On the Minimum Redundancy of Homophonic Coding

Valdemar C. da Rocha Jr., vcr@npd.ufpe.br

Comms. Res. Group - UFPE, P.O. Box 7800

50711-970 Recife PE, BRAZIL

Abstract—*This paper establishes the condition for the Rocha-Massey homophonic coding scheme to have lower redundancy than the Jendal-Kuhn-Massey scheme, for sources whose symbol probabilities are rational numbers. The condition is that each probability, in the probability distribution of the symbols of a discrete memoryless source, must contain only countably infinite binary expansions. Both schemes are briefly reviewed and the condition for achieving the minimum redundancy $h(2^{-m})$ in a perfect homophonic substitution scheme is determined, where $h(.)$ denotes the binary entropy function and m is the least positive integer for which all probability entries in the expanded alphabet of the Rocha-Massey scheme have denominators which are powers of 2.*

I. Introduction

Homophonic coding is a technique whereby a multiplicity of "homophones" are probabilistically substituted for each plaintext letter. Each homophone is one-to-one mapped to a binary word so as to hide the redundancy of the resulting new "plaintext". In secret-key cryptographic systems the use of homophonic coding increases the unicity distance of the cipher [1], and hence makes it harder to break, at the cost of some plaintext expansion.

In traditional homophonic coding, each letter of the original message is replaced by a substitute or *homophone* in a larger alphabet to form the plaintext message that is then encrypted. The homophone is chosen uniformly at random from a set of substitutes reserved in the larger alphabet for that letter in the original alphabet, the size of these sets being (roughly) proportional to the relative frequencies of the letters in the original alphabet. Consequently, the overall probability distribution of homophones is a uniform distribution and all homophones have the same length. In 1988, Günther [2] introduced *variable-length homophonic substitution*, in which the homophones for a particular letter can have different lengths and different probabilities of selection, and showed that this technique could be used to hide in the homophonically coded sequence all the redundancy in the original plaintext and thus can be used to construct what Shannon calls a *strongly ideal* cipher [1] while also reducing plaintext expansion.

The purpose of this paper is to prove a necessary and sufficient condition for the Rocha-Massey (RM) homophonic coding scheme [5] to have lower redundancy than the Jendal-Kuhn-Massey (JKM) scheme [3], for sources whose symbol probabilities are rational numbers. The condition is that each probability, in the probability distribution of the symbols of a discrete memoryless source, must contain only countably infinite binary expansions. Both schemes are briefly reviewed and the condition for achieving the minimum redundancy $h(2^{-m})$ in a perfect homophonic substitution scheme is determined, where $h(.)$ denotes the binary entropy function and m is a positive integer determined by the source probability distribution. Both homophonic schemes require finite memory in their respective implementations despite the fact that an unbounded number of coin tosses to select a homophone is required. The finite memory description of the JKL scheme follows that of [6].

II. Theoretical Background

For simplicity, we consider only the homophonic coding of the output sequence $U_1, U_2, U_3, \ldots$ of a K-ary discrete *memoryless* source (DMS). The homophonic coding problem then reduces to that for a single K-ary random variable U, but the theory is easily modified to handle general sources with memory simply by replacing the probability distribution for U_i with the conditional probability distribution for U_i given the observed values of $U_1, U_2, \ldots, U_{i-1}$. We assume that the source U has a probability distribution with rational entries only, $P_U(u_i) = m_i/n_i,\ 1 \leq i \leq K$, where m_i and n_i are positive integers and n_i is as small as possible. If n_i is an integer power of 2, JKM optimum homophonic coding operates with a finite upper bound on the number of fair coin flips required to select a homophone. Our interest is in the case where n_i is not a power of two so that there is no upper bound on the number of coin flips that may be required. We assume with no loss of essential generality that all K values of U have non-zero probability and that $K \geq 2$. The homophone V for U takes values in the set $\{v_1, v_2, \ldots \quad\}$, which may be finite or countably infinite, and is characterized by the fact that for each j there is exactly one i such that $P(V = v_j | U = u_i) \neq 0$. For a binary variable-length homophonic encoding, $V = (X_1, X_2, \ldots, X_W)$ where X_i is a binary random variable and where the length W of the homophone is in general also a random variable. It is required that the homophones be assigned in such a manner that $X_1 X_2 \ldots X_W$ is a prefix-free encoding of V, i.e., such that the homophones $v = (x_1, x_2, \ldots, x_w)$ are all distinct and none is the prefix of another. Homophonic coding is *perfect* if the new plaintext sequence is irredundant, i.e., if the components $X_1, X_2, \ldots, X_W$ of the homophone V are independent and uniformly distributed binary random variables, the homophonic coding is said to be *perfect*, and is said to be *optimum* if it is perfect and its *plaintext expansion* (defined as the average length of a homophone, $E[W]$, less the entropy of a source letter) is as small as possible [3]. The upper bound $E[W] - H(U) < 2$ bits on the redundancy with optimum binary homophonic substitution was stated in [3] and proved in [4].

