On optimal weighted-sum rates for the modulo sum problem

Chandra Nair and Yan Nan Wang
Dept. Of Information Engg.
The Chinese University of Hong Kong
Sha Tin, N.T., Hong Kong
Email: {chandra,dustin}@ie.cuhk.edu.hk

***Abstract*—In a seminal work Körner and Marton showed that for computing the module-two sum of doubly symmetric binary sources, linear codes achieved the optimal rates and outperformed random coding and binning based arguments. Körner also showed the optimality of Slepian-Wolf based random coding for the same problem for a different class of pairwise distributions. We show that the optimal sum-rate is given by linear codes for a larger class of binary distributions, thus extending the optimality results for this problem.**

I. Introduction

Let $p(x,y)$ denote the joint probability mass function of two random variables taking values in some finite alphabet space. Let (X^n, Y^n) be a sequence of random variables that are generated i.i.d. according to $p(x,y)$. A distributed source coding problem models communication from two senders, one who observes X^n and the other who observes Y^n, to a common receiver who wishes to decode $Z^n = (f(X_1,Y_1), f(X_2,Y_2), .., f(X_n,Y_n))$. An (n,R_X,R_Y)-code for this problem consists of two encoders: one that maps sequences X^n into symbols $M_X \in [1 : \lceil 2^{nR_X} \rceil]$ and another that maps sequences Y^n into symbols $M_Y \in [1 : \lceil 2^{nR_Y} \rceil]$; and a decoder that maps the received symbols (M_X, M_Y) into an estimate $\hat{Z}^n$ of the sequence Z^n. The probability of error for an (n,R_X,R_Y)-code, C, is defined as $\mathrm{P}(\hat{Z}^n \neq Z^n)$. A rate pair (R_X,R_Y) is said to be achievable for this problem if there exists a sequence of (n,R_X,R_Y)-codes such that the probability of error tends to zero as n tends to infinity. The closure of the set of all achievable rate pairs (R_X,R_Y) is called the capacity region, denoted as $\mathcal{C}$.

In [1], a remarkable result by Slepian and Wolf showed that when $Z = (X,Y)$ random binning ideas can be used to achieve the following rate region:

$$\begin{aligned} R_X &\geq H(X|Y) \\ R_Y &\geq H(Y|X) \\ R_X + R_Y &\geq H(XY) \end{aligned} \tag{1}$$

and hence this becomes an achievable region for any function $f(X,Y)$. We shall call this region the *Slepian-Wolf region*. Random coding and random binning ideas were used subsequently for many network information theory problems to yield the capacity results and still drives most of the achievable regions studied in the community.

Körner and Marton considered the following Doubly Symmmetric Binary Source (DSBS) distribution, i.e., for some $p \in [0,1]$,

$$p(x,y) = \begin{bmatrix} \frac{1-p}{2} & \frac{p}{2} \\ \frac{p}{2} & \frac{1-p}{2} \end{bmatrix}$$

where row index is $x \in \{0,1\}$, column index is $y \in \{0,1\}$. They investigated the capacity region when $Z = X \oplus Y$, i.e. the receiver wishes to compute the modulo-two sum of the sequences X^n, Y^n. In particular they showed that linear codes can be used to achieve the rate region:

$$\begin{aligned} R_X &\geq H(Z) \\ R_Y &\geq H(Z) \end{aligned} \tag{2}$$

and further that this matches the capacity region when $p(x,y)$ is DSBS distribution. We shall call this region the *Körner-Marton region*. For any $p \in (0,1)$ it is immediate that the above region is strictly larger than the region given by (1). Thus it became apparent that random coding ideas had its limitations and structured codes were needed for multiuser information theory problems. This has then led to development of lattice codes, coset codes, and other ideas that have spurred a sub-field of algebraic network information theory.

Returning back to the modulo-two sum problem, Körner showed the following result:

Theorem 1 (Exercise 16.23 in [2]). *When $H(Z) \geq \min\{H(X), H(Y)\}$, Slepian-Wolf's rate region characterizes the capacity region $\mathcal{C}$ for the Körner-Marton sum modulo two problem.*

Remark 1. To the best of the knowledge of the authors, these are all the collection of joint distributions $p(x,y)$ for which the capacity region has been determined. In this paper we show that linear codes minimize the sum-capacity for a larger class of distributions that include the DSBS as a special case.

