Performance Studies of a Coding Scheme for a Slepian -Wolf Source Network

Azizuddin Abdul Aziz

Electrical and Electronics Engineering Program
Universiti Teknologi PETRONAS, 31750 Tronoh, Perak, Malaysia
Tel: +605-368-7845, Fax: +605-365-7443, E-mail: azizuddin@petronas.com.my

John Kieffer

Department of Electrical and Computer Engineering
University of Minnesota, 200 Union Street S.E, Minneapolis, MN 55455 USA
Tel: +1-612-625-8574, E-mail: kieffer@ece.umn.edu

Abstract - *We investigate the simplest Slepian-Wolf source network, consisting of an information source correlated with a second source used as a side-information for the decoder but not the encoder. The information source and side-information source are modeled as binary random sequences* $X = \{X_i\}, Y = \{Y_i\}$, *respectively, for which the pairs* (X_i, Y_i) *are independent- identically distributed (i.i.d). In 1973, Slepian and Wolf characterized the optimum rate at which* X *can be encoded so that the decoder can recover* X *with the aid of* Y. *It is only in recent years, however, that practical coding schemes have been proposed which encode at a rate close to the Slepian-Wolf rate. One such coding scheme, proposed by Ramchandran et al.[2], uses a binary linear block channel code for encoding* X *while the decoder is furnished with the coset for each source block. We propose an alternate coding scheme, in which we also use a linear block channel code based encoder, but instead use a decoder which minimizes distortion as measure by the per-letter expected Hamming distance between source blocks and decoded source blocks. We provide the results of computer calculations indicating the distortion improvement over the Ramchandran et al.[2] coding scheme that can be expected with our coding scheme. The comparisons are done for several source models and for several linear block codes.*

Keywords: *Slepian-Wolf, linear block code*

1 Introduction

The lossless distributed source coding problem has been around since the 1970s when the Slepian-Wolf Coding Theorem was proved by David Slepian and Jack K. Wolf [1]. However, according to an article in the 50th year Commemorative Special Issue of the Transactions on Information Theory [3], the practicality of the theorem has not yet been realized in real-world data compression application. The reason for this is the following.

The Slepian-Wolf theorem establishes a certain rate region consisting of vectors in n-dimensional Euclidean space, where n is the number of sources in the distributed source network. Each vector $(R_1, R_2, ..., R_n)$ in the rate region is achievable in the sense that there exist lossless encoders for the distributed sources whose encoding rates are $R_1, R_2, ..., R_n$, respectively. However, Slepian and Wolf gave no constructive mechanism for finding these encoders. Only in the last 10 years people have made progress in showing how to construct the encoders achieving any given rate vector in the Slepian-Wolf rate region [1,3]. We may characterize the distributed source coding problem as the issue of designing codes for correlated and distributed sources which do not have access to one another. Since these sources cannot communicate with each other, the sources will produce redundancy when they work as a system. Slepian-Wolf coding determines the best way for each individual sensor in a sensor network to separately communicate its part of the data, making use of the fact that the data is correlated [1]. The Slepian-Wolf theorem potentially can be applied to many practical communication systems. One illustration is a cryptographic system to produce a secured shared key for two users via correlated sources and communication over a public channel. ``Secured" means that the key is robust to any eavesdropper who is trying to obtain information about the key.

1.1 Slepian-Wolf Coding Theory

As indicated above, the Slepian-Wolf Theorem in general applies to a source network consisting of any finite number of sources. We want to give the precise statement of the Slepian-Wolf Theorem here. For simplicity, we assume a source model consisting of just two correlated sources. We denote the two sources by X and Y, and we call the pair (X, Y) a correlated source. The X source output will consist of an infinite sequence

$$X_1, X_2, X_3, \ldots$$

of random samples from a finite alphabet A_X . Likewise, the Y source output will also consist of an infinite sequence

$$Y_1, Y_2, Y_3, \ldots$$

of random samples from a finite alphabet A_Y. We assume that (X, Y) is a *stationary ergodic* correlated source, meaning that the sequence of random pairs

$$(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \ldots$$

is a *stationary ergodic* random sequence. In the following, let X^n denote the random vector of X source samples $(X_1, X_2, \ldots, X_n)$, and let Y^n denote the random vector of Y source samples $(Y_1, Y_2, \ldots, Y_n)$. Note that X^n takes its values in the set $A_X{}^n$ of n-tuples with entries from A_X and that Y^n takes its values in the set $A_Y{}^n$ of n-tuples with entries from A_Y. We say that a quadruple $(\phi_x, \phi_y, \delta_x, \delta_y)$ is a n-block code for the correlated source (X,Y) if

