Zero-Error Sum Modulo Two with a Common Observation

Milad Sefidgaran and Aslan Tchamkerten
Telecom Paris, Institut Polytechnique de Paris
{milad.sefidgaran,aslan.tchamkerten}@telecom-paris.fr

Abstract—This paper investigates the classical modulo two sum problem in source coding, but with a common observation: a transmitter observes \((X,Z)\), the other transmitter observes \((Y,Z)\), and the receiver wants to compute \(X \oplus Y\) without error. Through a coupling argument, this paper establishes a new lower bound on the sum-rate when \(X - Z - Y\) forms a Markov chain.

I. INTRODUCTION

The problem of computing the modulo two sum of binary \(X\) and \(Y\) observed at different transmitters was introduced by Körner and Marton in 1979 [1]. Under the vanishing error probability criterion, they showed the optimality of a class of linear codes when the sources have symmetric probability distributions. This implied that the optimal sum-rate for computing the sum modulo two function can be less than the optimal sum-rate for the lossless recovery of the sources as characterized by Slepian and Wolf in 1973 [2]. Later, Körner showed that for general sources the rate region corresponds to Slepian-Wolf’s whenever \(H(X \oplus Y) \geq \min(H(X), H(Y))\) (see [3] Exercise 16.23 as this result refers to an unpublished reference).

In 1983, Ahlswede and Han [4] used a combination of the schemes of Slepian-Wolf and Körner-Marton to show that it was possible to improve over the convex hull of these schemes. More recent work [5] suggests that the Ahlswede-Han scheme cannot achieve a sum-rate lower than the minimum of the sum-rates by Slepian-Wolf and Körner-Marton.

Till recently, there was no better bound than the cut-set bound \(H(Y|X) + H(X|Y)\). In [5], it was shown that when \(H(X \oplus Y) < H(X|Y) + H(Y|X)\), the cut-set bound is not tight except for the cases of independent sources or sources with symmetric distribution. Recently, Nair and Wang [6] established a new lower bound on the weighted sum-rate of the transmitters which implied the optimality of the Slepian-Wolf scheme under some previously unknown conditions. Moreover, they provided sufficient conditions under which linear codes are “weighted-sum optimal.”

Finding a better lower bound than the cut-set bound amounts to quantify the potential penalty due to distributed processing. We note here that for the case of arithmetic sum under the zero-error performance criterion there exist better bounds than the cut-set bound as reported in [7]–[10].

In this paper, we consider the following variant of the sum modulo two problem for which we improve upon the cut-set bound. The transmitters have access to binary sources \((X,Z)\) and \((Y,Z)\), respectively, and use variable-length coding to send their information to the receiver who tries to compute \(X \oplus Y\) without error. The sources are assumed to satisfy the Markov chain \(X - Z - Y\). We establish a lower bound on the optimal sum-rate of this problem, which improves over the cut-set bound and over an extension of a bound by Nair and Wang [6] for this setup.

The rest of the paper is organized as follows. Section II contains preliminaries and the problem formulation. Section III presents the main result for a particular probability distribution of the sources. This result is proved in Section IV. Finally, Section V states the main result for arbitrary binary sources that satisfy \(X - Z - Y\), and provides a proof sketch.

II. PRELIMINARIES

Let \((x,z,y)\) be \(n\) realizations of finite support random variables \(X - Z - Y\) supposed to form a Markov chain. We consider the setup depicted in Fig. 1. Transmitter 1 assigns an index \(m_1(x,z) \in \{1,\ldots,L_1\}\) to its observation \((x,z)\) and transmits this index to the receiver using a uniquely decodable variable length code \(C_1 \subseteq \{0,1\}^*\). Transmitter 2 proceeds similarly, and assigns an index \(m_2(y,z) \in \{1,\ldots,L_2\}\) to its observation \((y,z)\) and transmits this index to the receiver using a uniquely decodable variable length code \(C_2 \subseteq \{0,1\}^*\). Upon receiving \((m_1,m_2)\), the receiver attempts to recover the component-wise modulo two sum \(x \oplus y\) using a decoder \(D(m_1,m_2)\). The encoding procedures (indices assignment and codebooks) and the decoder are referred to as a scheme \(S\).