In 1982, Ahlswede and Han [3] combined both the coding schemes above and obtained the following achievable rate region:

Theorem 2 (Ahlswede and Han [4]). *A rate pair (R_X, R_Y) is achievable if*

$$R_X \geq I(U;X|V) + H(Z|UV)$$

$$\begin{aligned} R_Y &\geq I(V;Y|U) + H(Z|UV) \\ R_X + R_Y &\geq I(UV;XY) + 2H(Z|UV) \end{aligned}$$

for some U and V that satisfy the Markov chain $U \to X \to Y \to V$.

Remark 2. The following remarks are worth noting.

1) Observe that when U, V are constant random variables, above rate region reduces to Slepian-Wolf's rate region; and when $U = X, V = Y$, it's reduced to Körner-Marton's rate region obtained using linear codes.
2) The multi-letter extensions of the above region tends to the capacity region. To see this, set $U = M_X$ and $V = M_y$ and apply Fano's inequality.
3) The above rate region remains achievable (and multi-letter extension tends to capacity) even if we assume that X, Y take some values in a finite field and Z is the modulo-sum in the field. See for instance Lemma 5 in [5].
4) It has been conjectured in [6], and verified by numerical simulations by different groups of researchers, that the smallest sum-rate yielded by the above region is indeed the minimum of $\{H(XY), 2H(Z)\}$, i.e. the minimum of the Slepian-Wolf region and the Körner-Marton region.
5) It is also known that for weighted sum-rate the region is strictly larger than the convex hull of the Slepian-Wolf region and the Körner-Marton region

The following is the cut-set lower bound which is rather immediate.

Theorem 3 ([7]). *Any achievable rate pair (R_X, R_Y) for the modulo sum problem must satisfy*

$$\begin{aligned} R_X &\geq H(Z|Y) = H(X|Y) \\ R_Y &\geq H(Z|X) = H(Y|X) \\ R_X + R_Y &\geq H(Z). \end{aligned}$$

Notation: $\bar{x} := 1 - x$.

II. Main Results

In this section, we derive a lower bound for the weighted sum-rate of the capacity region. We will then show that the lower bound is tight for several classes of distributions (including distributions for which the optimality was not known before). The following tensorization lemma will be used in the proof of the theorem.

Lemma 1. *Let $\lambda \geq 1$ and let (X^n, Y^n) be i.i.d distributed according to $p(x,y)$ where X, Y take values in a finite field. Let Z^n be obtained as $Z_i = X_i \oplus Y_i, i = 1, .., n$, i.e. the component-wise modulo sum on the field. Then for any $\lambda \geq 1$ the following holds:*

$$\begin{aligned} &\min_{\hat{U}:\hat{U}\to X^n \to Y^n} \lambda H(Z^n|\hat{U}) - H(Y^n|\hat{U}) \\ &\quad = n\left(\min_{U:U\to X\to Y} \lambda H(Z|U) - H(Y|U)\right). \end{aligned}$$

Proof. Clearly, by taking i.i.d. copies of the minimizer of the right hand side, it is immediate that the left-hand-side is at most the value of the right hand side. To show the other direction, observe that

$$\begin{aligned} &\lambda H(Z^n|\hat{U}) - H(Y^n|\hat{U}) \\ &= \sum_{i=1}^{n} \big((\lambda-1)H(Z_i|\hat{U}, Z^{i-1}) + H(Z_i|\hat{U}, Z^{i-1}) \\ &\qquad - H(Y_i|\hat{U}, Y_{i+1}^n)\big) \\ &= \sum_{i=1}^{n} \big((\lambda-1)H(Z_i|\hat{U}, Z^{i-1}) + H(Z_i|\hat{U}, Z^{i-1}, Y_{i+1}^n) \\ &\qquad - H(Y_i|\hat{U}, Z^{i-1}, Y_{i+1}^n)\big) \\ &\geq \sum_{i=1}^{n} \lambda H(Z_i|U_i) - H(Y_i|U_i), \end{aligned}$$

where $U_i = (\hat{U}, Y_{i+1}^n, Z^{i-1})$ and note that $U_i \to X_i \to (Y_i, Z_i)$ is Markov. The second equality above uses the Körner-Marton identity that $\sum_{i=1}^n I(Z^{i-1}; Y_i|\hat{U}, Y_{i+1}^n) = \sum_{i=1}^n I(Y_{i+1}^n; Z_i|\hat{U}, Z^{i-1})$. This completes the proof. □

We now state a lower bound to the capacity region, which we believe is new.