i. ϕ_x is a mapping from $A_X{}^n$ into some set S_X
ii. ϕ_y is a mapping from $A_Y{}^n$ into some set S_Y
iii. δ_x is a mapping from $S_X \times S_Y$ into some set $A_X{}^n$
iv. δ_y is a mapping from $S_X \times S_Y$ into some set $A_Y{}^n$

Furthermore, if S_X is of size at most 2^{nR_x} and S_Y is of size 2^{nR_Y} , then we say that $(\phi_x, \phi_y, \delta_x, \delta_y)$ is a rate (R_X, R_Y) code for the correlated source (X,Y). Also, we define

$$\text{Prob}\,[\delta_x(\phi_x(X^n), \phi_y(Y^n)) \neq X^n, \delta_y(\phi_x(X^n), \phi_y(Y^n)) \neq Y^n]$$

to be the *decoding error* of the n-block code $(\phi_x, \phi_y, \delta_x, \delta_y)$. We define a pair (R_x, R_y) to be an *admissible rate pair* if for any $\varepsilon > 0$, there exists for sufficiently large n a rate (R_x, R_y) n-block code for the source (X, Y) that has decoding error $\leq \varepsilon$. The problem posed by Slepian and Wolf was to determine all admissible rate pairs (R_x, R_y). Their solution to the problem involves joint entropy rates and conditional entropy rates computed for the correlated source (X, Y). Consequently, we need to define these concepts here. The joint entropy rate H(X, Y) of the correlated source (X,Y) is the following limit (which always exists):

$$H(X,Y) = \lim_{n \to \infty} \frac{1}{n} H(X^n, Y^n) \quad (1)$$

In the right side of (1), $H(X^n, Y^n)$ denotes the usual joint entropy of the random vectors X^n, Y^n, defined in any information theory textbook as

$$H(X^n, Y^n) = \sum_{x^n, y^n} -P(X^n = x^n, Y^n = y^n) \log_2 P(X^n = x^n, Y^n = y^n)$$

The number $H(X \mid Y)$, the conditional entropy rate of source X given source Y, is defined by

$$H(X \mid Y) = \lim_{n \to \infty} \frac{1}{n} H(X^n \mid Y^n) \qquad (2)$$

where you can obtain the conditional entropy $H(X^n \mid Y^n)$ of random vector X^n given random vector Y^n in terms of unconditional entropies as

$$H(X^n, Y^n) = H(X^n, Y^n) - H(Y^n)$$

Reversing the roles of X and Y in (2), you obtain the conditional entropy rate H(Y|X) of source Y given source X. We are now ready to state the Slepian-Wolf Theorem.

Slepian-Wolf Theorem: For the stationary ergodic correlated source (X,Y), (R_X, R_Y) is admissible if and only if

$$R_x \geq H(X \mid Y) \qquad (3)$$

$$R_Y \geq H(Y \mid X) \qquad (4)$$

$$R_x + R_y \geq H(X,Y) \qquad (5)$$

A sketch of the region consisting of all admissible rate pairs (R_X, R_Y) is given in Figure 1. Intuitively, the following is what the Slepian-Wolf Theorem is saying with respect to this region. Suppose (R_X, R_Y) lies in the achievable rate region. Then sources X and Y can be separately encoded at rates arbitrarily close to R_X, R_Y bits per sample, respectively, so that a decoder, provided with the encoded bits from both X and Y encoders, can approximately reconstruct the X source data and the Y source data with whatever precision is desired. (The block diagram in Figure 1 indicates this scenario). On the other hand, if R_X, R_Y is not in the region, it will not be possible to find a code with the properties we have described in this scenario.