It was shown by JKM in [3, *Proposition 3*] that a binary ho-

The author acknowledges partial support of this work by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant No. 304214/77-9.

mophonic coding scheme is optimum if and only if, for every value u_i of U, the conditional probabilities $P(V = v_j|U = u_i)$ of the homophones for u_i are such that the probabilities $P(V = v_j) = P(V = v_j, U = u_i) = P(V = v_j|U = u_i)P(U = u_i)$ of these homophones are equal (in some order) to the terms in the unique decomposition of $P(U = u_i)$ as a finite sum of distinct negative integer powers of 2 when this is possible, i.e., when $P(U = u_i)$ can be written as a ratio of integers in which the denominator is an integer power of 2, and as an infinite sum of distinct negative integer powers of 2 otherwise. In the former case, the homophone for a source letter can be selected using the results of at most as many flips of a fair coin as the exponent of 2 in the denominator of $P(U = u_i)$, while in the latter case there is no bound on the number of flips that may be required [3].

III. The Rocha-Massey Scheme

Rocha and Massey (RM) [5] introduced a perfect homophoning coding scheme for which better than "optimum" homophonic coding was possible, i.e., a homophonic coding scheme which in many cases outperformed the standard JKL scheme, under the assumption that the source U has a probability distribution with rational entries only. A lower bound on redundancy was then established and the conditions for meeting this bound were proved. Let n denote the least common denominator of these rational probabilities, i.e., let $P_U(u_i) = m_i/n,\ 1 \le i \le K$, where m_i and n are positive integers and n is as small as possible. If n is an integer power of 2, JKM optimum homophonic coding operates with a finite upper bound on the number of fair coin flips required to select a homophone. The RM scheme is focused in the case where n is not a power of two so that there is no upper bound on the number of coin flips that may be required. Hereafter then, we assume that $\lceil \log n \rceil > \log n$ where $\lceil \log n \rceil$ denotes the smallest integer at least equal to $\log n$ and where all logarithms are to the base 2. The "trick" in the RM scheme is to augment the source U with a "dummy" letter chosen so that all letters of the augmented source can be written as rational numbers with a common denominator that is a power of 2. Let Δ denote the "dummy" letter so that $\{u_1, u_2, \dots, u_K, \Delta\}$ is the output alphabet of the augmented source that we denote by $\tilde{U}$ and let $P_{\tilde{U}}(\Delta) = 2^{-s}$, where s is the least integer such that $2^s - 1$ is divisible by the product n' of the odd factors of n. This choice forces us then to choose $P_{\tilde{U}}(u_i) = (1 - 2^{-s})P_U(u_i) = (r_i/2^s)/(n'/n)$ for $1 \le i \le K$, where $r_i = (2^s - 1)m_i/n'$ is an integer. Thus, the letters of the augmented source all have probabilities that are rational numbers with a common denominator of $2^m n/n'$. It follows that at most $m + \log(n/n')$ fair coin flips will be required to choose the homophone if optimum standard homophonic coding is now applied to the output of this augmented source.