Theorem 4. *Any achievable rate pair (R_X, R_Y) for the modulo sum problem must satisfy the following constraints for any $\lambda \geq 1$:*

$$R_X + \lambda R_Y \geq H(XY) + \min_{U\to X\to Y} \lambda H(Z|U) - H(Y|U)$$

$$\lambda R_X + R_Y \geq H(XY) + \min_{V\to Y\to X} \lambda H(Z|V) - H(X|V)$$

Proof. In the following proof, we will use the bold alphabets to represent the random vectors of length n. As observed in Remark 2 the n-letter extension of Ahlswede and Han's region tends to the capacity region $\mathcal{C}$. Hence it suffices to show that any point (nR_X, nR_y) that belongs to the n-letter extension of the region in Theorem 2 satisfies the above constraints.

Note that for any achievable rate pairs (nR_X, nR_Y) in n-letter extension of the region in Theorem 2 we have,

$$\begin{aligned} &n(R_X + \lambda R_Y) \\ &= I(U;\mathbf{X}) + \lambda I(V;\mathbf{Y}|U) + (1+\lambda)H(\mathbf{Z}|UV) \\ &\overset{(a)}{=} I(U;\mathbf{X}) + \lambda I(V;\mathbf{Z}|U) + \lambda I(V;\mathbf{Y}|U\mathbf{Z}) \\ &\qquad + (1+\lambda)H(\mathbf{Z}|UV) \\ &= I(U;\mathbf{X}) + \lambda H(\mathbf{Z}|U) + \lambda I(V;\mathbf{Y}|U\mathbf{Z}) + H(\mathbf{Z}|UV) \\ &\overset{(b)}{=} I(U;\mathbf{X}) + H(\mathbf{Y}|U) + H(\mathbf{Z}|UY) + \lambda I(V;\mathbf{Y}|U\mathbf{Z}) \\ &\qquad + \lambda H(\mathbf{Z}|U) - H(\mathbf{Y}|U) + I(\mathbf{Z};\mathbf{Y}|UV) \\ &\overset{(c)}{=} I(U;\mathbf{X}) + H(\mathbf{Y}|U) + H(\mathbf{X}|U\mathbf{Y}) + \lambda I(V;\mathbf{Y}|U\mathbf{Z}) \\ &\qquad + \lambda H(\mathbf{Z}|U) - H(\mathbf{Y}|U) + I(\mathbf{Z};\mathbf{Y}|UV) \\ &= H(\mathbf{XY}) + \lambda I(V;\mathbf{Y}|U\mathbf{Z}) + \lambda H(\mathbf{Z}|U) - H(\mathbf{Y}|U) \\ &\qquad + I(\mathbf{Z};\mathbf{Y}|U,V) \\ &\geq nH(XY) + \lambda H(\mathbf{Z}|U) - H(\mathbf{Y}|U) \\ &\overset{(d)}{\geq} nH(XY) + n\left(\min_{U\to X\to Y} \lambda H(Z|U) - H(Y|U)\right) \end{aligned}$$

The equalities (a) (b) follows from Markov chain $V \to \mathbf{Y} \to (U, \mathbf{Z})$ and (c) is due to $H(\mathbf{Z}|U\mathbf{Y}) = H(\mathbf{XZ}|U\mathbf{Y}) = H(\mathbf{X}|U\mathbf{Y})$; while the last inequality (d) uses Lemma 1.

The other lower bound in the Theorem 4 follows in a similar manner. □

Remark 3. From [8] we can see that

$$\begin{aligned}&\min_{U\to X\to Y} \lambda H(Z|U) - H(Y|U)\\ &\quad= -\Big(\max_{U\to X\to Y} H(Y|U) - \lambda H(Z|U)\Big)\\ &\quad= -\mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)},\end{aligned}$$

where $\mathfrak{C}_x[f]\big|_{x_0}$ denotes the upper concave envelope of the function $f(x)$ with respect to x evaluated at $x = x_0$. Hence the lower bound in Theorem 4 can be written as

$$\begin{aligned}R_X + \lambda R_Y &\geq H(XY) - \mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)}\\ \lambda R_X + R_Y &\geq H(XY) - \mathfrak{C}_{\mu(y)}[H(X) - \lambda H(Z)]\big|_{p(y)} \qquad (3)\end{aligned}$$

for any $\lambda \geq 1$.

