On Adversarial Joint Source Channel Coding

Arya Mazumdar
Department of ECE
University of Minnesota – Twin Cities
Minneapolis, MN 55455
Email: arya@umn.edu.

Ankit Singh Rawat
Department of ECE
The University of Texas at Austin
Austin, TX 78712
Email: ankitsr@utexas.edu.

Abstract—In a joint source-channel coding scheme, a single mapping is used to perform both the tasks of data compression and channel coding in a combined way, rather than performing them separately. Usually for simple iid sources and channels, separation of the tasks is information theoretically optimal.

In an adversarial joint source-channel coding scenario, instead of a stochastic channel, an adversary introduces a set of bounded number of errors/erasures. It has been shown recently that, even in the simplest cases of such adversarial models, separation is suboptimal and characterizing the fundamental limits is difficult. In this paper, we study several properties of such adversarial joint source-channel schemes. We show optimality of separation in some situations, provide simple joint schemes that beat separation in others, and give new bounds on the rate of such coding.

I. Introduction

The characterization of the fundamental limits of information transmission over a stochastic channel, while ensuring the reconstruction of the information, is the most celebrated result in information theory. The requirement of the recovery of information within certain distortion (under a suitable distortion metric) is a natural generalization of the setting where one requires to recover the transmitted sequence exactly with vanishing probability of error [1]. It is known that for a large class of stochastic source and channel models, the separation scheme, which involves first compressing the source to a tolerable distortion with optimal source code and then transmitting compressed source encoded by an optimal error-correcting channel code, does not incur any loss in terms of the rate [2].

The problem of error free transmission over adversarial (combinatorial) channels has received a significant amount of attention from the coding theory community. As compared to the stochastic setting, the fundamental limits of the rate of the communication schemes for adversarial channels are still unknown [3]. Similarly to the stochastic setting, in certain applications, it is natural to study the problem of transmitting source information with certain fidelity constraint over adversarial channels. Recently, [4]–[6] has formally defined and studied this problem. In particular, these papers study the setting where one is interested in encoding a k-length binary source sequence into an n-length sequence and transmitting it over an adversarial channel that can corrupt upto δn symbols in the transmitted sequence. The objective is to recover the

source sequence within Dk Hamming distortion. The main take away from the work in [4] is that contrary to the stochastic case, the separation is not always optimal in adversarial setting.

In this paper we build on the work of [4], [5] and further the study of adversarial joint source-channel coding (JSCC) problem for Hamming distortion metric in the presence of adversarial errors (bit flips). In Section III, we derive an upper bound on the rate of an adversarial JSSC for a source that consists of binary sequences with $p \in (0, 0.5)$ fraction of 1s. For the general source, where we have $\mathbb{F}_2^k$ as the set of source sequences, we present multiple simple coding schemes in Section IV that outperform separation scheme for certain range of channel parameters δ. In Section V, we explore necessary conditions on the distance distribution of the transmitted codewords of an adversarial JSCC. These conditions are helpful in obtaining new upper bounds on the rate of adversarial JSCC schemes.

We then shift our focus to the setting where the adversarial channel introduces erasures in a transmitted codeword as opposed to bit-flips. The erasure setting models various applications including distributed storage systems and video delivery. For this setting, we consider both Hamming distortion metric and erasure distortion metric in Section VI-A and Section VI-B, respectively. Similarly to the case with adversarial errors [4], we obtain upper bounds on the rate of an adversarial joint source channel coding scheme which recovers a source sequence within Dk distortion in the presence of at most ϵn adversarial erasures. We present a replication based JSCC scheme for this setting. We conclude the paper in Section VIII with some comments on future work.

II. Definitions and preliminaries

Throughout this text $d_H(\cdot, \cdot)$ will denote the Hamming distance between the argument vectors and $wt(\cdot)$ will denote the Hamming weight of the argument.

A pair of mappings $f : \{0, 1\}^k \rightarrow \{0, 1\}^n$ and $g : \{0, 1\}^n \rightarrow \{0, 1\}^k$ are called together a (n, k, d_1, d_2)-JSCC, if $\forall x \in \{0, 1\}^k$ and $\forall z \in \{0, 1\}^n : wt(z) \leq d_2$,

$$d_H(g(f(x) + z), x) \leq d_1.$$

Sometime we refer to the image of f as the *code*.

In most of the parts of this paper we will be considering only asymptotic regime where $n \rightarrow \infty$ and k scaling linearly

with n. Define,

$$R \triangleq \frac{k}{n}; \quad \rho \triangleq \frac{n}{k}$$

as the rate of transmission and bandwidth expansion respectively. In the asymptotic regime we also assume $d_1 = Dk$ and $d_2 = \delta n$, $D, \delta \in [0, 1]$. In short we write an $(n, k, Dk, \delta n)$-JSCC as a (D, δ)-JSCC when the other parameters are clear from the context.