Example 1: Let U be the binary memoryless source with $P_U(u_1) = 1/3$ and $P_U(u_2) = 2/3$. The JKM homophonic coding uses an infinite decomposition of both $P_U(u_1)$ and $P_U(u_2)$ as sums of negative powers of 2 and results in an expected codeword length $E[W] = 2$. The redundancy is $E[W] - H(U) = E[W] - h(1/3) = 2 - 0.918 = 1.082$, where $h(.)$ denotes the binary entropy function. Augmenting this source with a dummy letter Δ of probability $P_{\tilde{U}}(\Delta) = 1/4$, we have $P_{\tilde{U}}(u_1) = (3/4)(1/3) = 1/4$ and $P_{\tilde{U}}(u_2) = (3/4)(2/3) = 1/2$ so that at most two fair coin flips are needed to select any homophone. The coding of $\tilde{U}$ is irredundant since its letters all have probabilities that are negative integer powers of 2 so that $E[\tilde{W}] = H(\tilde{U})$ in this example. The average number of letters from the source U that are encoded with the encoding of one letter of the source $\tilde{U}$ is p. The redundancy relative to the source U of the new scheme is thus $E[\tilde{W}] - pH(U) = H(\tilde{U}) - pH(U) = 3/2 - (3/4)h(1/3) = 0.811$, which is substantially better than the redundancy 1.082 for the JKM homophonic coding of U.

A. Implementing the RM Scheme

The RM scheme could be implemented as follows. One first uses a fair coin to test for the occurrence of an event of probability $P_{\tilde{U}}(\Delta) = 2^{-m}$, which requires at most m flips of the fair coin. If the event occurs, the dummy letter Δ becomes the output of $\tilde{U}$. Otherwise, one calls on the source U to emit a letter that then becomes the output of $\tilde{U}$. Decoding is simple–one just removes the dummy letters from the reconstructed output sequence of $\tilde{U}$ to obtain the output sequence of U.

B. Bounds on Redundancy

The following proposition was proved in [5] and is reproduced here for completeness.

Proposition 1: Let U be a K-ary discrete memoryless source whose letter probabilities are all rational numbers, let n be the least common denominator of these probabilities when written as reduced fractions, let $N = \lceil \log_2 n \rceil$, and let $p = n/2^N$. Then the homophonic coding of the augmented source $\tilde{U}$ as described above achieves a redundancy $E[\tilde{W}] - pH(U)$, when $p \ne 1$, satisfying

$$h(p) \le E[\tilde{W}] - pH(U) < h(p) + 2 - 2^{3-N},\ \text{all}\ N \ge 3, \quad (1)$$

and satisfying the equality

$$E[\tilde{W}] - pH(U) = h(p) \quad (2)$$

whenever (1) the letter probabilities of U written as fractions with denominator n all have numerators that are integer powers of 2 and (2) $n = 2^N - 2^i$ for some i with $0 \le i \le N - 2$ (which two conditions are always satisfied when $N = 2$).

Example 2: Consider the DMS U with letter probabilities $2/3, 1/6$ and $1/6$. Here $n = 6$ and $N = 3$. The two conditions for (2) to hold are satisfied so the redundancy is $E[\tilde{W}] - pH(U) = h(p) = h(3/4) = 0.8113$.

IV. The Modified JKM Scheme

So far the main complaint against the practical use of the JKM scheme was due to the apparent need to store a dictionary with an unbounded number of homophonic codewords, and less efficient but "more practical" solutions have appeared [7] [8] in the literature. We review next the modified JKM scheme as introduced in [6]. Essentially the modified JKM scheme sequentially constructs each homophonic codeword, as a concatenation of shorter codewords appropriately selected from a finite set of codewords derived from the source probabilities. It turns out that for any probability distribution with only rational number entries a sequential implementation of the JKM scheme exists with a finite number of binary codewords. The modified

JKM scheme avoids the memory problem of having to store a countably infinite number of homophonic codewords in the case where n_i is not a power of 2.