The sum-rate \(R_1 + R_2\) is said to be achievable if, for any \(\varepsilon > 0\) and all \(n\) sufficiently large, there exists a scheme \(S\) such that

\[
P(D(M_1,M_2) \neq X \oplus Y) = 0
\]
and such that
\[ n(R_1 + R_2 + \varepsilon) \geq \mathbb{E}[l(c_1(M_1)) + l(c_2(M_2))], \]
where \( l(c_i(M_i)) \) denotes the length of the \( M_i \)-th codeword of transmitter \( i \).

A first lower bound is the standard cut-set bound
\[ R_1 + R_2 \geq \max(H(X|Z) + H(Y|Z), H(X \oplus Y)). \quad (1) \]
A second lower bound is obtained through a straightforward extension of [6, Theorem 4], originally derived under the vanishing probability of error criterion:

**Proposition 1.** The sum-rate is lower-bounded as
\[
R_1 + R_2 \geq H(X|Z) + H(Y|Z) + H(Z) + \max \left( \min_{U - (X,Z)} H(X \oplus Y|U) - H(Y,Z|U), \right.
\min_{V - (Y,Z)} H(X \oplus Y|V) - H(X,Z|V) \right). 
\]

**III. MAIN RESULT**

The main result, Theorem 1 below, provides a new lower bound on the sum-rate. For ease of exposition this result is first stated and proved for the particular distribution
\[
P_{X|Z}(x = 0|z = 0) = P_{Y|Z}(y = 0|z = 1) = 1/2, \]
\[
P_{X|Z}(x = 0|z = 1) = P_{Y|Z}(y = 0|z = 0) = 0. \quad (2) \]
The sum-rate lower bound for the arbitrary binary sources that satisfy \( X - Z - Y \) is stated in Section IV.

**Theorem 1.** Suppose \( (X,Y,Z) \) satisfies (2). Then,
\[ R_1 + R_2 \geq 1 + L^*(p), \]
where
\[ L^*(p) \triangleq \min_{d \leq 2p} L(p, d), \]
\[ L(p, d) \triangleq h_b(p) + \left[ d - p \cdot h_b \left( \frac{d}{2p} \right) - \bar{p} \cdot h_b \left( \frac{d}{2\bar{p}} \right) \right], \]
where \( p \triangleq P(Z = 0) \), where \( \bar{p} \triangleq 1 - p \), and where \( h_b(p) \) denotes the binary entropy \( -p \log p - (1 - p) \log(1 - p) \).

As we can see in Fig. 2, the lower bound given by Theorem 1 improves upon both the cut-set bound and the bound given by Proposition 1 for all (non-trivial) values of \( p \)—the shape of this latter bound around \( p = 1/2 \) is somewhat unexpected.

**IV. PROOFS**

Throughout the proofs \( \varepsilon \) is an arbitrary constant in \((0, 1/2)\) that can be taken arbitrarily small for \( n \) sufficiently large.

**Proof of Theorem 1** We have
\[
R_1 + R_2 \geq (a) \frac{1}{n} H(M_1) + \frac{1}{n} H(M_2) - \varepsilon \\
\geq (b) \frac{1}{n} H(M_1, M_2) - \varepsilon \\
= (c) \frac{1}{n} H(M_1, M_2|Z) + \frac{1}{n} I(M_1, M_2; Z) - \varepsilon \\
= (d) H(X|Z) + H(Y|Z) + \frac{1}{n} I(M_1, M_2; Z) - \varepsilon \\
= H(X, Y, Z) - \frac{1}{n} H(Z|M_1, M_2) - \varepsilon. \quad (3) \]
Inequality (a) follows from the unique decodability of the codes which implies that \( \mathbb{E}[l(c_i(M_i))] \geq H(M_i) \) [3, Theorem 4.1]. Equality (b) follows from the following lemma whose proof is deferred to the end of this section.
Lemma 1. For any zero-error scheme and for sources that satisfy the Markov chain $X - Z - Y$ we have

