Converse and duality results for combinatorial source-channel coding in binary Hamming spaces

Andrew J. Young and Yury Polyanskiy

Abstract—This article continues the recent investigation of combinatorial joint source-channel coding. For the special case of a binary source and channel subject to distortion measured by Hamming distance, the lower (reverse) bounds on achievable source distortion are improved for all values of channel noise. Operational duality between coding with bandwidth expansion factors \( \rho > 1 \) is established. Although the exact value of the asymptotic noise-distortion tradeoff curve is unknown (except at \( \rho = 1 \)), some initial results on inter-relations between these curves for different values of \( \rho \) are shown and lead to statements about monotonicity and continuity in \( \rho \).

I. INTRODUCTION

The joint source-channel problem seeks to encode the user data in such a way that the noise applied to the encoded string did not lead to excessive distortion of the original data. The combinatorial joint source-channel (CJSCC) problem seeks to provide answers under the assumption that the noise is taken in the worst-case sense. Namely, for this paper we will focus on binary Hamming source/channel and correspondingly, the \((D, \delta)\) CJSCC is a pair of encoder and decoder such that addition of any string of (normalized) Hamming weight up to \( \delta \) does not lead to post-decoding distortion larger than \( D \), as measured by (normalized) Hamming distance.

The CJSCC problem and a framework for analysis were originally introduced in [1] and expanded in [2]. For thorough motivation and background on the problem we refer the reader to [2]. Here we only recall perhaps the more surprising observation from [1] about asymptotic sub-optimality of separated schemes in CJSCC.

In binary Hamming space, the adversarial source problem is a covering problem and the adversarial channel problem is a packing problem. For the covering problem, the asymptotically optimal covering has been found exactly, see e.g. [3]. The packing problem is addressed extensively in [4] and an exact asymptotic solution is still open. The best known lower bound is the Gilbert-Varshamov bound and the best known upper bound is the MRRW bound [5]. As such, these bounds, and the exact solution for the covering problem, characterize separation based schemes for the BSC. The observation that spurred our interest in this problem was made in [1], where it was shown that some simple CJSCC (such as repetition) achieve performance strictly better than any separated scheme for certain values of parameters.

In this paper, we extend the previous work in several directions. First, new converse bounds are proved that together improve the state-of-the-art for all values of bandwidth expansion factor \( \rho > 0 \) and channel parameters \( 0 < \delta < 1 \), cf. Fig. 2. One interesting implication is the following (perhaps counter-intuitive) conclusion: For certain values of \((D, \delta)\) increasing the redundancy factor \( \rho \) can lead to decrease in performance. For example, this holds for all \( \delta > \frac{1}{2} \), cf. Section III-A.

Second, an exact operational duality is established between the CJSCC problems at \( \rho \) and at \( \frac{1}{\rho} \). In particular, this allows us to extend our directory of basic CJSCC codes and close, for example, the question of the largest channel noise \( \delta \) for which distortion \( D \) is still less than 1. (This critical value, that is always greater than \( \frac{1}{2} \), establishes the threshold at which the adversarial binary channel becomes fully useless for conveying binary-coded information.) For more details, see (20).

The structure of the paper is as follows. Converse bounds occupy Section III. For \( \delta \geq 1/2 \), a converse relating coverings in the source and channel spaces, Section III-A, establishes that \( D \) is bounded below by a function monotonically increasing in \( \rho \) and \( \rho = 1 \) is strictly optimal in the region \( \rho \geq 1 \). For \( 1/4 \leq \delta < 1/2 \), a converse relating packings in the source and channel spaces, Section III-B, demonstrates that, for all \( \rho > 0 \), \( D \gtrsim \delta \) and, for linear codes, Section III-C, \( D \gtrsim \delta \). In Section III-D a further stronger converse bound is proved for all schemes based on the idea of repeating a small code, cf. [6]. Sections IV and V introduce the concepts of \( \rho \leftrightarrow \frac{1}{\rho} \) duality and the composition of CJSCCs, respectively. Finally, Section VI concludes with numerical comparisons and discussions.