Initially, each source probability is expanded as a sum of negative powers of 2 and then we identify in each such expansion the periodic and the non-periodic components, where by a periodic component we mean the set consisting of the terms of an infinite geometric series whose first term and ratio are both negative powers of 2, and by a non-periodic component we mean the set consisting of a finite number of negative powers of 2.

Example 3: The expansion of $P_U(u_1)$ is $P_U(u_1) = 1/5 = \sum_{i=0}^{\infty}(1/8)(1/16)^i + \sum_{i=0}^{\infty}(1/16)(1/16)^i$, i.e., $1/5$ is decomposed as a sum of two periodic components (two infinite geometric series) whose first terms are $1/8$ and $1/16$, respectively, and both series have the same ratio $1/16$.
Next we construct a binary prefix-free code C using as codeword probabilities the non-periodic components, and the first term and ratio of each periodic component, resulting from the decomposition of the source probabilities, with the remark that identical ratios are associated to the same codeword.

Example 4: Consider the $K = 2$ DMS with $P_U(u_1) = 1/5$ and $P_U(u_2) = 4/5$. We have earlier obtained the expansion of $P_U(u_1)$ and notice that $P_U(u_2) = 4/5 = \sum_{i=0}^{\infty}(1/2)(1/16)^i + \sum_{i=0}^{\infty}(1/4)(1/16)^i$. Since the ratio $1/16$ is the same for both infinite geometric series in the expansion of $P_U(u_1) = 1/5$ as well as for the two infinite geometric series in the expansion of $P_U(u_2) = 4/5$, only one codeword is allocated to this common ratio. Let S_1 denote the sum $\sum_{i=0}^{\infty}(1/8)(1/16)^i$ and let S_2 denote the sum $\sum_{i=0}^{\infty}(1/16)(1/16)^i$. It follows that $S_1 = 2/15$ and that $S_2 = 1/15$. Given that $U = u_1$, the probability of selecting a term from the sum S_1 is $(2/15)/(1/5) = 2/3$ and the probability of selecting a term from the sum S_2 is $1/3$. Next we use the non-periodic components, and the first term and ratio of each periodic component, resulting from the decomposition of the source probabilities, to construct a binary prefix-free code C, with the remark that identical ratios are associated to the same codeword. We have thus the set of probabilities $\{1/2, 1/4, 1/8, 1/16, 1/16\}$ containing the first terms and the common ratio, which is then used to construct the following set of codewords $\{0, 10, 110, 1110, 1111\}$, respectively.
For example, consider the $K = 2$ DMS with $P_U(u_1) = 1/5$ and $P_U(u_2) = 4/5$. We have earlier obtained the expansion of $P_U(u_1)$ and notice that $P_U(u_2) = 4/5 = \sum_{i=0}^{\infty}(1/2)(1/16)^i + \sum_{i=0}^{\infty}(1/4)(1/16)^i$. The ratio $1/16$ is common to the resulting infinite geometric series in the expansion of both $P_U(u_1) = 1/5$ and $P_U(u_2) = 4/5$, therefore only one codeword is allocated to this common ratio. We have thus the set of probabilities $\{1/2, 1/4, 1/8, 1/16, 1/16\}$ containing the first terms and the common ratio, which is then used to construct the following set of codewords $\{0, 10, 110, 1110, 1111\}$, respectively. Each term in the base 2 expansion of a given source probability is associated one-to-one with a homophone in a set denoted by V.