1.2 Coding with Side Information

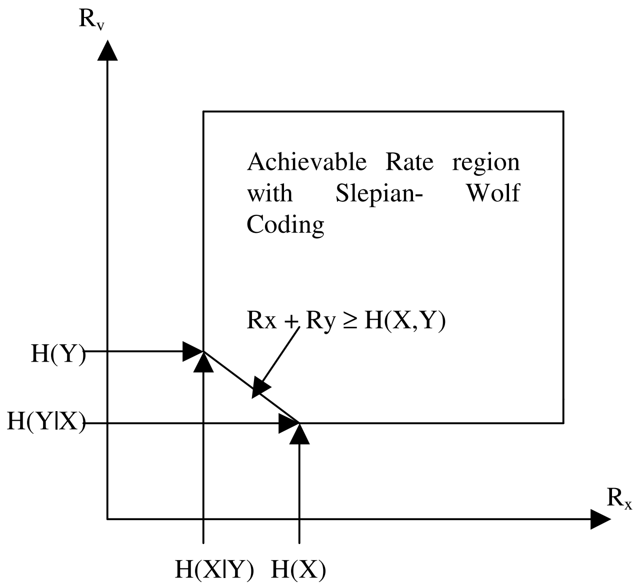

Figure 1 : The Slepian- Wolf rate region for two correlated sources

Let $|A_Y|$ denotes the cardinality of the Y source alphabet A_Y. Referring to inequalities (3)-(5), the side information which is the special case of the Slepian-Wolf Theorem is the case in which one seeks admissible rate pairs (R_X, R_Y) in which $R_Y = \log_2 |A_Y|$. In this case, inequalities (4)-(5) are automatically true once (3) is true, because

$$\log_2 |A_Y| \geq H(Y) \geq H(Y|X)$$

and

$$H(X|Y) + \log_2 |A_Y| \geq H(X|Y) + H(Y) = H(X,Y)$$

always hold. Also, encoding rate $R_y = \log_2 |A_Y|$ is automatically achieved for the Y source data if we simply encode the Y data into itself, thereby furnishing the decoder with the precise Y source data stream. In other words, in this special case, it is only the X source's data stream that is encoded; the Y source's data stream is provided directly to the decoder. The block diagram in Figure 2 denotes this ``source coding with side information'' scenario:

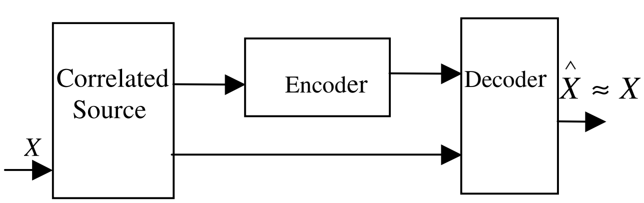

Figure 2: Source coding with side-information scenario

Formerly, in the general Slepian-Wolf problem treated in the preceding section, a n-block code consisted of a quadruple $(\phi_x, \phi_y, \delta_x, \delta_y)$ where ϕ_x, ϕ_y are the respective encoders for sources X, Y and δ_x, δ_y are the respective decoders. Now our n-block code is simply of the form (ϕ_x, δ_x). We want to specify the set of all R_x for which, given any $\varepsilon > 0$ there exists for n sufficiently large a rate $R_x + \varepsilon$ n-block code for which (ϕ_x, δ_x)

$$\text{Prob}[\delta_x(\varepsilon_x(X^n), (Y^n)) \neq X^n] \leq \varepsilon \qquad (6)$$

The $\hat{X}$ stream given as decoder output in Figure 2 is written $\hat{X} \approx X$ to denote that each of its n-blocks is within probability $\leq \varepsilon$ of the corresponding source n-block, as indicated by (6). We see from the Slepian-Wolf Theorem that the set of admissible R_X in the coding with side information problem is precisely the set of all R_X for which

$$R_X \geq H(X|Y)$$

is true. We can therefore interpret H(X|Y) as the smallest admissible rate for encoding source X in the ``coding with side information'' problem.

2 Source Model for X and Y

In this project, a source model is defined to create a correlation between a source X and side information which is a source Y.

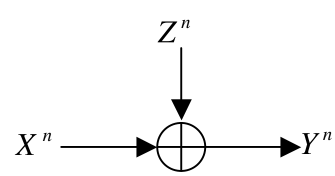

Figure 3: Source model for X and Y

Let $X^n = (X_1, X_2, \ldots, X_n)$ be a random sequence of binary inputs to the correlation channel. Let $Y^n = (Y_1, Y_2, \ldots, Y_n)$ be a sequence of binary outputs of the correlation channel. Y^n is the sum (modulo two) of X^n and Z^n. $Z^n = (Z_1, Z_2, \ldots, Z_n)$ can be regarded as the channel error resulting from the transmission of X^n. The components Z_i are binary and independent from each other and also independent from the random binary input X^n . This channel error occurs randomly in order to complete the model. This channel is constructed to model their correlation. Actually this is Binary Symmetric Channel but it is generally termed as correlation channel. The source model is chosen to a binary symmetric channel because it poses a more practical correlation model.