II. PRELIMINARIES

A. Hamming Space

For the BSC, the alphabets of interest are all binary Hamming space and the adversary is restricted to outputs whose hamming distance to the input is bounded according to the channel parameter. The notation for the \( n \) fold product of the field of two elements \( \mathbb{F}_2^n \) is used for \( n \) dimensional Hamming space and Euclidean notation \(|·|\) is used for both the Hamming distance and weight.

Given a set \( S \subset \mathbb{F}_2^n \) its Chebyshev radius is the radius of the smallest Hamming ball containing all of its points,

\[
\text{rad}(S) = \min_{y \in \mathbb{F}_2^n} \max_{x \in S} d(x, y),
\]

a point \( y_0 \) achieving this minimum is called a Chebyshev center, and its covering radius is the radius of the smallest covering by points in \( S \)

\[
r_{\text{cov}}(S) = \max_{y \in \mathbb{F}_2^n} \min_{x \in S} d(x, y).
\]
These two quantities satisfy an important relation
\begin{equation}
\text{rad}(S) = n - r_{\text{con}}(S),
\end{equation}
which follows from the following property of binary Hamming space: for all \( x \in \mathbb{F}_2^n \) and \( r \in \mathbb{R} \)
\begin{equation}
B(x, r)^c = B(\bar{x}, [n - r - 1]),
\end{equation}
where \( B(x, r) := \{ y \in \mathbb{F}_2^n : |x - y| \leq r \} \) and \( \bar{x} \) is the entrywise binary negation.

There are also some combinatorial quantities of interest:
- \( K(n, r) \) -- minimal number of points covering \( \mathbb{F}_2^n \) with radius \( r \) balls;
- \( A(n, d) \) -- maximal number of points with distance between any two points greater than \( d \);
- \( A_L(n, r) \) -- the maximal number of points such that any ball of radius \( r \) contains at most \( L \) points\(^1\).

### B. Basic Definitions

**Definition 1.** A pair of maps \( f : \mathbb{F}_2^k \to \mathbb{F}_2^n \) and \( g : \mathbb{F}_2^n \to \mathbb{F}_2^k \)

\( k \) is either an encoder or decoder, \( n \) the source or channel dimension.
\( f \) is a map from the source space to the channel space and \( g \) a map from the channel space to the source space. Moreover, the notation
\( D \) is either an encoder or decoder, \( \rho \) the source-channel sequence.

\( f \) and \( g \) achieve \( D \) if there exists a sequence of \( n \) and \( k \). Hence \( D \) is either an encoder or decoder, \( \rho \) the source-channel sequence. 

**Remark 4.** Note that \( D^*(\delta; f, g) \) is just a lower boundary of all asymptotically achievable \( (D, \delta) \) with bandwidth expansion factor \( \rho \).

**Definition 2.** The optimal distortion for a \( (k, n; D, \delta) \) CJSCE is
\[ D^*(\delta; k, n) := \min_{f, g} D(\delta; f, g), \]
with minimization over \( f : \mathbb{F}_2^k \to \mathbb{F}_2^n \) and \( g : \mathbb{F}_2^n \to \mathbb{F}_2^k \).

Asymptotically we allow the user to choose the optimal sequence of source and channel dimensions.

**Definition 3.** For bandwidth expansion factor \( \rho > 0 \), the asymptotically optimal CJSCE is
\[ D^*(\delta; \rho) := \lim_{m \to \infty} \inf \limits_{(k_m, n_m)} D(\delta; k_m, n_m), \]
with the infimum over subsequences of the natural numbers such that
\[ \lim_{m \to \infty} \frac{n_m}{k_m} = \rho. \]

**Remark 4.** Note that \( D^*(\delta; f, g) \) is just a lower boundary of all asymptotically achievable \( (D, \delta) \) with bandwidth expansion factor \( \rho \).

The following is a simplified characterization of the CJSCE performance of encoders and decoders, due to [1].