$$H(M_1, M_2 | Z) = nH(X | Z) + nH(Y | Z).$$

The main part of the proof consists in the derivation of a good upper bound on $H(Z | M_1, M_2)$ by introducing the following coupling of $(X, Y, Z)^n$. Given a coding scheme, let $(\tilde{X}, \tilde{Y}, Z)$ be an independent copy of $(X, Y, Z)$ conditioned on $(M_1, M_2)$. Hence, we have the Markov chain

$$\left( \tilde{X}, \tilde{Y}, \tilde{Z} \right) - (M_1, M_2) - (X, Y, Z)$$

and the marginal probabilities $P_{\tilde{X}, \tilde{Y}, \tilde{Z}}$ and $P_{X, Y, Z}$ coincide.

Associated with the above coupling is the $Z$-distance of the underlying scheme which we define as

$$d_{avg}(Z) \triangleq E\left[ d_H(Z, \tilde{Z}) \right] = \sum_{(m_1, m_2)} d_H(z_1, z_2) p(z_1 | m_1, m_2) p(z_2 | m_1, m_2),$$

where $d_H(\cdot, \cdot)$ denotes the normalized Hamming distance. Now,

$$H(Z | M_1, M_2) \leq H(\tilde{Z} | M_1, M_2, Z) \leq H(M_1, M_2 | Z, \tilde{Z}) + H(\tilde{Z} | Z) - nH(X, Y | Z),$$

where $(a)$ holds because of (4) and where $(b)$ follows from Lemma 1.

The rest of the proof is divided into four parts: $i.$ upper bound on $H(M_1, M_2 | Z, Z)$, $ii.$ upper bound on $H(\tilde{Z} | Z)$, $iii.$ upper bound $H(Z | M_1, M_2)$ using $i.$ and $ii.$, and finally lower bound the sum-rate in part iv.

i: For $H(M_1, M_2 | Z, \tilde{Z})$ we have

$$H(M_1, M_2 | Z, \tilde{Z}) \leq H(M_1, M_2 | Z) + H(\tilde{Z} | Z) \leq H(M_1, M_2 | Z) + H(Y | Z) = g(Z, \tilde{Z}),$$

where $(a)$ follows from Lemma 2 below (the proof is deferred to the end of this section) and Equality $(b)$ holds since $(M_1, M_2)$ is a function of $(X, Y, Z)$.

Lemma 2. For any zero-error scheme and for sources that satisfy the Markov chain $X - Z - Y$, there exists a function $g : \mathbb{Z}^n \times \mathbb{Z}^n \to \{0, 1\}^n$ such that

$$P \left( X \oplus \tilde{X} = g(Z, \tilde{Z}) = Y \oplus \tilde{Y} \right) = 1.$$
This contradicts the zero-error assumption. For any scheme we have

\[ P(X \oplus Y = \bar{X} \oplus \bar{Y}) = 1. \quad (11) \]

By contradiction suppose that such a function does not exist. Then, there exist \((x_1, y_1, z), (x_2, y_2, z), (\bar{x}_1, \bar{y}_1, \bar{z}), (x_2, \bar{y}_2, \bar{z})\) such that

\[ p(x_1, y_1, z, \bar{x}_1, \bar{y}_1, \bar{z}) > 0, \]
\[ p(x_2, y_2, z, x_2, \bar{y}_2, \bar{z}) > 0, \]

and such that \(x_1 \oplus \bar{x}_1 \neq x_2 \oplus \bar{x}_2\). From (11),

\[ x_1 \oplus y_1 = \bar{x}_1 \oplus \bar{y}_1 \]

and

\[ x_2 \oplus y_2 = \bar{x}_2 \oplus \bar{y}_2. \]

Moreover,

\[ p(x_1, y_2, z, \bar{x}_1, \bar{y}_2, \bar{z}) > 0, \]

and hence by (11) we have \(x_1 \oplus \bar{x}_1 = y_2 \oplus \bar{y}_2\). The equalities

\[ x_2 \oplus y_2 = \bar{x}_2 \oplus \bar{y}_2, \quad x_1 \oplus \bar{x}_1 = y_2 \oplus \bar{y}_2, \]

contradicts \(x_1 \oplus \bar{x}_1 \neq x_2 \oplus \bar{x}_2\).