The following two converse bounds [4] will be referenced frequently throughout the text.

Proposition 1 (Information Theoretic Converse): For any (n, k, d_1, d_2)-JSCC,

$$2^k \sum_{i=0}^{d_2} \binom{n}{i} \leq 2^n \sum_{i=0}^{d_1} \binom{k}{i}.$$

Asymptotic form of the bound for any (D, δ)-JSCC is

$$1 - h_B(D) \leq \rho(1 - h_B(\delta)).$$

Proposition 2 (Coding Theoretic Converse): For any (n, k, d_1, d_2)-JSCC,

$$A(k, 2d_1 + 1) \leq A(n, 2d_2 + 1),$$

where $A(n, d)$ is the maximum size of an error-correcting code of length n and minimum distance d. Suppose,

$$\lim_{n \to \infty} \frac{A(n, \tau n)}{n} = R_0(\tau),$$

though it is not known whether the limit exists. Asymptotically, for any (D, δ)-JSCC,

$$R_0(2D) \leq \rho R_0(2\delta).$$

It is known that,

$$R_{GV}(\tau) \leq R_0(\tau) \leq R_{MRRW}(\tau), \tag{1}$$

where the MRRW II bound [7] is

$$R_{MRRW}(\tau) \equiv \min_{0 < \alpha \leq 1 - 2\tau} 1 + \hat{h}(\alpha^2) - \hat{h}(\alpha^2 + 2\tau\alpha + 2\tau), \tag{2}$$

with $\hat{h}(x) = h_B(1/2 - 1/2\sqrt{1-x})$, and the Gilbert-Varshamov bound [8] is

$$R_{GV}(\tau) \equiv 1 - h_B(\tau). \tag{3}$$

A *separation scheme* for JSCC is a scheme where a covering code with covering radius d_1 is first used to compress k-bit vectors and then an error correcting code with distance $2d_2 + 1$ is used for noise protection. For any separation scheme, asymptotically, we have,

$$1 - h_B(D) = \rho R_0(2\delta). \tag{4}$$

One of the main contribution of [4] is to show that separation is not optimal in general.

III. OPTIMALITY OF SEPARATION

In this section we primarily focus on a very common scenario, where, instead of considering a source that produces all possible binary sequences, we consider p-biased source (for $p < \frac{1}{2}$). This source consists of all vectors of Hamming weight pk in $\mathbb{F}_2^k$. We show that, when the source is very sparse, separation is close to being the optimal method.

If we want a (D, δ)-JSCC for this source, then analogous to the coding theoretic converse, we have

$$A(k, 2Dk + 2, pk) \leq A(n, 2\delta n + 1), \tag{5}$$

where $A(n, d, w)$ is the maximum possible size of a constant weight binary code with length n, distance d and weight w. The claim follows from the fact that, for any (D, δ)-JSCC, any two source vectors that are distance more than $2Dk$ apart, must map to vectors that are more than $2\delta n$ apart. Also, note that, a constant weight code must have only even distances between codewords.

The main claim of this section is the following.

Theorem 3: If we have a (D, δ)-JSCC for a p-biased source, then,

$$R \leq \frac{R_0(2\delta)}{h_B(p) - p h_B\left(\frac{D}{p}\right) - (1-p) h_B\left(\frac{D}{1-p}\right)}. \tag{6}$$

Proof: Simple Gilbert-Varshamov argument gives [9],

$$A(n, d, w) \geq \frac{\binom{n}{w}}{\sum_{i=0}^{d/2} \binom{w}{i}\binom{n-w}{i}}. \tag{7}$$

Therefore, using (5) we must have,

$$nR_0(2\delta) \geq k\Big(h_B(p) - \max_{0 \leq \alpha \leq D} \Big\{ p h_B\Big(\frac{\alpha}{p}\Big) + (1-p) h_B\Big(\frac{\alpha}{1-p}\Big) \Big\} \Big).$$

The term within the maximization is a concave function of α and it is maximized for $\alpha = p(1-p)$. Assuming $D < p(1-p)$, the term will maximize at $\alpha = D$. We substitute this value to obtain the theorem. $\blacksquare$

Note that, separation achieves,

$$R \geq \frac{R_0(2\delta)}{h_B(p) - h_B(D)}, \tag{8}$$

as there exists a covering code for pk-biased source with rate $h_B(p) - h_B(D)$. Comparing the above with (6), we can say that separation is close to being optimal as $p \to 0$ (as D is dominated by p, it also goes to 0) for all δ. Also note that, optimality of separation for $p \to 0$ is not evident when we use the information theoretic converse, $R \leq \frac{1 - h_B(\delta)}{h_B(p) - h_B(D)}$, as the numerators are dependent on δ.