In general, one first calls on the source U to emit a letter, say $u_i,\ 1 \leq i \leq K$, and then performs an experiment to select a homophone associated with u_i. Here we introduce a slight but significant variation on the approach followed in the JKM scheme. Instead of considering from the start the set of all homophones corresponding to $U = u_i$, we partition this set into subsets. Each non-periodic term corresponds to a subset with a single homophone and each periodic term corresponds to a subset with a countably infinite number of homophones. Our selection of a homophone is done in two parts, first we select a subset and then we select a homophone within this subset. Suppose that the subset selected corresponds to a non-periodic component.The corresponding codeword in C is made the output of V. Suppose now that the subset selected corresponds to a periodic component, say the j^{th} periodic component, $j = 1, 2, \ldots, J$, in the base 2 expansion of $P_U(u_i)$.

A. Implementing the Modified JKM Scheme

We assume that the source U has a probability distribution with rational entries only, $P_U(u_i) = m_i/n_i,\ 1 \leq i \leq K$, where m_i and n_i are positive integers and n_i is as small as possible. If n_i is an integer power of 2, standard optimum homophonic coding operates with a finite upper bound on the number of fair coin flips required to select a homophone. Our interest is in the case where n_i is not a power of two so that there is no upper bound on the number of coin flips that may be required. We recall from [3] that for achieving lowest redundancy with binary homophonic coding all the terms in the base 2 expansion of any given source probability must be distinct and that each such term is associated one-to-one with a homophone in a set denoted by V. In general, one first calls on the source U to emit a letter, say $u_i,\ 1 \leq i \leq K$, and then performs an experiment to select a homophone associated with u_i. Suppose that the homophone to be selected corresponds to one non-periodic component. The corresponding codeword in C is made the output of V. Suppose now that the homophone to be selected corresponds to one of the terms of the j^{th} periodic component, $j = 1, 2, \ldots$, in the base 2 expansion of $P_U(u_i)$. A binary experiment is then performed whose outcomes E_j and $\overline{E_j}$ have probabilities $P(E_j) = 1 - P(\Delta_{ij})$ and $1 - P(E_j) = P(\Delta_{ij})$, respectively, where $P(\Delta_{ij})$ is a negative power of 2 equal to the ratio in the corresponding geometric series. We shall refer to the Δ_{ij}'s as dummy symbols and will write Δ if only one dummy symbol is required. If E_j occurs then the codeword in C associated with the first term of the j^{th} periodic component becomes the output of V, corresponding to the homophone denoted by v_{ij}. Otherwise, i.e., if $\overline{E_j}$ occurs, u_i is stored, the codeword in C associated with the ratio (dummy symbol) becomes the first symbol of the homophone under construction. The binary experiment is repeated as many times as necessary until the event E_j occurs and the codeword in C associated with v_{ij} becomes the last symbol of the homophone $\Delta_{ij}\Delta_{ij}\ldots\Delta_{ij}v_{ij}$ which is output as the value assumed by V. For example, if $\overline{E_1}$ occurs three times before E_1 occurs, the homophone for u_1 will be $\Delta_{11}\Delta_{11}\Delta_{11}v_{11}$.

Example 5: Consider once more the $K = 2$ DMS with $P_U(u_1) = 1/5$ and $P_U(u_2) = 4/5$. The two periodic components in $P_U(u_1)$ suggest that we represent u_1 by two countably infinite subsets of homophones, denoted as V_{11} and V_{12}, respectively, where $P(V = V_{11}) = \sum_{i=0}^{\infty}(1/8)(1/16)^i = 2/15$ and $P(V = V_{12}) = \sum_{i=0}^{\infty}(1/16)(1/16)^i = 1/15$. Given that $U = u_1$, it follows that V_{11} is selected with probability $2/3$, or V_{12} is selected with probability $1/3$. The homophones corresponding to u_1 are the elements in the subsets V_{11} and V_{12}, where $V_{11} = \{v_{11}, \Delta v_{11}, \Delta\Delta v_{11}, \ldots\}$ and $V_{12} = \{v_{12}, \Delta v_{12}, \Delta\Delta v_{12}, \ldots\}$.