X^n and Y^n equiprobably distributed where $P(X_i = 0) = P(X_i = 1) = P(Y_i = 0) = P(Y_i = 1) = \frac{1}{2}$. $Z^n = (Z_1, Z_2, \ldots, Z_n)$ is randomly generated by specifying a parameter p which is known as crossover probability. The crossover probability p is characterized by probability distribution $P(Z_i = 1) = p$ and $P(Z_i = 0) = 1 - p$. The event $\{Z_i = 1\}$ means that a transmission error with probability p has occurred precisely at the i-th transmission. Whereas, the event $\{Zi = 0\}$ means that the transmission is successful with probability (1-p).

3 Project Implementation

Different binary linear block (n, k) codes were considered. This yields Rate, R = (n - k) / n. Then, correlated source models were determined by codes with rate, R satisfying :
$R \geq Hp(X|Y) = -p \log2 p-(1-p) \log2(1-p)$[Entropy of X given Y as a function of p](p is crossover probability for Binary Symmetric Channel). Finally, Ramchandran and the proposed decoding schemes performance were compared. Below, Ramchandran Decoding and the proposed decoding were discussed.

3.1 Optimal Decoder for Ramchandran Coding Scheme

Approach 1: Optimized decoder that minimizes the distortion measure Prob[$X \neq \hat{X}^n$].

Slepian-Wolf Decoding Algorithm for Ramchandran Coding scheme :

1. From syndrome of x, find the coset of x.
2. Pick any a ∈ this coset.
3. Compute y' = y + a (a is arbitrarily chosen)
4. Find the coset of y'
5. Compute u = y' + coset leader of the coset of y'
6. The estimated codeword = u + a

In [2], Ramchandran et al. showed that their optimized decoder was Minimum Hamming distance decoder. Their optimized decoder minimized with respect to the distortion criterion Prob[$X \neq \hat{X}^n$].

3.2 Optimal Decoder for the Proposed Coding Scheme

Approach 2: Optimized decoder that minimizes distortion measure $E[d_n(X^n, \hat{X}^n)]$. The decoding process is the same as Ramchandran method. However, instead of minimizing the distortion criterion Prob[$X \neq \hat{X}^n$], we minimized $E[d_n(X^n, \hat{X}^n)]$.

Slepian-Wolf Decoding Algorithm for the proposed coding scheme :

1. From syndrome of x, find the coset of x.
2. Pick any a ∈ this coset.
3. Compute y' = y + a (a is arbitrarily chosen)
4. Find the coset of y' which will C_i

$$g_i = \arg\min_{g \in Ci} \sum_{v \in C1} w(v) p^{w(g+v)} (1-p)^{n-(g+v)}$$

5. In C_i, find g_i for which
6. Compute u = y' + g_i
7. The estimated codeword = u + a

In order to minimize $E[d_n(X^n, \hat{X}^n) \mid Y^n = y^n]$ over $\hat{x}^n$, it is similar to

$$\sum_{y^n} \min_{\hat{x}^n} E[d_n(X^n, \hat{X}^n) \mid Y^n = y^n] P[Y^n = y^n]$$

$$\sum_{y^n} \min_{\hat{x}^n} \sum_{x^n} d_n(X^n, X^n) P[X^n = \hat{x}^n \mid Y^n = y^n]$$

The inner summation is over each coset and compare from each coset to coset. This reduces to Channel Symbol Error Rate (CSER) that uses the distortion function

$$\delta(x^n, \hat{x}^n) = d(x^n, \hat{x}^n) / n$$

The decoder optimized expected distortion for CSER is given by

$$\sum_{i}^{r} \min_{g \in Ci} \sum_{v \in C1} \frac{w(v)}{n} p^{w(g+v)} (1-p)^{n-(g+v)}$$

$C_1, C_2, \ldots, C_r$ are the cosets, with C_1 being the codeword coset.

