Compressing a long range dependent renewal process

Barlas Oğuz, Venkat Anantharam
Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA 94720
Email: {barlas, ananth}@eecs.berkeley.edu

***Abstract*—Analysis of variable bit-rate video data has shown that long range dependence persists across a wide variety of codecs. While codecs are generally lossy, one may conjecture, as a partial explanation for this fact, that there exist information sources for which any lossless code results in a bit-rate process that eventually dominates a long range dependent random process. We prove this to be true for discrete time long range dependent renewal processes under a mild technical assumption.**

I. INTRODUCTION

In compressing a discrete time, finite valued, stationary, ergodic process (X_n), viewed as a model for a source of information, the random variables

$$\rho_n := \rho_n(X_{-\infty}^n) = -\log P(X_n | X_{-\infty}^{n-1})$$

are of central importance. The behavior of (ρ_n) restricts the minimum code length of lossless compression algorithms by the following lemma, [1], which is also proved in [2].

Lemma 1.1 (Barron's Lemma): Given $\{c(n), n \geq 1\}$, positive constants with $\sum_n 2^{-c(n)} < \infty$, we have

$$L_n(X_1^n) \geq -\log P(X_1^n | X_{-\infty}^0) - c(n), \text{ eventually, a.s. .} \quad (1)$$

Here $L_n(X_1^n)$ is the code length for the first n symbols of the source for some lossless coding algorithm that produces bit strings. $c(n)$ can be made logarithmic in n.

By the ergodic theorem, the limit of $\frac{1}{n}\sum_{i=1}^n \rho_i$ as $n \to \infty$ exists a.s. and equals $\eta := E[-\log P(X_1|X_{-\infty}^0)]$, i.e. the entropy rate of (X_n). This implies the following well known first order converse source coding theorem for such sources.

Theorem 1.2:

$$\liminf_n \frac{1}{n} L_n(X_1^n) \geq \eta, \text{ a.s. .}$$

Lemma 1.1 is strong enough to permit second order refinements to theorem 1.2 once we know more about the process (ρ_n). For example, in [2], it is shown that for certain short range dependent classes of sources (e.g. finite state Markov chains), $(L_n - n\eta)$ satisfies a central limit theorem.

In this paper, we are interested in the case when the source (X_n) is a stationary, long range dependent renewal process. Our main result is a second order converse source coding theorem, stating that the bit length process (L_n) will eventually dominate a long range dependent process the growth of whose variance is identical to that of (X_n), so that, in particular, it has the same Hurst exponent as (X_n). We prove this here under a mild technical assumption; the statement is made precise in section II. This result provides partial theoretical justification to existing empirical work in the field of variable bit-rate (VBR) video traffic ([3], [4], [5], [6] to cite a few). A conclusion resulting from this work is that long range dependence is omnipresent in VBR video traffic, and persists across a wide variety of codecs. Combined with these observations, our theorem backs the intuition that for many information sources long range dependence persists under compression.

Our main result, theorem 3.1, is proved in section III. Most of the work is done in proving lemma 3.8, which is itself a novel limit theorem of purely probabilistic interest. The implications for source coding, stated in theorem 2.4, follow directly from Barron's lemma.

In section IV, we provide a counter example indicating some of the difficulties in generalizing a theorem of this sort beyond renewal processes. We conclude by discussing some future directions.

II. MAIN RESULT

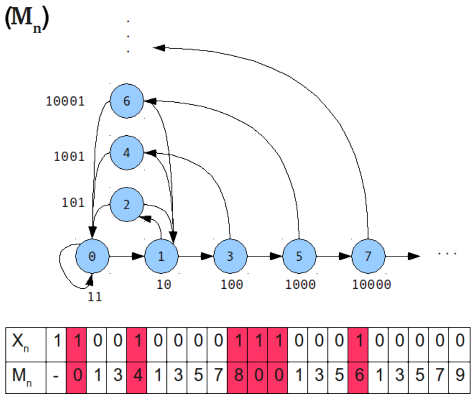
Throughout the paper (X_n) denotes a stationary renewal process with inter-arrival times having the law of $T + 1$. We write $f_T(k)$ for $P(T = k)$ and $F_T(k)$ for $P(T \leq k)$. We will assume that $f_T(k) > 0$, i.e. that the renewal process is aperiodic. We let