We notice further that the probability of the i^{th} element in V_{11} is $P_{V_{11}}^{i-1}(\Delta)P_{V_{11}}(u_1) = (1/16)^{i-1}(1/8)$ and, correspondingly in V_{12} is $P_{V_{12}}^{i-1}(\Delta)P_{V_{12}}(u_1) = (1/16)^{i-1}(1/16)$, respectively, for $1 \leq i < \infty$. Similarly, if $U = u_2$, it follows that $P_U(u_2) = 4/5 = \sum_{i=0}^{\infty}(1/2)(1/16)^i + \sum_{i=0}^{\infty}(1/4)(1/16)^i = 8/15+4/15$ and thus $P(V = V_{21}) = 8/15$ and $P(V = V_{22}) = 4/15$, etc. With a little abuse of notation we write

$$\begin{aligned}
V_{11} &= \{110, 1111|110, 1111|1111|110, \ldots\} \\
V_{12} &= \{1110, 1111|1110, 1111|1111|1110, \ldots\} \\
V_{21} &= \{0, 1111|0, 1111|1111|0, \ldots\} \\
V_{22} &= \{10, 1111|10, 1111|1111|10, \ldots\},
\end{aligned}$$

where v_{11} is mapped to 110 in V_{11}, v_{12} is mapped to 1110 in V_{12}, v_{21} is mapped to 0 in V_{21} and v_{22} is mapped to 10 in V_{22}. Furthermore, Δ is mapped to 1111, and the symbol $a|b$ denotes the concatenation of a and b.

The success of the sequential implementation of the JKM scheme relies on the truth of the following proposition [6].

Proposition 2: In any probability distribution with rational entries only, a probability can always be decomposed in base 2 as a sum of a finite number of distinct non-periodic components plus a finite number of distinct periodic (infinite geometric series) components whose first term and ratio are both negative powers of 2.

In general, an unbounded number of coin tosses may still be required to produce a homophone in our implementation of the JKM scheme, however that has no influence in the memory size for storing homophonic codewords, and as shown in [3] the expected number of bits needed to select a homophone is 4. Decoding is immediate: in the received binary sequence (concatenation of codewords from C) one just deletes the codewords representing the dummy symbols and maps back to the corresponding u_i's the remaining codewords.

Example 6: Consider the $K = 2$ DMS with $P_U(u_1) = 7/10$ and $P_U(u_2) = 3/10$. For the binary expansion of 7/10 it follows that $n_1' = 5$ and $s_1 = 4$, thus

$$P_U(u_1) = \frac{(15/16)(7/10)}{21/32} = \frac{21/32}{1-1/16} = \frac{1/2+1/8+1/32}{1-1/16},$$

which we write as

$$\begin{aligned}
P_U(u_1) &= (1/2)\textstyle\sum_{i=0}^{\infty}(1/16)^i + (1/8)\sum_{i=0}^{\infty}(1/16)^i \\
&+ \ (1/32)\textstyle\sum_{i=0}^{\infty}(1/16)^i \\
&= 1/2 + (1/8)\textstyle\sum_{i=0}^{\infty}(1/16)^i + (1/16)\sum_{i=0}^{\infty}(1/16)^i,
\end{aligned}$$

where in the last step we used the fact that

$$(1/2)\sum_{i=1}^{\infty}(1/16)^i + (1/32)\sum_{i=0}^{\infty}(1/16)^i = (1/16)\sum_{i=0}^{\infty}(1/16)^i.$$

Similarly we obtain $P_U(u_2) = (1/4)\sum_{i=0}^{\infty}(1/16)^i + (1/32)\sum_{i=0}^{\infty}(1/16)^i$. If we apply the JKM scheme to this source we obtain the redundancy $E(W) - H(U) = 42/25 - h(3/10) = 0.799$ while for the RM scheme the corresponding redundancy is $E(\tilde{W}) - (15/16)H(U) = 31/16 - 0.826 = 1.111$ with $P(\Delta) = 1/16$.