The following lemma exhibits two conditions under which the lower bound is tight. A similar statement also holds when the roles of X and Y are interchanged.

Lemma 2. *The lower bound for the weighted sum-rate $R_X + \lambda R_Y$, for $\lambda \geq 1$ given in Theorem 4 is optimal, i.e. matches the weighted sum-rate of the capacity region, if either of the following conditions hold:*

(i) $\mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)} = H(Y) - \lambda H(Z)$ and Y is independent of Z,

(ii) $\mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)} = H(Y|X) - \lambda H(Z|X)$.

Further if condition (i) holds for some $\lambda_1 > 1$, then it will also hold for $1 \leq \lambda \leq \lambda_1$; and if condition (ii) holds for some $\lambda_2 \geq 1$, then it will also hold for $\lambda \geq \lambda_2$.

Remark 4. A relatively easier condition to verify is the following: For a fixed $p_{Y|X}$ (and hence $p_{Z|X}$), if $H(Y) - \lambda H(Z)$ is concave in the distribution of X, $\mu(x)$, then condition (i) above holds. On the other hand if $H(Y) - \lambda H(Z)$ is convex in the distribution of X, $\mu(x)$, then condition (ii) above holds.

Proof. If condition (i) holds: we have from (3)

$$\begin{aligned}R_X + \lambda R_Y &\geq H(XY) - H(Y) + \lambda H(Z)\\ &= H(X|Y) + \lambda H(Z)\\ &= (\lambda + 1)H(Z)\end{aligned}$$

where the last equality uses $H(X|Y) = H(Z|Y) = H(Z)$. Note that $R_X = H(Z), R_Y = H(Z)$ belongs to the Körner-Marton achievable region, thus showing the achievability of this optimal weighted sum-rate using linear codes.

If condition (ii) holds: we have from (3)

$$\begin{aligned}R_X + \lambda R_Y &\geq H(XY) - H(Y|X) + \lambda H(Z|X)\\ &= H(X) + \lambda H(Y|X).\end{aligned}$$

Note that $R_X = H(X), R_Y = H(Y|X)$ belongs to the Slepian-Wolf achievable region, thus showing the achievability of this optimal weighted sum-rate using random binning.

To show the second part, note that condition (i) is equivalent to

$$H(Y|U) - \lambda H(Z|U) \leq H(Y) - \lambda H(Z)\ \forall U - X - Y.$$

Hence if condition (i) holds for some λ_1 then for $1 \leq \lambda \leq \lambda_1$, we have

$$\begin{aligned}&H(Y|U) - \lambda H(Z|U)\\ &= H(Y|U) - \lambda_1 H(Z|U) + (\lambda_1 - \lambda)H(Z|U)\\ &\leq H(Y) - \lambda_1 H(Z) + (\lambda_1 - \lambda)H(Z)\\ &= H(Y) - \lambda H(Z).\end{aligned}$$

Similarly, note that condition (ii) is equivalent to

$$H(Y|U) - \lambda H(Z|U) \leq H(Y|X) - \lambda H(Z|X)\ \forall U - X - Y.$$

Hence if condition (ii) holds for some λ_2 then for $\lambda \geq \lambda_2$, we have

$$\begin{aligned}&H(Y|U) - \lambda H(Z|U)\\ &= H(Y|U) - \lambda_2 H(Z|U) - (\lambda - \lambda_2)H(Z|U)\\ &\leq H(Y|X) - \lambda_2 H(Z|X) - (\lambda - \lambda_2)H(Z|X)\\ &= H(Y|X) - \lambda H(Z|X),\end{aligned}$$

where we have used $U \to X \to Z$ being Markov in the last inequality, apart from condition (ii). □

Remark 5. The conditions for optimality in the lemma is remniscent of the essentially less noisy condition for broadcast channel in [9].

Corollary 1. *The Slepian-Wolf rate region is optimal for the modulo-sum problem if $\mathfrak{C}_{\mu(x)}[H(Y) - H(Z)]\big|_{p(x)} = H(Y|X) - H(Z|X) = 0$. Similarly, it is optimal if $\mathfrak{C}_{\mu(y)}[H(X) - H(Z)]\big|_{p(y)} = H(X|Y) - H(Z|Y) = 0$.*

Proof. If $\mathfrak{C}_{\mu(x)}[H(Y) - H(Z)]\big|_{p(x)} = H(Y|X) - H(Z|X)$, then we have from Equation (3) that

$$R_X + R_Y \geq H(XY).$$

The constraints $R_X \geq H(X|Y)$ and $R_Y \geq H(Y|X)$ follow from Theorem 3. The other condition follows similarly. □

A. Application to binary alphabets

In this section we will study distributions over pairs of binary alphabets and determine conditions under which one of the conditions in Lemma 2 hold. We will see that we can recover all the previously determined cases as well as recover new distributions from the results listed below.