$$\alpha_n := \sup_{k \geq 1} \sup_{A \in \mathcal{F}_{-\infty}^k, B \in \mathcal{F}_{n+k}^\infty} |P(AB) - P(A)P(B)|$$

denote the strong mixing coefficients of (X_n). Finally, for the main result, we will assume that the stationary residual time distribution of the process has finite entropy. Note that, for any $\epsilon > 0$, this is automatic if the $(1+\epsilon)$ moment of T is finite, so this is a rather mild assumption. The aperiodicity assumption is made for convenience. The need for the finite entropy assumption appears to be an artifact of the structure of our proof.

Definition 2.1: A real valued, stationary, random process (Y_n) is said to be long range dependent (LRD) if

- $EY_1^2 < \infty$, and
- $$\limsup_n \frac{\text{var}(\sum_1^n Y_i)}{n} = \infty$$



X_n	1	1	0	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0
M_n	-	0	1	3	4	1	3	5	7	8	0	0	1	3	5	6	1	3	5	7	9

Fig. 1. Construction of the Markov chain, with an example sequence showing the correspondence with X_n

Definition 2.2: The Hurst index H, $\frac{1}{2} \le H \le 1$, of a real valued, stationary random process (Y_n) is defined by

$$H := inf\left\{h : \limsup_n \frac{\text{var}(\sum_1^n Y_i)}{n^{2h}} < \infty\right\}$$

The following lemma establishes some equivalent characterizations of long range dependence for aperiodic stationary renewal processes, viewed as $\{0,1\}$-valued processes.

Lemma 2.3: The following are equivalent:

(i) (X_n) is LRD.
(ii) $\sum \alpha_n = \infty$.
(iii) F_T has infinite variance.

Proof: We take the fact that $(ii) \iff (iii)$ from [7], section 6. $(i) \iff (iii)$ is proved in [8]. ■

Theorem 2.4: Let (X_n) be an aperiodic, long range dependent, stationary renewal process whose stationary residual time distribution has finite entropy. Then, there exists a long range dependent random process (γ_n) such that

$$L_n(X_1^n) \ge \gamma_n, \text{ eventually, a.s.}$$

for all uniquely decodable source codes. Moreover, (γ_n) has the same Hurst index as (X_n).

Proof: This immediately follows from Barron's lemma once we prove theorem 3.1 below. ■

III. PROOF OF THE MAIN RESULT

We now state and prove our main result.

Theorem 3.1: The ratio of $\text{var}(\sum_1^n X_i)$ to $\text{var}(\sum_1^n \rho_i)$ converges to a nonzero constant as $n \to \infty$.

Proof:

For the proof, we construct the following Markov chain (M_n) from the renewal process (X_n) (fig. 1):

- $M_n \in \{0,1,2,3,\ldots\}$
- $\{M_n = 0\} = \{X_{n-1}^n = 11\}$
- For $k \in \{1,2,\ldots\}$
  - $\{M_n = 2k-1\} =$ $\{X_n = 0 \text{ and } k \text{ zeros since last arrival in } X_n\}$
  - $\{M_n = 2k\} =$ $\{X_n = 1 \text{ and } k \text{ zeros since last arrival in } X_n\}$

The following definitions are from [8]:

Definition 3.2: A stationary irreducible Markov chain $(Z_n) \in \{0,1,2,3,\ldots\}$ is said to be LRD, if $1(Z_n = 0)$ is LRD in the above defined sense.

Definition 3.3: The Hurst index of a stationary irreducible Markov chain $(Z_n) \in \{0,1,2,3,\ldots\}$ is defined to be the same as the Hurst index of $1(Z_n = 0)$.

Lemma 3.4: (M_n) is long range dependent if and only if (X_n) is long range dependent. Moreover, (M_n) has the same Hurst index as (X_n).

