of the asymptotic optimality of the Elias code [1, Theorem 1] or other universal codes [1].

To conclude this section, we consider the problem of determining the set G_n of good sequences of lengths. From the viewpoint of the user of the random number generator, it would be convenient to know which output lengths ought to be discarded because they do not correspond to almost-fair coin flips. As we saw in Definition 1, the choice of G_n is not unique. However, this choice is not crucial since the probability that the output length belongs to G_n is guaranteed to converge to 1 as $n\to\infty$; thus a user of the random number generator who would not discard any generated sequence would fail to obtain almost-fair coin flips with at most asymptotically vanishing probability.

For Shannon–Fano codes operating with a known memoryless source with entropy H, the sequence of sets (cf. (13))

$$G_n=\{r\in I_n\colon |r-nH|\le n\delta_n\} \tag{57}$$

satisfies the conditions in Definition 1 for any δ_n going to zero sufficiently slowly (Appendix II). In this case, (57) gives a simple rule for the user of the Shannon–Fano code to discard output lengths that do not belong to G_n. Note that if an output length belongs to G_n for a fixed n, Definition 1 does not offer any guarantee that it is a "good" length, since this notion is only defined asymptotically.

The choice in (57) does not satisfy the condition in Definition 1 for every optimal random number generator. For example, consider the trivial case where the source already generates fair coin flips and the random bit generator is such that it appends a 0 to the input string unless it is the all-zero string, in which case it leaves the string unchanged. This mapping generates random bits at the optimal rate. The sequence $\{r_n=n\}$ which satisfies the typicality condition in (57) cannot belong to G_n for all n because

$$\frac{1}{r_n}D(\phi_n(Z_{r_n}^n)\|B^{r_n})=1$$

if $r_n=n$.

Even if we restrict ourselves to the Shannon–Fano code we cannot expect that the choice in (57) remains good beyond the class of stationary ergodic sources. For example, consider a coin with bias probability uniformly distributed in $[0,1]$. Then, (57) evaluated at $H=0.5\log_2 e$ fails to satisfy [C1] in Definition 1 for any $\delta_n\to 0$ as $n\to\infty$. We can still find a simple description of a valid G_n (cf. (13))

$$G_n=\{r\in I_n\colon r\ge\sqrt{n}\}. \tag{58}$$

The validity of this choice of G_n is shown in Appendix II.

In the universal setting when the source is unknown, and an optimal universal source code (e.g., the Lempel–Ziv algorithm) is used, a good on-line estimate of the entropy of the source is r_n/n where r_n is the length of the output sequence due to an input of length n. If we were to substitute H with r_n/n in (57), then G_n would include all output lengths, thereby possibly violating condition [C3] in Definition 1.

V. Kolmogorov Complexity of the Output of an RBG

We will use the same notation for Kolmogorov complexity as in [2]. $K(z^n)$ denotes the Kolmogorov complexity of the sequence z^n and $K(z^n|n)$ denotes the the Kolmogorov complexity of z^n given its length n. We will show that if we are generating random bits according to Definition 1 then with high probability the output sequences have a Kolmogorov complexity per unit length close to 1. This is interesting because from [2] we have that if

$$\lim_{n\to\infty}\frac{K(z^n|n)}{n}=1$$

then the sequence z will pass all computable statistical tests for randomness.

Theorem 8: If a code ϕ is a rate-H RBG for a stationary ergodic source with entropy rate H then

$$\frac{K(\phi(Z^n)|l(\phi(Z^n)))}{l(\phi(Z^n))} \to 1$$

in probability as $n \to \infty$.

Proof: Fix $\gamma > 0$. We will first show that

$$P_{Z^n}(\{z^n: l(\phi(z^n)) > n(H+\gamma)\}) \to 0 \tag{59}$$

as $n \to \infty$. First of all, since ϕ is a random bit generator we can, for any $\epsilon > 0, \delta > 0$ find a sequence of sets G_n as in Definition 1 with the required properties. We fix $\epsilon = \gamma/2(H+\gamma)$ and $\delta > 0$ and consider the corresponding G_n.