4 Results

4.1 A study on Binary Linear Block (3,1) code

The result in Figure 4 shows that the CSER decoding scheme yields the same performance as Ramchandran decoder. The rate of transmission,

$$R = H_p(X \mid Y) = h(p) = \frac{n-k}{k} = 0.6667$$

The cross 'X' in Figure 4 marks the value of p when $R = H(X \mid Y)$. The according value of p = 0.17412. Beyond this point, we have to consider Wyner-Ziv distortion-rate curve in order to compare the results.

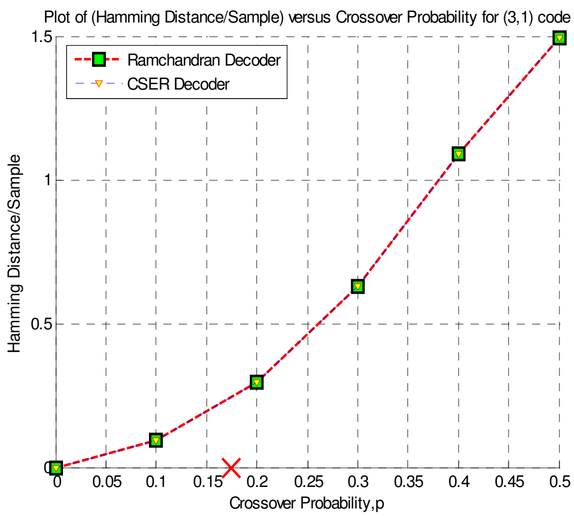

Figure 4: The plot of Hamming Distance per samples versus Crossover Probability, p of Ramachandran and CSER decoding schemes for (3,1) code.

4.2 A study on Binary Linear Block (5,2) Code

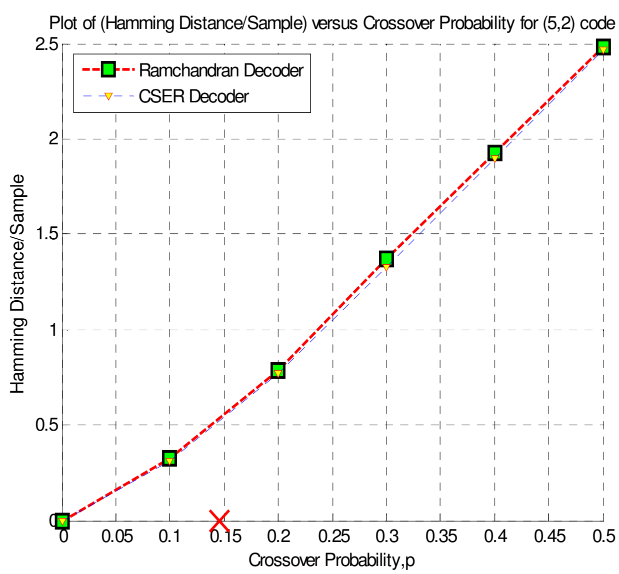

Figure 5: The plot of Hamming Distance per samples versus Crossover Probability,p of Ramachandran and CSER decoding schemes for (5,2) code.

The result shows that the CSER decoding scheme has a better performance than Ramchandran decoder for (5,2) BCH code. It can be seen that the thin broken line curve representing the CSER decoding scheme is lower than the curve of Ramchandran decoding scheme. The rate of transmission,

$$R = H_p(X \mid Y) = h(p) = \frac{n-k}{k} = 0.6000$$

The cross 'X' in Figure 5 marks the value of p when $R = H(X \mid Y)$. The according value of p =0.14612. Beyond this point, we have to consider Wyner-Ziv distortion-rate curve in order to compare the results.

4.3 Study on Binary Linear Block (5,3) Code

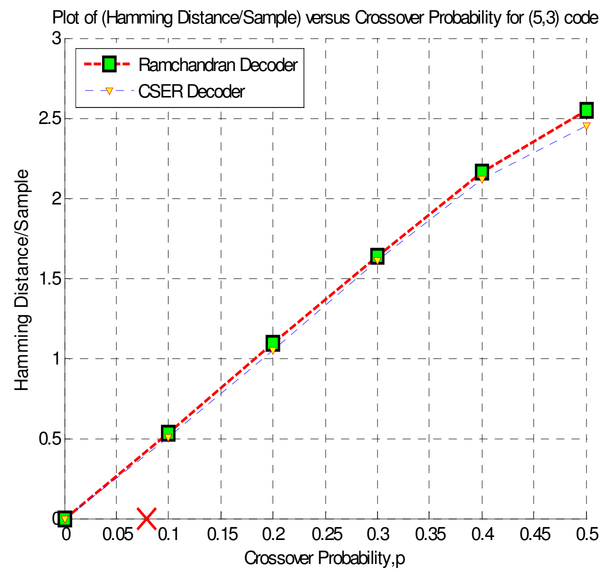

Figure 6: The plot of Hamming Distance per samples versus Crossover Probability,p of Ramachandran and CSER decoding schemes for (5,3) code.