Proof: Note that the events $\{X_{n-1} = 1\}$ and $\{M_n = 0\} \cup \{M_n = 1\}$ are identical. The claim follows from [8], section 3, prop. 1. ■

An outline of the rest of the proof is as follows. We first establish $\rho_n = \rho(M_n)$ as an instantaneous function of (M_n). This allows us to write an explicit expression for $\text{var}(\sum_1^n \rho_i)$ in terms of the transition probability matrix of (M_n), following [8]. We need to understand the asymptotic behavior of this expression, which requires an exchange of $\lim$ and sum; this is done in lemma 3.8. Since $1(X_{n-1} = 1)$ is also an instantaneous function of (M_n), as in the proof of lemma 3.4, we can similarly write an explicit expression of $\text{var}(\sum_1^n X_i)$; the asymptotics of this expression are easier to characterize, since the function involved is of finite support. The result follows from these two asymptotic characterizations.

We now proceed to flesh out this proof sketch.

A. ρ_n is L_2

One can check that $\rho_n = \rho(M_n)$ where:

- $\rho(0) = -\log f_T(0)$,
- $\rho(2k-1) = -\log P(T > k-1 | T \ge k-1)$,
- $\rho(2k) = -\log P(T = k | T \ge k)$.

Lemma 3.5: ρ_n is an L_2 functional of M_n.

Proof: Let π_i be the stationary distribution of (M_n). Note that $\pi_i > 0 \implies \rho_i < \infty$. We need to prove that

$$\sum \rho(i)^2 \pi_i < \infty \ .$$

Note that $\pi_{2k+1} = \pi_{2k-1} P(T > k | T \ge k)$, and $\pi_{2k} = \pi_{2k-1} P(T = k | T \ge k)$ for $k = 1, 2, \ldots$. This gives

$$\begin{aligned}\sum \rho(i)^2 \pi_i &= \pi_0 \rho(0)^2 + \pi_1 \rho(1)^2 \\ &\quad + \sum_{k=1}^{\infty} \pi_{2k-1} P(T = k | T \ge k) \log^2 P(T = k | T \ge k) \\ &\quad + \sum_{k=1}^{\infty} \pi_{2k-1} P(T > k | T \ge k) \log^2 P(T > k | T \ge k) \ , \\ \pi_0 \rho(0)^2 &= (\sum_{k=1}^{\infty} \pi_{2k}) f_T(0) \log^2 f_T(0) \ , \\ \pi_1 \rho(1)^2 &= (\sum_{k=1}^{\infty} \pi_{2k})(1 - f_T(0)) \log^2 (1 - f_T(0)) \ .\end{aligned}$$

Since the $p \log^2 p$ terms are bounded above by 1, $\sum \rho_i^2 \pi_i \le 4$. ■

For future reference, also note the following,

Lemma 3.6:

$$\sum_{j=2k-1}^{\infty} \rho(j)\pi_j \le \sum_{l=k}^{\infty} \pi_1 P(T > l)$$

Proof:

$$\begin{aligned}
&\sum_{j=2k-1}^{\infty} \rho(j)\pi_j = \\
&\qquad - \sum_{l=k}^{\infty} \pi_{2l-1} P(T = l | T \ge l) \log P(T = l | T \ge l) \\
&\qquad - \sum_{l=k}^{\infty} \pi_{2l-1} P(T > l | T \ge l) \log P(T > l | T \ge l) \\
&\qquad \le \sum_{l=k}^{\infty} \pi_{2l-1} = \frac{\pi_1}{P(T \ge 1)} \sum_{l=k}^{\infty} P(T \ge l)
\end{aligned}$$

since $-p \log p \le \frac{1}{2}$. ∎

B. Variance of the partial sums

Following [9] for the (standard) notation, for the Markov chain (M_n), let $p_{ij}^{(n)}, n \ge 0$, denote the probability of going from state i to j in exactly n steps, where $p_{ij}^{(0)} := \delta_{ij}$, and $f_{ij}^{(n)}, n \ge 1$, the probability of going from state i to j for the first time in exactly n steps. Subscripts to the left of a probability indicate taboo states, e.g. ${}_1p_{ij}^{(n)}, n \ge 1$, is the probability of going from i to j in n steps without visiting 1 at times 1 through $n-1$. m_{ij} is the mean time of getting from i to j. We also use the following notation from [8]:

$$Q_{ij}^{(n)} := \sum_{r=1}^{n} (p_{ij}^{(r)} - \pi_j) \ , \quad n \ge 1 \ .$$

$$R_{ij}^{(n)} := \sum_{r=1}^{n} Q_{ij}^{(r)} \ , \quad n \ge 1 \ .$$

From [8] and because (X_n) is aperiodic and LRD, one can conclude that the following holds for (M_n).

$$\lim_{n\to\infty} Q_{ij}^{(n)} = \infty \tag{2}$$

$$\lim_{n\to\infty} \frac{R_{ij}^{(n)}}{n} = \infty \tag{3}$$

$$\lim_{n\to\infty} \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} = 1 \tag{4}$$

(4) is eq. 8 in [8]. (2) follows from eqs. 8 and 5, [8]. (3) is implied by (2).

Lemma 3.7: (Carpio & Daley [8], 3.13) For any L_2 instantaneous function $\psi_n := \psi(M_n)$, the following holds,

$$\begin{aligned}
&\frac{\text{var}(\psi_0 + \ldots + \psi_n) - (n+1)\text{var}(\psi_0)}{2R_{kk}^{(n)}/\pi_k} = \\
&\qquad \sum_i \sum_j \psi(i)\psi(j)\pi_i\pi_j \frac{R_{ij}^{(n)}/\pi_j}{R_{kk}^{(n)}/\pi_k}
\end{aligned} \tag{5}$$

For our purposes, we set $k = 1$, and we want to consider both the instantaneous functions ρ_n and $1(M_n \in \{0, 1\})$. Using (4) for fixed i and j, when ψ has finite support, as in the latter case, we can conclude, as in [8], that the limit of the left hand side of equation (5) is $\sum_i \sum_j \psi(i)\psi(j)\pi_i\pi_j$. For the former case we would like to show:

Lemma 3.8:

$$\lim_{n\to\infty} \frac{\text{var}(\rho_0 + \ldots + \rho_n)}{2R_{11}^{(n)}/\pi_1} = \sum_i \sum_j \rho(i)\rho(j)\pi_i\pi_j$$

Proof: Unfortunately, the convergence in (4) is far from uniform in i and j. Thus, there is a difficulty in exchanging the limit in n with the infinite sum. For this we bound $\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}$ from above and below and use the dominated convergence theorem. Before getting into the details, note that the proof of lemma 3.8 will complete the proof of theorem 3.1, as discussed in the earlier proof sketch.

1) Upper bound: We will deal with even and odd states differently. Also, the states 0 and 1 are special, and need to be treated separately. We cover all these cases below. The cases $\{i = 0, j = 0\}$, $\{i = 0, j = 1\}$, $\{i = 1, j = 0\}$ and $\{i = 1, j = 1\}$ need not be considered, since we know $\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}$ is bounded in n for fixed i and j by (4).