Now consider

$$E_n = \{z^n: l(\phi(z^n)) > n(H+\gamma), l(\phi(z^n)) \in G_n\}$$

and for fixed $\alpha > 0$

$$F_n = \left\{ z^n: l(\phi(z^n)) \le n(H+\gamma),\ \frac{1}{n}\log\frac{1}{P_{Z^n}(z^n)} > H - \alpha \right\}.$$

Then we have

$$\begin{aligned} H(Z^n) &\ge P_{Z^n}(E_n)H(Z^n_{E_n}) + P_{Z^n}(F_n)H(Z^n_{F_n}) && (60)\\ &\ge P_{Z^n}(E_n)H(\phi(Z^n_{E_n})) + P_{Z^n}(F_n)H(Z^n_{F_n}) && (61)\\ &\ge P_{Z^n}(E_n)n(H+\gamma)(1-\epsilon) \\ &\quad + P_{Z^n}(F_n)(\log P_{Z^n}(F_n) + (H-\alpha)n). && (62) \end{aligned}$$

Thus (substituting in the value of ϵ)

$$\begin{aligned} \frac{H(Z^n)}{n} &\ge P_{Z^n}(E_n)\left(H + \frac{\gamma}{2}\right) \\ &\quad + P_{Z^n}(F_n)(H-\alpha) + \frac{P_{Z^n}(F_n)\log P_{Z^n}(F_n)}{n}. \end{aligned}$$

Now since $P_{L_n}(G_n) \to 1$ as $n \to \infty$, we have

$$P_{Z^n}(E_n) + P_{Z^n}(F_n) \to 1$$

as $n \to \infty$. Since $\lim_{n\to\infty}(H(Z^n)/n) = H$, if

$$\limsup_{n\to\infty} P_{Z^n}(E_n) > 0$$

we can choose α small enough so as to have a contradiction. Thus

$$\lim_{n\to\infty} P_{Z^n}(E_n) = 0.$$

Since $P_{L_n}(G_n) \to 1$ as $n \to \infty$ we have (59).

Now again fix $\gamma > 0$. We will show that any set $B_n \subseteq \phi(A^n)$ with $|B_n| \le 2^{n(H-\gamma)}$ satisfies

$$P_{\phi(Z^n)}(B_n) \to 0 \tag{63}$$

as $n \to \infty$.

Now fix $\epsilon > 0$, $\delta > 0$. Let G_n be the sequence of sets as in Definition 1. Then we have

$$\begin{aligned} H(\phi(Z^n)) &\ge \sum_{r\in G_n} P_{L^n}(r)H(\phi(Z^n_r)) && (64)\\ &= \sum_{r\in G_n} rP_{L^n}(r)\frac{H(\phi(Z^n_r))}{r} && (65)\\ &\ge (1-\epsilon)\sum_{r\in G_n} rP_{L_n}(r) && (66)\\ &\ge (1-\epsilon)n(H-\delta). && (67) \end{aligned}$$

Fix $\alpha > 0$. Let

$$S_n = \{\phi(z^n): l(\phi(z^n)) \le n(H+\alpha)\}.$$

We also have

$$H(\phi(Z^n)) = \sum_{\phi(A^n)} P_{\phi(Z^n)}(\phi(z^n))\log\frac{1}{P_{\phi(Z^n)}(\phi(z^n))}. \tag{68}$$

We split the the right side of (68) into sums over three sets $B_n, B_n^c \cap S_n$ and $B_n^c \cap S_n^c$. We will now apply Lemma 1 to each of the terms. Thus we have

$$\begin{aligned} &\sum_{B_n} P_{\phi(Z^n)}(\phi(z^n))\log\frac{1}{P_{\phi(Z^n)}(\phi(z^n))} \\ &\quad \le P_{\phi(Z^n)}(B_n)n(H-\gamma) \\ &\qquad - P_{\phi(Z^n)}(B_n)\log P_{\phi(Z^n)}(B_n) && (69) \end{aligned}$$

$$\begin{aligned} &\sum_{B_n^c\cap S_n} P_{\phi(Z^n)}(\phi(z^n))\log\frac{1}{P_{\phi(Z^n)}(\phi(z^n))} \\ &\quad \le P_{\phi(Z^n)}(B_n^c \cap S_n)n(H+\alpha) + 1 \\ &\qquad - P_{\phi(Z^n)}(B_n^c \cap S_n)\log P_{\phi(Z^n)}(B_n^c \cap S_n) && (70) \end{aligned}$$

and

$$\begin{aligned} &\sum_{B_n^c\cap S_n^c} P_{\phi(Z^n)}(\phi(z^n))\log\frac{1}{P_{\phi(Z^n)}(\phi(z^n))} \\ &\quad \le P_{\phi(Z^n)}(B_n^c \cap S_n^c)n \\ &\qquad - P_{\phi(Z^n)}(B_n^c \cap S_n^c)\log P_{\phi(Z^n)}(B_n^c \cap S_n^c). && (71) \end{aligned}$$

Combining inequalities (69)–(71) we have for sufficiently large n

$$\begin{aligned} \frac{H(\phi(Z^n))}{n} &\le P_{\phi(Z^n)}(B_n)(H-\gamma) \\ &\quad + P_{\phi(Z^n)}(B_n^c \cap S_n)(H+\alpha) + \alpha. && (72) \end{aligned}$$