Notation: We will parameterize the space of distributions over pairs of binary alphabets, $p(x, y)$ as follows: $\mathrm{P}(X = 0) = x, \mathrm{P}(Y = 0|X = 0) = c, \mathrm{P}(Y = 1|X = 1) = d$.

Proposition 1. *The optimal weighted sum-rate of the capacity region is given by the Slepian Wolf region if any of the following conditions hold:*

(i) *For any λ, if $(c-\frac{1}{2})(d-\frac{1}{2}) \leq 0$, or*

(ii) $\lambda \geq \left(\frac{c-\bar{d}}{c-d}\right)^2$*, $c \neq d$, and $(c-\frac{1}{2})(d-\frac{1}{2}) > 0$.*

where $\bar{d} = 1-d$.

Proof. If condition (i) holds: then it suffices to show by Corollary 1 that $H(Y) - H(Z)$ is convex in $\mu(x)$, which will then imply that $\mathfrak{C}_{\mu(x)}[H(Y) - H(Z)]\big|_{p(x)} = H(Y|X) - H(Z|X)$. Denoting $\mu(X=0) = u$, we need to show that

$$g(u) := H_2(uc + \bar{u}\bar{d}) - H_2(uc + \bar{u}d)$$

is convex in u, when $(c-\frac{1}{2})(d-\frac{1}{2}) \leq 0$. Here $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ denotes the binary entropy function. Elementary calculations show that $g(u)$ is convex for $u \in [0,1]$ if and only if $(c-\frac{1}{2})(d-\frac{1}{2}) \leq 0$.

If condition (ii) holds: then it suffices to show by Lemma 2 that for $\lambda_2 = \left(\frac{c-\bar{d}}{c-d}\right)^2$, we have $\mathfrak{C}_{\mu(x)}[H(Y) - \lambda_2 H(Z)]\big|_{p(x)} = H(Y|X) - \lambda_2 H(Z|X)$. As before it suffices to show that

$$g(u) := H_2(uc + \bar{u}\bar{d}) - \lambda_2 H_2(uc + \bar{u}d)$$

is convex in u. This is again verifiable by elementary calculations. □

Remark 6. The following points are worth noting:

(i) The condition (i) above is already known and stated as exercise 16.23 page 390 of Csiszár and Körner's book [2]. One can verify that that $H(Z) \geq H(Y)$ is equivalent to $(c-\frac{1}{2})(d-\frac{1}{2}) \leq 0$.

(ii) Note that an equivalent Proposition can also be stated for the alternate parameterization: $\mathrm{P}(Y=0) = y, \mathrm{P}(X=0|Y=0) = \hat{c}, \mathrm{P}(X=1|Y=1) = \hat{d}$.

The next proposition determines conditions under which the optimal weighted sum-rate is given by the Körner-Marton region, i.e. satisfy the first constraint of Lemma 2. Continuing with the same notation $\mathrm{P}(X=0) = x, \mathrm{P}(Y=0|X=0) = c, \mathrm{P}(Y=1|X=1) = d$, since we require Y to be independent of c, we need to restrict to $x = \frac{\sqrt{d\bar{d}}}{\sqrt{d\bar{d}}+\sqrt{c\bar{c}}}$.

Proposition 2. *Let $\mathrm{P}(X=0) = x, \mathrm{P}(Y=0|X=0) = c, \mathrm{P}(Y=1|X=1) = d$ where $x = \frac{\sqrt{d\bar{d}}}{\sqrt{d\bar{d}}+\sqrt{c\bar{c}}}$. The optimal weighted sum-rate of the capacity region is given by the Körner-Marton region, i.e. using linear codes, if any of the following conditions hold:*

(i) *For any λ, if $c = d$, or*

(ii) *$1 \leq \lambda \leq \lambda_1$, $c \neq d$, and $(c-\frac{1}{2})(d-\frac{1}{2}) > 0$, where λ_1 is the larger root of the quadratic equation*

$$\lambda^2(c-d)^2 + \lambda\big(2(c-d)(c-\bar{d}) - 4d\bar{d}(c-\bar{c})^2\big) + (c-\bar{d})^2 = 0.$$

where $\bar{d} = 1-d, \bar{c} = 1-c$.