- $\{$Case $i > 1, j = 1\}$: For this case, we simply have $\sum_{r=1}^{n} p_{i1}^{(r)} = \sum_{r=1}^{n} \sum_{m=1}^{r} f_{i1}^{(m)} p_{11}^{(r-m)} = \sum_{m=1}^{n} f_{i1}^{(m)} \sum_{r=m}^{n} p_{11}^{(r-m)} = \sum_{m=1}^{n} f_{i1}^{(m)} \sum_{r=0}^{n-m} p_{11}^{(r)} \le \sum_{r=1}^{n} p_{11}^{(r)} + 1$.
- $\{$Case $i > 1, j = 0\}$: Note that for any state i, $p_{i0}^{(r)} = \frac{f_T(0)}{1-f_T(0)} p_{i1}^{(r)} = \frac{\pi_0}{\pi_1} p_{i1}^{(r)}$. Therefore we get $\sum_{r=1}^{n} p_{i0}^{(r)} = \frac{\pi_0}{\pi_1} \sum_{r=1}^{n} p_{i1}^{(r)} \le \frac{\pi_0}{\pi_1} \sum_{r=1}^{n} p_{11}^{(r)} + \frac{\pi_0}{\pi_1}$.
  For the rest, we write,
  $$\sum_{r=1}^{n} p_{ij}^{(r)} = \sum_{r=1}^{n} {}_1p_{ij}^{(r)} + \sum_{r=1}^{n} \sum_{m=1}^{r} f_{i1}^{(m)} \sum_{l=1}^{r-m} p_{11}^{(r-m-l)} \, {}_1f_{1j}^{(l)} \ . \tag{6}$$
- $\{$Case $i = 0, j > 1\}$: Since it is impossible to go to j from 0 without going through 1, $\sum_{r=1}^{n} {}_1p_{0j}^{(r)} = 0$. Also, there is only a single path to go from state 1 to state j (say in l' steps), and ${}_1f_{1j}^{(l')} = \frac{\pi_j}{\pi_1}$. Now (6) becomes $\sum_{r=1}^{n} p_{0j}^{(r)} = \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} \sum_{m=1}^{r-l'} f_{01}^{(m)} p_{11}^{(r-m-l')} = \frac{\pi_j}{\pi_1} \sum_{m=1}^{n-l'} f_{01}^{(m)} \sum_{r=m+l'}^{n} p_{11}^{(r-m-l')} = \frac{\pi_j}{\pi_1} \sum_{m=1}^{n-l'} f_{01}^{(m)} \sum_{r=0}^{n-m-l'} p_{11}^{(r)} \le \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} p_{11}^{(r)} + \frac{\pi_j}{\pi_1}$.
- $\{$Case $i = 2k - 1, j > 1\}$ $(k = 1, 2, \ldots)$: Going to j from i without passing 1 is only possible if $j > i$, with probability $\frac{\pi_j}{\pi_i}$. Therefore we have $\sum_{r=1}^{n} {}_1p_{ij}^{(r)} = \frac{\pi_j}{\pi_i} 1(j > i)$ for all sufficiently large n. Also, following the first case, $\sum_{m=1}^{r} f_{i1}^{(m)} \sum_{l=1}^{r-m} p_{11}^{(r-m-l)} \, {}_1f_{1j}^{(l)} = \frac{\pi_j}{\pi_1} \sum_{m=1}^{r-l'} f_{i1}^{(m)} p_{11}^{(r-m-l')}$. The bound becomes $\sum_{r=1}^{n} p_{ij}^{(r)} \le \frac{\pi_j}{\pi_i} 1(j > i) + \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} \sum_{m=1}^{r-l'} f_{i1}^{(m)} p_{11}^{(r-m-l')} \le \frac{\pi_j}{\pi_i} 1(j > i) + \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} p_{11}^{(r)} + \frac{\pi_j}{\pi_1}$.

- {Case $i = 2k, j > 1$} ($k = 1, 2, \ldots$): Here $\sum_{r=1}^{n} {}_1 p_{ij}^{(r)} = 0$. Everything else is identical to the last case, so $\sum_{r=1}^{n} p_{ij}^{(r)} \le \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} p_{11}^{(r)} + \frac{\pi_j}{\pi_1}$.

To summarize, for $(i, j) \neq (0,0), (0,1), (1,0), (1,1)$,

$$\sum_{r=1}^{n} p_{ij}^{(r)} \le \frac{\pi_j}{\pi_i} 1(j > i, i \text{ odd}) + \frac{\pi_j}{\pi_1} + \frac{\pi_j}{\pi_1} \sum_{r=1}^{n} p_{11}^{(r)} \ . \tag{7}$$