If

$$\limsup_{n\to\infty} P_{\phi(Z^n)}(B_n) > 0$$

then for sufficiently small $\alpha > 0$, (67) and (72) are contradictory since $P_{\phi(Z^n)}(S_n) \to 1$ as $n \to \infty$. Thus we must have

$$\lim_{n\to\infty} P_{\phi(Z^n)}(B_n) = 0.$$

Now fix $\gamma>0$ and consider the sets

$$H_n = \{z^n\colon K(\phi(z^n)) \geq n(H-\gamma), n(H-\gamma) \leq l(\phi(z^n)) \leq n(H+\gamma)\}$$
$$D_n = \{z^n\colon K(\phi(z^n)) < n(H-\gamma) \text{ or } l(\phi(z^n)) < n(H-\gamma)\}.$$

Now $|D_n| \leq 2(2^{n(H-\gamma)+1})$ and so using (59) and (63) we have

$$P_{Z^n}(H_n) \to 1 \tag{73}$$

as $n \to \infty$. From [2, eq. (7.8)] we have

$$K(\phi(z^n)|l(\phi(z^n))) \geq K(\phi(z^n)) - 2\log l(\phi(z^n)) - c$$

where c is a constant independent of z^n. Thus for $z^n \in H_n$

$$\frac{K(\phi(z^n)|l(\phi(z^n)))}{l(\phi(z^n))} \geq \frac{n(H-\gamma) - 2\log(n(H+\gamma)) - c}{n(H+\gamma)}.$$

Since $\gamma>0$ was arbitrary we have for any $\beta>0$

$$P_{Z^n}\left(\left\{z^n\colon \frac{K(\phi(z^n)|l(\phi(z^n)))}{l(\phi(z^n))} > 1-\beta\right\}\right) \to 1$$

as $n \to \infty$. □

Appendix I

We now give an example of a source whose entropy rate is $\frac{1}{2}$ bit and its maximum randomness rate is zero.

Consider a source $\boldsymbol{Z}$ such that Z^n takes values in A^n where $A = \{a,b,c\}$. Let

$$P_{Z^n}(a,a,\cdots,a) = \frac{1}{2}$$

and all the strings with only b's and c's have the same probability $2^{-(n+1)}$. All other strings occur with probability 0. The entropy rate of this nonergodic source is

$$\lim_{n\to\infty} \frac{H(Z^n)}{n} = \frac{1}{2}.$$

Thus the intrinsic randomness rate of this source as defined in [11] is $1/2$.

Consider an arbitrary random number generator ϕ_n. For Condition [C1] in the definition to be satisfied it is clear that for all sufficiently large n, we must have $l(\phi_n(a,a,\cdots,a)) \in G_n$. Select any $r_n \in G_n$. The maximum of $H(\phi_n(Z^n_{r_n}))$ over all ϕ_n such that $l(\phi_n(a^n)) = r_n$ is

$$\frac{1}{2} + \frac{1}{2}\log 2(2^{r_n}-1).$$

Thus if $\{\phi_n\}$ is to satisfy Condition [C3] we must have for any $\epsilon>0$ and for sufficiently large n

$$\frac{1}{r_n} + \frac{\log(2^{r_n}-1)}{2r_n} \geq 1-\epsilon. \tag{74}$$

Since r_n is a positive integer, (74) implies that $r_n = 1$ for all but finitely many n.

Let the number of sequences that are mapped to a string of length r_n (other than from $(a,a,\cdots,a)$) be m_n. To make sure that $r_n \in G_n$ for sufficiently large n we need for any $\delta>0$, for sufficiently large n

$$\frac{m_n}{2^{n+1}} \geq \frac{1}{2} - \delta.$$

If this is the case then for sufficiently large n

$$E[l(\phi_n(Z^n))] \leq 1 + (\log(2^n - m_n) + 1 + \delta)\frac{2^n - m_n}{2^{n+1}} \tag{75}$$
$$\leq 1 + (n + 2 + \log\delta + \delta)\delta. \tag{76}$$

Thus for sufficiently large n

$$\frac{E[l(\phi_n(Z^n))]}{n} \leq 2\delta.$$

$\delta>0$ was arbitrary and thus there is no RBG that generates bits at the required rate.

Note that for this source, neither Theorem 3 nor Theorem 4 apply, as there is no optimal source code (in the sense of Definition 1) for it and $\underline{H}(\boldsymbol{Z}) = 0$.

Appendix II

In this appendix we consider Shannon–Fano codes for two sources where we can give explicit descriptions of valid sequences of sets G_n.