Proof. If condition (i) holds: then Z is independent of X and $H(Y) - \lambda H(Z)$ is concave in $\mu(x)$, therefore

$$\mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)} = H(Y) - \lambda H(Z).$$

Therefore Condition (i) in Lemma 2 (see (3)) is satisfied and we are done. Note that this is precisely the DSBS source whose capacity region was established by Körner and Marton in [7].

If condition (ii) in the proposition holds: define

$$g(u) := H_2(uc + \bar{u}\bar{d}) - \lambda_1 H_2(uc + \bar{u}d)$$

where λ_1 is the larger root of the quadratic equation

$$\lambda^2(c-d)^2 + \lambda\big(2(c-d)(c-\bar{d}) - 4d\bar{d}(c-\bar{c})^2\big) + (c-\bar{d})^2 = 0.$$

Then elementary calculations can be used to verify that $g(u)$ is concave for $u \in [0,1]$ and hence

$$\mathfrak{C}_{\mu(x)}[H(Y) - \lambda H(Z)]\big|_{p(x)} = H(Y) - \lambda H(Z).$$

As before Condition (i) in Lemma 2 (see (3)) is satisfied and we are done. □

Remark 7. The following points are worth noting:

(i) As long as, $(c-\frac{1}{2})(d-\frac{1}{2}) > 0$, we can see that $\lambda_1 > 1$, and hence the optimal sum-rate, will be given by the Körner-Marton region, i.e. using linear codes. Note that we still need $x = \frac{\sqrt{d\bar{d}}}{\sqrt{d\bar{d}}+\sqrt{c\bar{c}}}$. Thus linear coding strategy of Körner-Marton are optimal for some larger class of parameters.

(ii) As before, an equivalent Proposition can also be stated for the alternate parameterization: $\mathrm{P}(Y=0) = y, \mathrm{P}(X=0|Y=0) = \hat{c}, \mathrm{P}(X=1|Y=1) = \hat{d}$.

B. Comparison of the bounds

In [3] Ahlswede and Han chose the following $p(x,y)$ given by

$$p(x,y) = \begin{bmatrix} 0.003920 & 0.019920 \\ 0.976080 & 0.000080 \end{bmatrix}$$

where row index is $x \in \{0,1\}$, column index is $y \in \{0,1\}$, to show that their achievable rate region performs strictly better than both Körner and Marton's rate region and Slepian and Wolf's rate region. It turns out that for this distribution Y is indeed independent of Z. Therefore from Remark 7 we already know that the optimal sum-rate is given by the Körner-Marton linear coding region.

The figure 1 plots Ahlswede-Han's rate region, the lower bound from Theorem 4, and the cut-set lower bound for the above example.

As one can see readily and as established in Proposition 2, the lower bound in Theorem 4 yields the optimal sum-rate of $2H(Z)$ for this example. By numerical simulations: the largest λ for which the hyperplane of the lower bound passes through the $(H(Z), H(Z))$ point is $\lambda_1^* = 5.253$ (matches, curiously, the sufficient condition established in Proposition 2), while that for the Ahlswede-Han region is $\lambda_1^\dagger = 5.338$. Then the largest λ for which the hyperplane of the lower bound passes through the $(H(X), H(Y|X))$ point is $\lambda_2^* = 25.844$ (matches the sufficient condition established in Proposition 1), while, by numerical simulations, that for the Ahlswede-Han region is $\lambda_2^\dagger = 6.620$.