Subtracting $n\pi_j$ from both sides and dividing by $\frac{\pi_j}{\pi_1} \sum (p_{11}^{(r)} - \pi_1)$ (using (2) take n large enough s.t. this term is positive),

$$\frac{\sum_{r=1}^{n} (p_{ij}^{(r)} - \pi_j)/\pi_j}{\sum_{r=1}^{n} (p_{11}^{(r)} - \pi_1)/\pi_1} \le \frac{\frac{\pi_1}{\pi_i} 1(j > i, i \text{ odd}) + 1}{\sum_{r=1}^{n} (p_{11}^{(r)} - \pi_1)} + 1 \ . \tag{8}$$

Note that we derived a bound for $\frac{Q_{ij}^{(n)}/\pi_j}{Q_{11}^{(n)}/\pi_1}$. However, since the denominator is positive for sufficiently large n, the same bound holds for $\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}$.

We substitute this bound into lemma 3.7, to verify that it yields a finite sum (assume n large enough s.t. $Q_{11}^{(n)} > 1$). Here C denotes a sufficiently large constant, which can change from line to line.

$$\sum_i \sum_j \rho(i)\rho(j)\pi_i\pi_j \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} \le \tag{9}$$

$$\sum_i \sum_j \rho(i)\rho(j)\pi_i\pi_j \Big(\frac{\frac{\pi_1}{\pi_i} 1(j > i, i \text{ odd}) + 1}{\sum_{r=1}^{n} (p_{11}^{(r)} - \pi_1)} + 1\Big) + C \tag{10}$$

$$= \sum_{k=1}^{\infty} \sum_{j=2k-1}^{\infty} \rho(2k-1)\rho(j)\pi_j \frac{\pi_1}{\sum_{r=1}^{n} (p_{11}^{(r)} - \pi_1)} \tag{11}$$

$$+ \sum_i \sum_j \rho(i)\rho(j)\pi_i\pi_j \Big(1 + \frac{1}{\sum (p_{11}^{(r)} - \pi_1)}\Big) + C \tag{12}$$

$$= C \sum_{k=1}^{\infty} \rho(2k-1) \sum_{j=2k-1}^{\infty} \rho(j)\pi_j + C \ , \tag{13}$$

$$\le C \sum_{k=1}^{\infty} \rho(2k-1) \sum_{l=k}^{\infty} P(T \ge l) + C \tag{14}$$

$$= C \sum_{l=0}^{\infty} P(T > l) \sum_{k=1}^{l} \rho(2k-1) + C \tag{15}$$

$$= C \sum_{l=0}^{\infty} -P(T > l) \log P(T > l) + C \ , \tag{16}$$

and the right hand side is finite by the assumption of finite entropy for the stationary residual time. The step from (13) to (14) used lemma (3.6).

2) Lower bound: The upper bound implies a uniform integrability condition,

$$\sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \left[\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}\right]_+ \to 0, \quad \text{as } M \to \infty$$

uniformly in n. Define $\overline{\rho}^M(i) = \rho(i) 1(i \le M)$ and $\underline{\rho}^M(i) = \rho(i) - \overline{\rho}^M(i)$. From lemma 3.7 we have

$$\frac{\mathrm{var}(\sum_{t=0}^{n} \rho_t) - (n+1)\mathrm{var}(\rho_0)}{2R_{11}^{(n)}/\pi_1} - \frac{\mathrm{var}(\sum_{t=0}^{n} \overline{\rho}_t^M) - (n+1)\mathrm{var}(\overline{\rho}_0^M)}{2R_{11}^{(n)}/\pi_1}$$

$$= \sum_{i,j} \rho(i)\rho(j)\pi_i\pi_j \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} - \sum_{i,j \le M} \rho(i)\rho(j)\pi_i\pi_j \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}$$

$$= \sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} \tag{17}$$

Briefly adopt the shorthand notation:

$$\phi_n := \frac{(\rho_0 + \ldots + \rho_n) - (n+1)E\rho_0}{\sqrt{2R_{11}^{(n)}/\pi_1}}$$

$$\overline{\phi}_n^M := \frac{(\overline{\rho}_0^M + \ldots + \overline{\rho}_n^M) - (n+1)E\overline{\rho}_0^M}{\sqrt{2R_{11}^{(n)}/\pi_1}}$$

$$\underline{\phi}_n^M := \phi_n - \overline{\phi}_n^M$$

In this notation (17) gives

$$\left| \sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} \right| < |\mathrm{var}(\phi_n) - \mathrm{var}(\overline{\phi}_n^M)| + \frac{\pi_1(n+1)}{2R_{11}^{n}} |\mathrm{var}(\rho_0) - \mathrm{var}(\overline{\rho}_0^M)| \tag{18}$$