It can be seen from Figure 6 that the curve (indicated by the think broken line) resulting from the CSER decoding scheme is lower than the curve of Ramchandran decoding. Hence, CSER decoding scheme has an improvement in performance than Ramchandran decoder for (5,2) BCH code. The rate of transmission,

$$R = H_p(X \mid Y) = h(p) = \frac{n-k}{k} = 0.4000$$

The cross 'X' in the figure above marks the value of p when $R = H(X \mid Y)$. The according value of p = 0.0794. Beyond this point, we have to consider Wyner-Ziv distortion-rate curve in order to compare the results.

4.4 A study on Binary Linear Block (7,4) Code

The analytical result is compared with the simulation result. We can se that the simulation results match the analytical result exactly. The result shown in Figure 7 shows that the CSER decoding scheme yields the same performance as Ramchandran decoder. The rate of transmission,

$$R = H_p(X \mid Y) = h(p) = \frac{n-k}{k} = 0.42857$$

The cross 'X' in the Figure 7 above marks the value of p when $R = H(X \mid Y)$. The according value of p = 0.08766. Beyond this point, we have to consider Wyner-Ziv distortion-rate curve in order to compare the results.

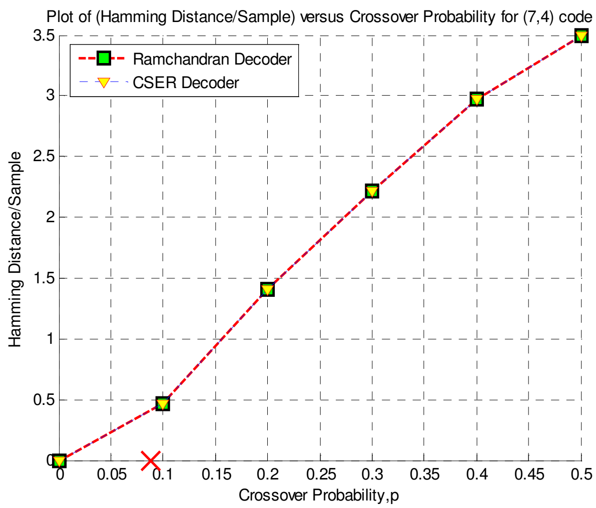

Figure 7: The plot of Hamming Distance per samples versus Crossover Probability,p of Ramachandran and CSER decoding schemes for (7,4) code.

4.5 A Study on Binary Linear Block (15, 11) Code

The Hamming (15,11) code is a special type of BCH code. This is a single-error correcting BCH code since BCH with length $2^m - 1$ is a single-error correcting and thus a Hamming code [13]. The result shows that the CSER decoding scheme yields the same performance as Ramchandran decoder.
The rate of transmission,

$$R = H_p(X \mid Y) = h(p) = \frac{n-k}{k} = 0.2667$$

The cross 'X' in Figure 8 marks the value of p when $R = H(X \mid Y)$. The according value of p = 0.04545. Beyond this point, we have to consider Wyner-Ziv distortion-rate curve in order to compare the results. We do not need to study the other Hamming code with longer block length such as (31,26),(63,57),(127,120) and (255,247) because they have the similar performance with Hamming (15,11) code. These codes are single-error correcting, perfect codes and have minimum distance equals to three [14]. The analytical and simulation methods are the same with (15,11) code. Furthermore, the computation for longer block length will be more time consuming and more complex.