This example shows that not always the RM scheme produces lower redundancy than the JKL scheme. The reason for the JKL outperforming the RM scheme in this case is the presence of a non-periodic component in the binary expansion of one of the probabilities. As we show next, in general, for any K-ary DMS with a probability distribution with rational entries only, the RM scheme will give a lower redundancy than the JKM scheme if and only if each one of the DMS probability entries have periodic components only.

V. The Main Result

Suppose that the binary expansion of $P_U(u_i),\ 1 \leq i \leq K$, has periodic components only, i.e., suppose that

$$P_U(u_i) = \sum_{j=1}^{J_i}\sum_{l=0}^{\infty} 2^{-r_j - lm} = \frac{\sum_{j=1}^{J_i} 2^{-r_j}}{1-2^{-m}}.$$

It follows for the RM scheme that $P_{\tilde{U}}(u_i) = (1 - 2^{-m})P_U(u_i) = \sum_{j=1}^{J_i} 2^{-r_j},\ 1 \leq i \leq K$. For standard homophonic substitution it follows that

$$H(V) = H(U) + H(V|U) \tag{3}$$

and similarly, for the augmented source of the RM scheme it follows that

$$H(\tilde{V}) = H(\tilde{U}) + H(\tilde{V}|\tilde{U}). \tag{4}$$

Since in the RM scheme

$$H(\tilde{U}) = (1 - 2^{-m})H(U) + h(2^{-m}), \tag{5}$$

by combining (3),(4) and (5), and making $\beta_m = 1 - 2^m$, it follows that

$$\begin{aligned}
H(\tilde{V}) &= \beta_m H(V) + h(2^{-m}) \\
&- [\beta_m H(V|U) - H(\tilde{V}|\tilde{U})] \qquad (6) \\
&= \beta_m H(V) + h(2^{-m}) \\
-\beta_m &\sum_{i=1}^{K} P_U(u_i)[H(V|U = u_i) - H(\tilde{V}|\tilde{U} = u_i)], \qquad (7)
\end{aligned}$$

where in (7) we used the fact that $P_{\tilde{U}}(u_i) = (1 - 2^{-m})P_U(u_i)$. We observe in (7), however, that the terms in the expression for $H(\tilde{V}|\tilde{U} = u_i)$ are contained in the expression for $H(V|U = u_i)$, and since both entropies $H(\tilde{V}|\tilde{U} = u_i)$ and $H(V|U = u_i)$ are nonnegative, it follows that $H(V|U = u_i) \geq H(\tilde{V}|\tilde{U} = u_i)$. Since $P_{\tilde{U}}(u_i) = \sum_{j=1}^{J_i} 2^{-r_j}$, where J_i and the r_i are positive integers, it follows for each $j,\ 1 \leq j \leq J_i$, that $\tilde{V} = v_{ij}$ with probability $P(\tilde{V} = v_{ij}) = 2^{-r_j}$, and thus we can write $P(\tilde{V} = v_{ij}|\tilde{U} = u_i) = 2^{-r_j}/[(1 - 2^{-m})P_U(u_i)] = \alpha_{ij}$. It follows for the RM scheme that

$$H(\tilde{V}|\tilde{U} = u_i) = -\sum_{j=1}^{J_i} \alpha_{ij}\log\alpha_{ij} \tag{8}$$

$$= \sum_{j=1}^{J_i} r_j\alpha_{ij} - \log\frac{1}{P_U(u_i)} + \log(1 - 2^{-m}). \tag{9}$$

Similarly, since $P_U(u_i) = \sum_{j=1}^{J_i} \sum_{l=0}^{\infty} 2^{-r_j - lm}$, it follows that $P(V = v_{i,j+lm}|U = u_i) = 2^{-r_j - lm}/P_U(u_i)$, and we can write for standard homophonic coding

$$H(V|U = u_i) = -\sum_{l=0}^{\infty} \sum_{j=1}^{J_i} \frac{2^{-r_j - lm}}{P_U(u_i)} \log \frac{2^{-r_j - lm}}{P_U(u_i)}$$

$$= \sum_{j=1}^{J_i} r_j \alpha_{ij} - \log \frac{1}{P_U(u_i)} + \frac{m2^{-m}}{1 - 2^{-m}}. \qquad (10)$$