**Theorem 5.** (Covering Converse) If a \( (k, n; D, \delta) \) CJSCE exists, then
\[ K(k, (1 - D)k - 1) \geq K(n, (1 - \delta)n - 1) \]
\[ K(k, Dk) \leq K(n, \delta n) \]

**Proof:** Let \( C \subset \mathbb{F}_2^n \) be a minimal \( K(n, \delta n) \) covering. Partition \( \mathbb{F}_2^n \) into \( \{ U_c : c \in C \} \) with \( \text{rad}(U_c) \leq \delta n \) for all \( c \). By the CJSCE condition, \( \{ f^{-1}U_c \} \) is a partition of \( \mathbb{F}_2^k \) with \( \text{rad}(f^{-1}U_c) \leq Dk \). For each \( c \) choose \( c' \) to be the minimizer achieving \( \text{rad}(f^{-1}U_c) \). Let \( C' = \{ c' \} \), then \( r_{\text{con}}(C') \leq Dk \)

1Note: \( A_1(n, r) = A(n, 2r + 1) \).
and thusly $K(k, Dk) \leq |C'| = |C| = K(n, \delta n)$. The second statement follows by Theorem 9.

Asymptotically this yields a lower-bound on $D^*(\delta; \rho)$ given by the following function:

$$D_{cov}(\delta; \rho) = \begin{cases} \frac{1}{m} \delta - \left(1 - \rho - (1 - \delta)|^+ \right) \delta < \frac{1}{2} \\ \frac{1}{m} \left(1 - \delta - (1 - \delta)|^+ \right) \delta \geq \frac{1}{2} \end{cases}.$$

It should be noted that, for $1/2 < \delta \leq 1$, $D_{cov}(\delta; \rho)$ is monotonically increasing in $\rho$ with $D_{cov}(\delta; \rho) > 1$ for $\rho > 1$ and

$$\lim_{\rho \to 0} D_{cov}(\delta; \rho) = \frac{1}{2} \quad \lim_{\rho \to \infty} D_{cov}(\delta; \rho) = \begin{cases} 0 & \frac{1}{2} \leq \delta < 1 \\ \frac{1}{2} & 1/2 < \delta \leq 1 \end{cases}.$$

Combined with and (18) this shows

$$\lim_{\rho \to 0} D^*(\delta; \rho) = \frac{1}{2} \quad \forall \delta < \frac{1}{2} \quad \lim_{\rho \to \infty} D^*(\delta; \rho) = 1 \quad \forall \delta > \frac{1}{2}.$$

B. Packing Converse

The coding converse has a natural extension to multiple packings.

**Theorem 6.** Let $f$ be a $(k, n; D, \delta)$ CJSCC. If an $L$-multiple packing of radius $Dk$ exists in $F_2^n$, then its image under $f$ is an $L$-multiple packing of radius $\delta n$ and

$$A_L(k, Dk) \leq LA_L(n, \delta n).$$

**Proof:** Let $C$ be an $L$-multiple packing of radius $Dk$. Suppose $f(C)$ is not an $L$-multiple packing of radius $\delta n$. Then there exists $y_0 \in F_2^n$ such that $|f(C) \cap B_{\delta n}(y_0)| > L$. By construction $\text{rad}(f(C) \cap B_{\delta n}(y_0)) \leq \delta n$. Thus there exists $x_0$ such that $f^{-1}(f(C) \cap B_{\delta n}(y_0)) \subset B_{Dk}(x_0)$. For all $c_0 \in C$,

$$f(c_0) \in f(C) \cap B \implies c_0 \in f^{-1}(f(C) \cap B).$$

Hence $|C \cap B_D(x_0)| \geq |C \cap f^{-1}(f(C) \cap B_{\delta n}(y_0))| \geq |f(C) \cap B_{\delta n}(y_0)| > L$, a contradiction. The bound follows from $|f^{-1}(f(c_0))| \leq L$.

With $L = 1$, Theorem 6 is asymptotically equivalent to (4), and the novelty here is using it for $\delta > 1/4$ or $L > 1$. Blinovsky showed explicit upper and lower bounds

$$R_{ach}(L, \delta) + o(1) \leq \frac{1}{n} \log A_L(n, n\delta) \leq R_{con}(L, \delta) + o(1).$$

The upper bound was improved in [7] for $L = 2$ and [8] for odd $L$. As per the numerical evaluations given in Section VI, for $0 \leq \delta < 1/4$ the best bound is given by $L = 2$.