A similar argument shows that there exists a function \(g(Z, \bar{Z})\) such that \(P(Y \oplus \bar{Y} = g(Z, \bar{Z})) = 1\). From (11) it then follows that \(g(Z, \bar{Z}) = g(Z, \bar{Z})\). □

**Proof of Lemma 2**

The Markov chain \(Z - (M_1, M_2) - \bar{Z}\) implies that \(Z_i\) has the same conditional distribution as \(\bar{Z}_i\) given \((m_1, m_2)\). Hence, define

\[ p(i|m_{1,2}) \triangleq \sum_{m_1, m_2} p(i|m_{1,2}) \]

By the Markov chain \(Z_i - (M_1, M_2) - \bar{Z}_i\),

\[ \sum_{m_1, m_2} p(i|m_{1,2}) = 2p(i|m_{1,2})(1 - p(i|m_{1,2})). \]

The \(Z\)-distance can then be written as

\[ d_{avg}(Z) = \sum_{m_1, m_2} p(m_1, m_2) d_H(Z, \bar{Z}) \]

\[ = \frac{1}{n} \sum_{m_1, m_2} p(m_1, m_2) d_H(Z, \bar{Z}) \]

\[ = \frac{2}{n} \sum_{m_1, m_2} p(m_1, m_2) \sum_{i=1}^{n} (\bar{Z}_i)(1 - P(Z_i = 0)) \]

\[ \leq \frac{2}{n} \sum_{i=1}^{n} P(Z_i = 0)(1 - P(Z_i = 0)) \]

\[ = 2p \bar{p}, \]

where (a) follows from Jensen’s inequality applied to the concave function \(x - x^2\) and by noting that

\[ P(Z_i = 0) = \sum_{m_1, m_2} p(m_1, m_2) p(i|m_{1,2}). \]

**V. Generalization of Theorem 1**

In this section, we state our result for general sources and provide a sketch of its proof.
Theorem 2. For any zero-error scheme and for sources that satisfy the Markov chain \( X - Z - Y \), we have
\[
R_1 + R_2 \geq H(X|Y) + H(Y|X) + L^*(p),
\]
where
\[
L^*(p) \triangleq \min_{d \leq 2p} L(p, d).
\]
Here, \( L(p, d) \) is redefined as:
\[
L(p, d) \triangleq H(X, Y, Z) - p h_b\left(\frac{d'}{p}\right) - \frac{p}{1-p} h_b\left(\frac{d'}{p}\right).
\]
For the probability distribution given by (2), Definition [13] reduces to the definition of \( L(p, d) \) in Theorem [1].

It can be verified that the sum-rates lower bound in Theorem 2 is tight in the cases of independent \((X, Y, Z)\), constant \(X\), and identical sources \(X = Y = Z\). The optimal sum-rates are equal to \(H(X, Y), H(Y, Z)\), and 0, respectively.

Proof sketch of Theorem 2. All the steps up to (6) carry to the general case. We now show how to bound
\[
H\left(X|Z, \tilde{Z}, X \oplus \tilde{X} = g(Z, \tilde{Z})\right) + H\left(Y|Z, \tilde{Z}, Y \oplus \tilde{Y} = g(Z, \tilde{Z})\right)
\]
for general sources. Pick strongly typical sequences \(z, \tilde{z} \in \mathcal{T}_n^p(Z)\) such that \(d_H(z, \tilde{z}) = d\) and consider the function \(g(z, \tilde{z})\) given by Lemma 2. Note that this Lemma implies that \(P(X = X \oplus g(z, \tilde{z})|Z, \tilde{Z}) = 1\). Define
\[
A_X(z, \tilde{z}, g(z, \tilde{z})) \triangleq \left\{ x : (x, z), (x \oplus g(z, \tilde{z}), \tilde{z}) \in \mathcal{T}_n^p(X, Z) \right\},
\]
\[
\mathcal{P}_X(z, \tilde{z}, g(z, \tilde{z})) \triangleq \hat{p}_{x[z, \tilde{z}, g(z, \tilde{z})]}(\cdot) : x \in A_X(z, \tilde{z}, g(z, \tilde{z}))\].