Then by the Cauchy-Schwartz inequality,

$$(\mathrm{var}(\phi_n) - \mathrm{var}(\overline{\phi}_n^M))^2 = (E(\phi_n - \overline{\phi}_n^M)(\phi_n + \overline{\phi}_n^M))^2 \tag{19}$$

$$\le \mathrm{var}(\underline{\phi}_n^M)\mathrm{var}(\phi_n + \overline{\phi}_n^M).$$

We know by the upper bound that $\mathrm{var}(\phi_n + \overline{\phi}_n^M)$ is bounded above by a constant. Also, since $\frac{n}{R_{11}^n}$ is bounded and we have

$$\mathrm{var}(\underline{\phi}_n^M) = \frac{(n+1)\mathrm{var}(\underline{\rho}_0^M)}{2R_{11}^{(n)}/\pi_1} + \sum_{i,j>M} \pi_i\pi_j\rho(i)\rho(j) \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}$$

we have $\mathrm{var}(\underline{\phi}_n^M) \to 0$ as $M \to \infty$ uniformly in n, again by the upper bound.

We have, by definition

$$\sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \left[\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}\right]_+$$

$$- \sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \left[\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}\right]_- =$$

$$\sum_{\{i>M\} \cup \{j>M\}} \pi_i\pi_j\rho(i)\rho(j) \frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1} \ .$$

Thus,

$$\sum_{\{i>M\}\cup\{j>M\}} \pi_i\pi_j\rho(i)\rho(j)\left[\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}\right]_-$$

$$< \sum_{\{i>M\}\cup\{j>M\}} \pi_i\pi_j\rho(i)\rho(j)\left[\frac{R_{ij}^{(n)}/\pi_j}{R_{11}^{(n)}/\pi_1}\right]_+$$

$$+C_1\sqrt{\mathrm{var}(\underline{\phi}_n^M)} + C_2|\mathrm{var}(\rho_0) - \mathrm{var}(\bar{\rho}_0^M)| \ .$$

All three terms on the right hand side go to zero as $M \to \infty$, uniformly in n. ■

This concludes the proof of theorem 3.1. ■

IV. GENERALIZATIONS: A COUNTER-EXAMPLE

Fix $1 < p < 2$. One might hope for a general result like:

Conjecture 4.1: Let (X_n) be a finite real valued, aperiodic, stationary, ergodic process with mixing coefficients α_n. Then

$$\sum n^{p-2}\alpha_n = \infty \implies (\rho_n) \text{ is LRD with } H \geq \frac{1}{2}(3-p)$$

The conjectured relation between H and p is motivated by theorem 6.1 of [7], theorem 1 of [10] and our theorem 3.1 above. The following counterexample kills this conjecture.

Let $X_n \in \{A_1, A_2, B_1, B_2\}$. Define events at each n,

$$e_{ab}^n = \{\text{A transition occurs between } \{A_1, A_2\}, \{B_1, B_2\} \text{ at } n\}$$

$$e_{12}^n = \{\text{A transition occurs between } \{A_1, B_1\}, \{A_2, B_2\} \text{ at } n\}$$