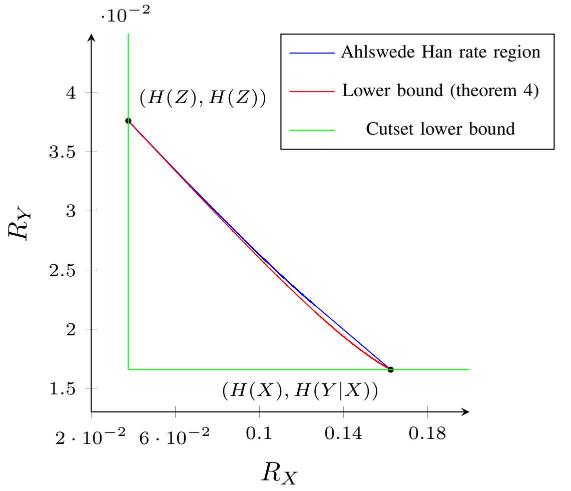

Fig. 1. Compare Ahlswede Han rate region and our lower bound

C. Application to higher alphabet fields

The modulo-sum problem for binary alphabets has a peculiar structure that was exploited in the Exercise 16.23 of [2]. If $H(Z) \geq H(Y)$, then $P_{Y|X}$ was a stochastic degradation of $p_{Z|X}$, and the reverse held if $H(Y) \geq H(Z)$. In general we know that for higher alphabets the above dichotomy does not hold. Hence Lemma 2 establishes that a better comparision between the channels $p_{Z|X}$ and $p_{Y|X}$ for obtaining the optimal weighted sum-rate is related to (essentially) less noisy comparison.

Below we provide two examples in $GF(3)$ for which the results in Lemma 2 yield optimality.

For $GF(3)$, one instance of $p(x, y)$ satisfying that Z is independent of Y and $\mathfrak{C}_{\mu(x)}[H(Y) - H(Z)]\big|_{p(x)} = H(Y) - H(Z)$ is given by the following distribution:

$$p(x, y) = \begin{bmatrix} 0.08 & 0.06 & 0.18 \\ 0.08 & 0.18 & 0.06 \\ 0.24 & 0.06 & 0.06 \end{bmatrix}$$

where row index is $x \in \{0, 1, 2\}$, column index is $y \in \{0, 1, 2\}$.

One can check that Z is independent of Y, and verify that $f(p(x)) = H(Y) - H(Z)$ is concave with respect to $p(x)$. So the first constraint of Lemma 2 is satisfied for $\lambda = 1$, and thus Körner-Marton rate region is sum rate optimal.

And another instance of $p(x, y)$ satisfying $\mathfrak{C}_{\mu(x)}[H(Y) - H(Z)]\big|_{p(x)} = H(Y|X) - H(Z|X)$ is given by the following distribution:

$$p(x, y) = \begin{bmatrix} 0.02 & 0.02 & 0.48 \\ 0.02 & 0.06 & 0.16 \\ 0.06 & 0.02 & 0.16 \end{bmatrix}$$

where row index is $x \in \{0, 1, 2\}$, column index is $y \in \{0, 1, 2\}$.

One can verify that $f(p(x)) = H(Y) - H(Z)$ is convex with respect to $p(x)$. So the second constraint of Lemma 2 is satisfied for $\lambda = 1$, thus Slepian-Wolf rate region is sum rate optimal.

III. Summary

In this paper we established that linear coding strategy of Korner and Marton [7] yields the optimal sum-rate for pairs of distributions outside the doubly symmetric binary source. This was shown by developing a lower bound and identifying sufficient conditions when the lower bound is tight. The ideas and results are applicable to larger fields as well.

Acknowledgements

The authors wish to thank Prof. Shun Watanabe who kindly told us about exercise 16.23 page 390 in Csiszár and Körner's book [2] and hence prompted us to investigate the optimality for other parameter regimes.

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19(4), pp. 471 – 480, July 1973.

[2] I. Csiszar and J. Korner, *Information theory: Coding theorems for discrete memoryless systems*. Cambridge University Press, 1 2011.

[3] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *Information Theory, IEEE Transactions on*, vol. 29, pp. 396 – 412, 06 1983.

[4] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, no. 5, pp. 805–814, 10 1974. [Online]. Available: http://dx.doi.org/10.1214/aop/1176996549

[5] T. S. Han and K. Kobayashi, "A dichotomy of functions f(x,y) of correlated sources (x,y) from the viewpoint of the achievable rate region," *IEEE Transactions on Information Theory*, vol. 33, no. 1, pp. 69–76, Jan. 1987.

[6] M. Sefidgaran, A. Gohari, and M. R. Aref, "On korner-marton's sum modulo two problem," in *2015 Iran Workshop on Communication and Information Theory (IWCIT)*, May 2015, pp. 1–6.

[7] J. Korner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Transactions on Information Theory*, vol. 25, no. 2, pp. 219–221, Mar 1979.

[8] C. Nair, "Upper concave envelopes and auxiliary random variables," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 1, pp. 12–20, 2013. [Online]. Available: http://dx.doi.org/10.1007/s12572-013-0081-7

[9] ——, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4207–4214, Sept 2010.