First we consider the case of a biased coin with bias $p<1/2$. Let the sequence of sets G_n be as defined in (57). If an output length r_n belongs to G_n then the input sequences which map to that length r_n must have m 0's where m satisfies

$$\left|\left\lceil \log \frac{1}{p^m(1-p)^{n-m}} \right\rceil - nh(p)\right| \leq n\delta_n. \tag{77}$$

This implies

$$m \in T_n \triangleq \left\{k\colon \left|\frac{k}{n} - p\right| \leq \frac{\delta_n + 1/n}{\log\frac{1-p}{p}}\right\}. \tag{78}$$

We may have input strings with different number of 0's mapping to the same length r_n but the number of 0's has to belong to T_n. Thus using the concavity of the entropy function and the fact that sequences with the same number of 0's are equally likely we have

$$H(\phi_n(Z^n_{r_n})) \geq \min_{k\in T_n} \log \binom{n}{k}. \tag{79}$$

Since we have taken $p<1/2$ we have for sufficiently large n

$$H(\phi_n(Z^n_{r_n})) \geq \log \binom{n}{n(p - c\delta_n)} \tag{80}$$

for some $c>0$. Applying Stirling's formula to the above equation we have that

$$\lim_{n\to\infty} \frac{H(\phi_n(Z^n_{r_n}))}{r_n} = 1 \tag{81}$$

for any sequence $r_n \in G_n$. Using the remark after [3, Lemma 2.12] we also have condition [C1] in Definition 1 satisfied if

we take δ_n so that $\sqrt{n}\delta_n \to \infty$ as $n \to \infty$. Thus the rule for choosing G_n in (57) works when $\delta_n \to 0$ sufficiently slowly.

We now consider a coin with bias probability uniformly distributed in $[0,1]$. We will show that the following choice of G_n is valid for this source

$$G_n = \{r \in I_n \colon r \geq \sqrt{n}\,\}. \tag{82}$$

The probability of any particular sequence of length n with m 0's is

$$\frac{1}{(n+1)\binom{n}{m}}.$$

Thus if we have an output length r_n any sequence which maps to this length must have m 0's where m satisfies

$$r_n - 1 < \log(n+1)\binom{n}{m} \leq r_n.$$

Using the concavity of the entropy function and the fact that sequences with the same number of 0's are equally likely we have

$$H(\phi_n(Z_{r_n}^n)) \geq r_n - 1 - \log(n+1).$$

Thus if $r_n \in G_n$, then

$$\lim_{n\to\infty} \frac{H(\phi_n(Z_{r_n}^n))}{r_n} = 1. \tag{83}$$

It is easy to verify that with this choice of G_n, condition [C1] in Definition 1 is satisfied. Thus we have a valid choice of G_n.

REFERENCES

[1] R. Ahlswede, T. S. Han, and K. Kobayashi, "Universal coding of integers and unbounded search trees," *IEEE Trans. Inform. Theory*, vol. 43, pp. 669–682, Mar. 1997.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley Series in Telecommunications). New York: Wiley, 1991.

[3] I. Csiszár, and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[4] P. Elias, "The efficient construction of an unbiased random sequence," *Ann. Math. Statist.*, vol. 43, pp. 865–870, 1972.

[5] T. S. Han and M. Hoshi, "Interval algorithm for random number generation," *IEEE Trans. Inform. Theory*, vol. 43, pp. 599–611, Mar. 1997.

[6] W. Hoeffding and G. Simons, "Unbiased coin tossing with a biased coin," *Ann. Math. Statist.*, vol 41, pp. 341–352, 1970.

[7] Y. Peres, "Iterating von Neumann's procedure for generating random bits," *Ann. Statist.*, vol. 20, no. 1, pp. 590–597, 1992.

[8] R. Shack, "The length of a typical Huffman codeword," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1246–1247, July 1994.

[9] Q. Stout and B. Warren, "Three algorithms for unbiased coin tossing with a biased coin," *Ann. Probab.*, vol. 12, pp. 212–222, 1984.

[10] S. Vembu, "Information theory without ergodicity assumptions: Some new results," Ph.D. dissertation, Dept. Elec. Eng., Princeton Univ., Princeton, NJ, Mar. 1994.

[11] S. Vembu and S. Verdú, "Generating random bits from an arbitrary source: Fundamental limits," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1322–1332, Sept. 1995.

[12] J. von Neumann, "Various techniques used in connection with random digits" *Nat. Bur. Stand. Appl. Math. Ser.*, vol. 12, pp. 36–38, 1951. Reprinted in *Collected Works of von Neumann*, vol. 5.

[13] J. Ziv, "Coding theorems for individual sequences," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 389–394, July 1978.

[14] ———, "Compression, tests for randomness and estimating the statistical model of an individual sequence," *Sequences: Combinatorics, Compression, Security, and Transmission*, R. M. Capocelli, Ed. New York: Springer-Verlag, 1990.