The following “staircase” converse shows that coding with greater than unit bandwidth expansion factor probably yields no improvement in the region $1/4 < \delta < 1/2$.

**Proposition 7.** Let $\rho > 0$.

i) (Plotkin-Levenshtein) Provided an infinite sequence of Hadamard matrices exists in $k$-space, for all $m \in \mathbb{N}$,

$$D^* \left( \frac{1}{2m-1} m \right) \geq \frac{1}{2m-1}. \quad (7)$$

ii) (Blinovsky) For all $\ell \in \mathbb{N}$,

$$D^* \left( \frac{1}{2} \left( 1 - \left( \frac{2\ell}{\ell} \right)^2 \right)^{-1} \right) \geq \frac{1}{2} \left( 1 - \left( \frac{2\ell}{\ell} \right)^2 \right).$$

**Proof:** (Sketch)

i) Evaluate the coding converse using the Plotkin-Levenshtein solution to $A(n, d)$, [4] ch. 7.3.

ii) Evaluate the endpoint for Blinov’s upper and lower bounds for ranging values of $L$.

C. Linear Encoder Converse

A linear $(k, n; D, \delta)$ CJSCC is a $n \times k$ matrix $A \in F_2^{n \times k}$ and satisfies, for all $x \in F_2^k$,

$$\text{wt}(x) > 2D + 1 \implies \text{wt}(Ax) \geq 2[\delta n] + 1. \quad (8)$$

For linear encoders we can sharpen the double staircase result of the previous section:

**Theorem 8.** For all $\rho > 0$ and $1/4 \leq \delta \leq 1/2$, the asymptotic distortion for linear encoders satisfies

$$D^*_{lin}(\delta; \rho) \geq \delta.$$

**Proof.** Boundary cases are handled by (7), so assume $\delta \in (1/4, 1/2)$. Let $A_k$ be a sequence of linear $(k, [\delta k]; Dk, \delta)$ CJSCC with $D = \lim_{n \to \infty} D_k$. Let $X^k$ be i.i.d. Bernoulli$(q)$ vector with $q > \max(2D, 1/2)$. By the law of large numbers

$$\lim_{n \to \infty} \mathbb{P}[|X^k| \geq 2Dk + 1] = 1. \quad (9)$$

On the other hand, we have

$$\mathbb{P}[|X^k| \geq 2Dk + 1] \leq \mathbb{P}[|A_k X^k| \geq 2[\delta k] + 1] \leq \frac{\mathbb{E}[|A_k X^k|]}{2[\delta k] + 1} \leq \frac{qpk}{2\delta pk - 3}, \quad (10)$$

where (10) is from (8), (11) is by Chebyshev, and (12) follows because for arbitrary coefficients $a_j$ we have

$$\mathbb{P} \left[ \sum a_j X_j = 1 \right] \leq q$$

when $q \geq 1/2$ and thus the average Hamming weight of the vector $A_k X^k$ does not exceed $qpk$. Putting together (9) and (12) we conclude that $q \geq 2\delta$, and thusly $D \geq \delta$. \hfill □

D. Repetition Scheme Converse

An $L$-repetition scheme $f^{\otimes L} : F_2^k \to F_2^{Lk}$ is the $L$ times concatenation of a based code $f : F_2^k \to F_2^L$. Previous results [1] demonstrated that repeating a small base code may yield good CJSCC. The asymptotic performance of $L$-repetition schemes is characterized in [6, Thm. 2] where it is shown that, for all $\rho > 0$, $k \in \mathbb{N}$ and $f : F_2^k \to F_2^{[\delta k]}$, the limit function

$$D(\delta; f^\infty) := \lim_{L \to \infty} D(\delta; f^{\otimes L})$$
Duality

Let the optimization is defined as follows:

\[ \text{admissible channel parameter for a given distortion. This parameter. Conversely, the dual problem asks for the largest problem seeks the minimal distortion for a given channel bandwidth.} \]

Speaking asymptotically, the point \( \delta \) is achievable at \( C_{\text{JSCC}} \) if and only if there exists a CJSCC if and only if there exists a

\[ \delta \rho \leq \frac{1}{k} \] 