Subtracting (8) from (10) and multiplying the result by $(1 - 2^{-m})P_U(u_i)$ we have

$$(1 - 2^{-m})P_U(u_i)[H(V|U = u_i) - H(\tilde{V}|\tilde{U} = u_i)]$$
$$= P_U(u_i)h(2^{-m}). \qquad (11)$$

Finally, taking (11) into (6) it follows that

$$H(\tilde{V}) = (1 - 2^{-m})H(V). \qquad (12)$$

However, since all probabilities in both V and $\tilde{V}$ are negative powers of 2, it follows that $E(W) = H(V)$ and that $E(\tilde{W}) = H(\tilde{V})$, and thus we rewrite (12) as

$$E(\tilde{W}) = (1 - 2^{-m})E(W). \qquad (13)$$

Under the assumption we made on $P_U(u_i), 1 \le i \le K$, at the beginning os this section, it follows from (13) that the redundancy $\tilde{\rho} = E(\tilde{W}) - (1 - 2^{-m})H(U)$ of the RM scheme is related to the redundancy $\rho = E(W) - H(U)$ of the JKM scheme as $\tilde{\rho} = (1 - 2^{-m})\rho$. We have thus proved the following.

Proposition 3: For a DMS source with a probability distribution with rational entries only, and such that each probability entry has a binary expansion with periodic components only, the redundancy $\tilde{\rho} = E(\tilde{W}) - (1 - 2^{-m})H(U)$ of the RM scheme is never greater than the corresponding redundancy $\rho = E(W) - H(U)$ of the JKM scheme. The redundancies $\tilde{\rho}$ and ρ are related as $\tilde{\rho} = (1 - 2^{-m})\rho$. Furthermore, the least redundancy of the RM scheme is achieved for the least m for which all probability entries in the expanded alphabet $\tilde{U}$ have denominators which are powers of 2.

References

[1] C. E. Shannon, ”Communication Theory of Secrecy Systems”, *Bell System Tech. J.*, vol. 28, pp. 656-715, Oct., 1949.

[2] Ch. G. Günther, “A Universal Algorithm for Homophonic Coding”, pp. 405-414 in *Advances in Cryptology- Eurocrypt’88*, Lecture Notes in Computer Science, No.330. Heidelberg and New York: Springer, 1988.

[3] H. N. Jendal, Y. J. B. Kuhn and J. L. Massey, “An Information-Theoretic Approach to Homophonic Substitution”, pp. 382-394 in *Advances in Cryptology-Eurocrypt’89* (Eds. J.-J. Quisquater and J. Vandewalle), Lecture Notes in Computer Science, No. 434. Heidelberg and New York: Springer, 1990.

[4] V. C. da Rocha Jr. and J. L. Massey, “On the Entropy Bound for Optimum Homophonic Substitution”, *IEEE Int. Symp. on Info. Theory*, Ulm, Germany, 29 June - 4 July, 1997, p. 93.

[5] V. C. da Rocha Jr. and J. L. Massey, “Better than “Optimum” Homophonic Substitution”, *IEEE Int. Symp. on Info. Theory*, Sorrento, Italy, 25 - 30 June, 2000, p. 241.

[6] V. C. da Rocha Jr., “Perfect Homophonic Substitution with Finite Memory”, *IEEE Int. Symp. on Info. Theory*, Lausanne, Switzerland, 30 June - 05 July, 2002.

[7] B. Ryabko and A. Fionov, “Efficient Homophonic Coding”, *IEEE Trans. on Info. Theory*, vol.45, no.6, pp.2083-2091, Sept. 1999.

[8] M. Hoshi and T.S. Han, “Interval Algorithm for Homophonic Coding”, *IEEE Trans. on Info. Theory*, vol.47, no.3, pp.1021-1031, March 2001.