Fix $1 < p < 2$ and construct (X_n) such that $R_n = 1(X_n \in \{A_1, A_2\})$ defines a LRD renewal process with Hurst exponent strictly bigger $\frac{1}{2}(3-p)$. This defines the events e_{ab}^n. Theorem 6.1 of [7], theorem 1 of [10] imply that the mixing coefficients of (R_n) satisfy the hypothesis of the conjecture, hence the mixing coefficients of (X_n) (which is to be constructed) must also satisfy this hypothesis, by inheritance of strong mixing, see e.g. [11]. We will define e_{12}^n to be independent of e_{ab}^n given $X_{-\infty}^{n-1}$. Moreover, set $P(e_{12}^n|X_{-\infty}^{n-1}) = h^{-1}(1 - E[-\log P(e_{ab}^n|X_{-\infty}^{n-1})|X_{-\infty}^{n-1}])$. Here $h^{-1}(\eta)$ is the inverse of $h(p) = -p\log p - (1-p)\log(1-p)$ on $p \in [0, 0.5]$. Then $E[-\log P(e_{ab}^n|X_{-\infty}^{n-1})|X_{-\infty}^{n-1}] + E[-\log P(e_{12}^n|X_{-\infty}^{n-1})|X_{-\infty}^{n-1}] = 1$. Thus, we have forced $(\rho_n - 1)$ to be a martingale difference sequence. Therefore the conclusion of the conjecture cannot hold.

In this example, we have constructed (X_n) from two LRD processes, coupled in such a way that their effect on the entropy rate balance each other. Clearly, this is a contrived example. For many useful models such a construction ought not to be possible. For example, we expect our theorem to extend to semi-Markov and generalized semi-Markov processes.

V. CONCLUSION

We have shown for the class of long range dependent renewal sources, under a mild technical assumption, that the bit-rate process of every source code must dominate a long range dependent random process.

Our proof can be thought of as a counterpart to the numerous results in the probability literature which show that a random process derived from a short range dependent process must itself exhibit short range dependence. Therefore, it would be of considerable interest to generalize our results further to complement those already existing. To this end, we have argued that an analogous result in full generality might not exist, but extensions to many other important classes of stochastic models should be possible.

Clearly, our result would not extend to codes with distortion without qualification. An interesting problem would be to analyze the LRD bit-rate vs. LRD distortion trade-off in this setting, in some suitably defined sense, i.e. to quantify "how much" of the long range dependence must persist under lossy coding, as a function of the target distortion.

ACKNOWLEDGMENT

The research of the authors was supported by the ARO MURI grant W911NF-08-1-0233, "Tools for the Analysis and Design of Complex Multi-Scale Networks" and by the NSF grants CCF-0500234, CCF-0635372 and CNS-0627161.

REFERENCES

[1] A. Barron, "Logically smooth density estimation." Ph.D. dissertation, Stanford University, Department of Electrical Engineering, 1985.

[2] I. Kontoyiannis, "Second-order noiseless source coding theorems," *IEEE Transactions on Information Theory*, vol. 43, no. 4, pp. 1339–1341, 1997.

[3] J. Beran, R. Sherman, M. Taqqu, and W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 1566–1579, 1995.

[4] M. Garrett and W. Willinger, "Analysis, modeling and generation of self-similar VBR video traffic," *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 4, pp. 269–280, 1994.

[5] O. Rose, "Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems," in *Conference on Local Computer Networks: Proceedings.* Institute of Electrical and Electronics Engineers, 1995, p. 397.

[6] F. Fitzek and M. Reisslein, "Mpeg4 and h. 263 video traces for network performance," *IEEE Network*, vol. 15, no. 6, pp. 40–54, 2001.

[7] H. Berbee, "Convergence rates in the strong law for bounded mixing sequences," *Probability Theory and Related Fields*, vol. 74, no. 2, pp. 255–270, 1987.

[8] K. Carpio and D. Daley, "Long-range dependence of Markov chains in discrete time on countable state space," *Journal of Applied Probability*, vol. 44, no. 4, pp. 1047–1055, 2007.

[9] K. Chung, *Markov chains with stationary transition probabilities*. Berlin: Springer-Verlag, 1967.

[10] D. Daley, "The Hurst index of long-range dependent renewal processes," *Annals of Probability*, pp. 2035–2041, 1999.

[11] R. Bradley, "Basic properties of strong mixing conditions. A survey and some open questions," *Probability Surveys*, vol. 2, pp. 107–144, 2005.