Define the new encoder/decoder pair as follows:

\[ f_1(s) \triangleq g(s), \quad g_1(x) \triangleq 1^n + f(x + 1^k), \quad s \in \mathbb{F}_2^n, x \in \mathbb{F}_2^k, \]

where \( 1^k \) and \( 1^n \) are the all-one vectors with respective dimensions. Let \( e \in \mathbb{F}_2^k \) s.t. \( |e| < k - D \), then we have

\[ |g_1(f_1(s) + e)| = |1^n + f(g(s) + e + 1^k) - s| = n - |f(g(s) + e') - s| < n - n\delta, \]

where we defined \( e' \triangleq e + 1^k \) and applied (14). Clearly, (15) shows that \( (f_1, g_1) \) defines a \( (n, k; D_1, \delta_1) \) CJSCC.

V. \( D - \delta \) Trade off as a Function of \( \rho \)

In the information theoretic setting there is both monotonicity and continuity in \( \rho \). This section partially extends these properties to the combinatorial setup. A basis for this analysis is the performance of CJSCCs combined by composition.

**Lemma 11. (Composition of Encoders)** Let \( k, m, n \in \mathbb{N} \) and \( 0 \leq \delta \leq 1 \). For all \( f_1 : \mathbb{F}_2^k \to \mathbb{F}_2^n \) and \( f_2 : \mathbb{F}_2^n \to \mathbb{F}_2^m \)

\[ D(\delta; f_2 \circ f_1) \leq D(D(\delta; f_2); f_1) \]

and

\[ D^*(\delta; k, n) \leq D^*(D^*(\delta; m, n); k, m). \]

**Proof:** Let \( g_1 \) and \( g_2 \) be the optimal Chebyshev decoders. Then \( d((f_2 \circ f_1)(x), y) \leq \delta n \) implies \( d(f_1(x), g_2(y)) \leq D(\delta; f_2)m \) implies \( d(x, (g_1 \circ g_2)(y)) \leq D(D(\delta; f_2); f_1)k \). The second statement follows immediately from the first by using the optimal encoders.

Of particular interest is the canonical admissible \( \rho \) source-channel sequence \( (k, |\rho k|) \). To facilitate in the analysis of such sequences we define upper and lower limits

\[ \overline{E}(\delta; \rho) := \limsup_{k \to \infty} D(\delta; k, |\rho k|) \]

\[ \underline{E}(\delta; \rho) := \liminf_{k \to \infty} D(\delta; k, |\rho k|). \]

The notation \( E(\delta; \rho) \) is used in statements that apply to both. An immediate application of the composition Lemma shows that \( E(\delta; \rho) \) is more or less impervious to small deviations in \( \rho \) and provides a limited monotonicity result.

- For all \( \rho > 0 \) and \( a, b \in \mathbb{N} \)

\[ \underline{E}(\delta^-; \rho) \leq \limsup_{k \to \infty} D(\delta; k + a, |\rho k| + b) \leq \overline{E}(\delta^+; \rho) \]

\[ \underline{E}(\delta^-; \rho) \leq \liminf_{k \to \infty} D(\delta; k + a, |\rho k| + b) \leq \overline{E}(\delta^+; \rho). \]

- If \( \rho, \tau > 0 \) and \( \limsup_{k \to \infty} D(\delta; |\tau k|, |\rho k|) \leq \delta \), then \( \overline{E}(\delta; \rho) \leq \overline{E}(\delta^+; \tau) \).

- If \( \rho, \tau \in \mathbb{Q} \), then

\[ \limsup_{k \to \infty} D(\delta; |\tau k|, |\rho k|) \leq \overline{E}(\delta^+; \rho/\tau). \]

VI. Discussion

In this section we discuss how our converse results compare against simple achievability results.
A. Basic CJSCCs

The following is a collection of basic CJSCCs that we will compare our converse bounds against:

- (Pseudo)-identity code $I_{k,n}$: maps $k \to \min\{k, n\}$ bits followed by $n - \min\{k, n\}$ zeros. The distortion of the (pseudo)-identity map is
  \[ D(\delta; I_{k,n}) = (\delta n + \max\{0, k-n\})/k. \tag{16} \]
- Repetition code $R_{\rho,k}$: for $\rho \in \mathbb{N}$ each bit is repeated $\rho$ times. For odd $\rho$ the distortion of the $\rho$-repetition code is better than (16):
  \[ D(\delta; R_{\rho,k}) = \frac{\ceil{\delta / \rho}k}{\rho} - 1 = \frac{\delta_\rho k}{\rho}. \tag{17} \]
- Dual repetition codes: For $\rho$ equal to reciprocal of the odd integer, one may define a small code $f_1 : \mathbb{F}_2^\frac{k}{2} \to \mathbb{F}_2$ to be a majority vote. Repeating this code asymptotically achieves
  \[ D = 1 - (1 - \delta) \frac{1 + \rho}{2} \text{.} \]

This is an improvement over the pseudo-identity for large $\delta$.

- Separated code $S_{M,k,n}$: Given a covering $C_1 \subset \mathbb{F}_2^k$ (of radius $kD$) and a packing $C_2 \subset \mathbb{F}_2^n$ (of radius $\delta n$) of equal cardinality $M$, the separation code takes $x \in \mathbb{F}_2^k$, finds the closest point in $C_1$ and outputs a corresponding point from $C_2$. Asymptotically, these codes achieve [1, Sec. III-C],
  \[ D = \begin{cases} \frac{1}{2} - h^{-1}(1 - \rho(1 - h(2\delta)^+)) & 0 \leq \delta < \frac{1}{2}, \\ \frac{1}{2} \leq \delta < 1. \end{cases} \tag{18} \]
- Dual separated codes $S_{M,k,n}^*$: Given the packing $C_1 \subset \mathbb{F}_2^k$ of radius $k(1-D)$ and covering $C_2 \subset \mathbb{F}_2^n$ of radius $n(1-\delta)$ of the same cardinality $M$, the encoder takes $x \in \mathbb{F}_2^k$, finds the closest point in $C_1$ and outputs a corresponding point in $C_2$.

To verify that this construction indeed yields a $(k, n; D, \delta)$ CJSCC we will use (3). Indeed, by (2) every ball of radius $n\delta$ must miss at least one point of $C_2$. Thus, $f^{-1}(B(y, n\delta))$ must exclude a ball of radius $k(1-D)$, and thus again by (2) is contained in a ball of radius $kD$, QED.

Asymptotically, these codes achieve:
  \[ D = 1 - \frac{1}{2}h^{-1}((1 - \rho(1 - h(\delta)^+)), \frac{1}{2} \leq \delta \leq 1. \tag{19} \]

B. Comparison for $\rho = 3$

Figure 2 gives the best known converse and achievability bounds for bandwidth expansion factor $\rho = 3$. The dotted black line represents the uncoded or $\rho = 1$ case where the identity scheme is optimal. Deviation from this line is of interest.

The achievability bound is given as follows:
- for $0 \leq \delta < 0.185$ the best code is the separated code (18). 

Fig. 2. State of the art for achievability and converse bounds when $\rho = 3$.

- for $0.184 \leq \delta < 1/3$ – the [3, 1, 3]-repetition code (17).
- for $1/3 \leq \delta < 1/2$, – the separated code with $M = 2$, see (18))
- for $1/2 \leq \delta < 1$, – the dual separated code (19).

The converse bound is given as follows:
- for $0 \leq \delta \leq 1/4$, the best bound is Theorem 6 using $L = 2$ and the upper bound from [7],
- for $1/4 < \delta \leq 1/2$, the interlacing of the bounds in Proposition 7 (double staircase) and,
- for $1/2 < \delta \leq 1$, Theorem 5 the covering converse.

We also note that together the dual-separated codes and the covering converse establish that the supremum of $\delta$ for which distortion $D < 1$ is asymptotically achievable at bandwidth expansion factor $\rho$ is given by
  \[ \delta^*(1; \rho) = 1 - h^{-1} \left( \left| 1 - \frac{1}{\rho} \right|^+ \right) \tag{20} \]

REFERENCES


3Added in print: The bound in the first interval is superceded by [9, Theorem 1], while the double staircase can be replaced by the straight line by [9, Theorem 3].