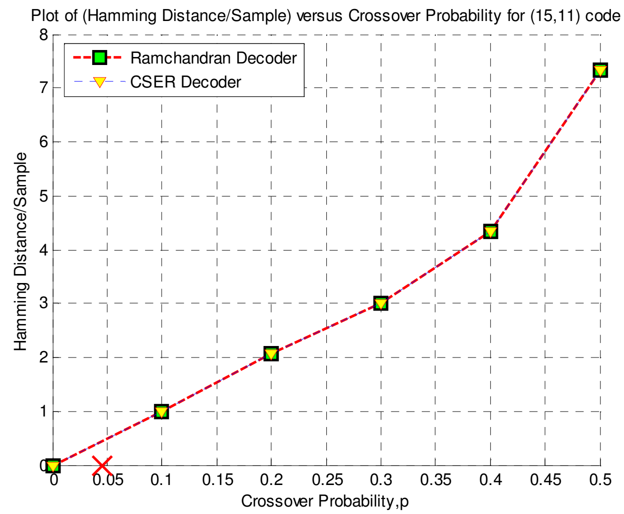

Figure 8 : The plot of Hamming Distance per samples versus Crossover Probability, p of Ramachandran and CSER decoding schemes for (15,11) code.

5 Conclusions

In this paper, we presented the comparison between Ramchandran and his group's optimized decoder for Slepian-Wolf coding and our optimized decoder. From the observation of the simulation results, it can be concluded that the optimized decoder which minimizes with respect to the distortion criterion $E[d_n(X^n, \hat{X}^n)]$ gave a similar or better performance if compared to Ramchandran and his group optimized decoder. Ramchandran and his group had constructed an optimized decoder which minimized with respect to the distortion criterion Prob[$X \neq \hat{X}^n$]. This is optimized decoder is Minimum Hamming Distance Decoder. We also have showed that Ramchandran and his group's optimized decoder gives the same performance as our optimized decoder for some linear block codes especially Hamming codes. The other linear block codes which were tested by using our decoder yield a better performance than Ramchandran and his group's optimized decoder.

6 Acknowledgments

The authors would like to thank University Teknologi PETRONAS for the support and other individuals who had contributed to this project.

7 References

[1] D.Slepian and J.K. Wolf,"Noiseless coding of correlated information sources," *IEEE Transaction on Information Theory*, vol.19,no.4,pp.471-480,Jul 1973.

[2] S.S.Pradhan and K.Ramchandran,"Distributed source coding using syndromes (DISCUS):Design and construction," *IEEE Transaction on Information Theory*, vol.49,issue 3,March 2003.page(s)620-643.

[3] S.Verdu,"Fifty years on Shannon theory," *IEEE Transaction on Information Theory*, vol.44, pp.2057-2078, October 1998.

[4] T. Uyematsu," An algebraic construction of codes for Slepian-Wolf source networks," *IEEE Transactions on Information Theory 47(7*): 3082-3088 (2001).

[5] A.D.Liveris, C.-F.Lan, K.R. Narayanan, Z. Xiong,C.N.Georghiades," Slepian-Wolf Coding of Three Binary Sources Using LDPC Codes," *Proc. 3rd International Symposium on Turbo Codes*, pp.63-66,Brest,France,September 2003.

[6] M.Elia, "Some Results on the Existence of Binary Linear Codes,"*IEEE 'Transactions on Information Theory,* Vol.IT-29, No.6, November 1983.

[7] Kieffer,J.C.," More about Conditional Entropy "Unpublished Class Notes for ECE5581,University of Minnesota,2004,http://www.ece.umn.edu/kieffer/ece5581.html

[8] Z.Xiong,A.D.Liveris,S.Cheng,"Compression via Channel Coding :Distributed Source Coding for Sensor Networks",*IEEE Signal Processing Magazine September* 2004.

[9] T.M.Cover,"A proof of the Data Compression Theorem of Slepian and Wolf for Ergodic Process", *IEEE Transactions on Information Theory,*vol.IT-21,pp.226-228,Mar.1975.

[10] T. Cover and J.Thomas, *Elements of Information Theory*,New York :Wiley,1991.

[11] A.Wyner,"Recent results in the Shannon Theory,"*IEEE Transactions on Information Theory*, vol.20,no.1,pp.2-10,1974.

[12] P.Garrett,The Mathematics of Coding Theory-Information,Compression,Error Correction,and Finite Fields, New Jersey:Pearson Prentice Hall,2004.

[13] Shu Lin and Daniel.J.Castello,*Error Control Coding:Fundamnentals and Applications*, Prentice Hall 1983.

[14] H.Tong,"Hamming Code Performance Analysis",*IEEE Transactions on Information Theory,* 2001