In the following example, we highlight another setting where separation is optimal.

Example: Here we assume the general source with all possible binary sequences in $\mathbb{F}_2^k$ and focus on small number of errors. Consider an $(n, k, 1, 2)$-JSCC. Assume, $k + 1$ and $n + 1$ are both powers of two. For a separation scheme,

we can use a Hamming code (that is the optimal code with covering radius 1) of length k and dimension $k - \log(k+1)$ as the covering code and a 2-error correcting BCH code (of length n and dimension $n - 2\log(n+1)$) as the channel code. Hence, separation achieves $\frac{2^k}{k+1} \approx \frac{2^n}{(n+1)^2}$. As 2-error correcting BCH codes are in general optimal (quasi-perfect) family for large n, we have from the coding theoretic converse, $\frac{2^k}{k+1} \leq \frac{2^n}{(n+1)^2}$. This shows that separation is optimal in this case. For the general case when d_2 is a constant, we may use a d_2-error correcting BCH code of length n and dimension $n - d_2 \log(n+1)$. Again, the converse result helps us prove that all such separation schemes provide order optimal $(n, k, 1, d_2)$-JSCC.

IV. SCHEMES THAT BEAT SEPARATION

In this section, we present simple (D, δ)-JSCC schemes that outperform the separation scheme.

A. Systematic code for Joint Source Channel Coding

Assume that $\rho = \frac{n}{k} \geq 1$. Given a source sequence in $\mathbb{F}_2^k$, we encode it using a systematic linear code that attain the GV bound, *i.e.*, the code with minimum distance $2\Delta n$ and the rate $R = \frac{1}{\rho} = 1 - h_B(2\Delta)$. Assume that the first k bits in a codeword constiue the systematic part. Now, given $\delta n \leq \Delta n$ errors in the received sequence, we can correct all these errors leading to the Hamming distortion $Dk = 0$. In the case with $\delta n > \Delta n$ errors, we can simply output the first k bits (systematic part) of the received sequence as an estimate of the original source sequence. In the worst case, these k bits contain δn errors, leading to the Hamming distortion $Dk = \delta n$ or $D = \rho\delta$.

On the other hand the separation scheme, where we use an optimal source code with Hamming distortion Dk and a code attaining the GV bound with minimum distance $2\delta n$, gives us $D = h_B^{-1}(1 - \rho(1 - h_B(2\delta)))$ [4]. Therfore, combining both these coding schemes, we can design a (D, δ)-JSSC with

$$D = \begin{cases} 0, & \text{if } \delta \leq \frac{1}{2}h_B^{-1}(1 - 1/\rho); \\ \min\{\rho\delta, h_B^{-1}(1 - \rho(1 - h_B(2\delta)))\}, & \text{otherwise.} \end{cases}$$

Fig. 1 illustrates the tradeoff between D and δ achieved by the JSCC scheme described above. As it is clear from the figure, this scheme outperforms the separation scheme starting moderate values of δ.

This naive systematic scheme can actually be improved. We illustrate this via an example. Suppose, $\rho = \frac{20}{11}$. We simply repeat $\frac{9k}{11}$ bits twice and leave $\frac{2k}{11}$ bits uncoded to obtain a total of $n = \frac{20k}{11}$ bits. As long as $\delta n \leq \frac{2k}{11}$ or $\delta \leq \frac{1}{10}$, the adversary will target only uncoded bits and as a result we will have $D = \rho\delta$. However, when $\delta > \frac{1}{10}$, the adversary has to introduce errors in repeated bits. In this case $D = 1 - \frac{\rho(1-\delta)}{2}$ (this is the average number of remaining errors, see the randomized decoder below for optimum adversary strategy and decoding). In particular, for $\rho \geq 2$ we can use the following repetition code and the randomized decoding.
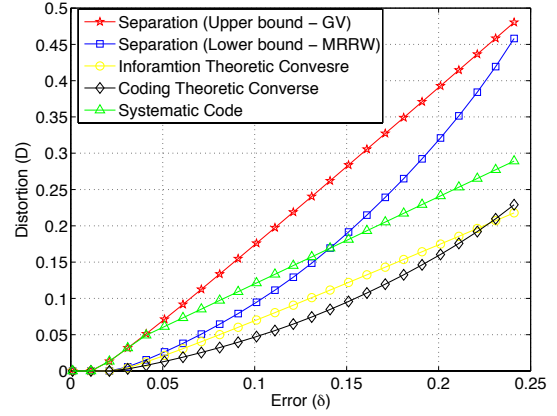

Fig. 1. Performance of systematic code approach at $\rho = 1.2$.

B. Repetition with randomized decoder

For an integer valued ρ, we can utilize ρ-repetition as transmission scheme. Here, we present such a repetition scheme with a randomized decoder that allows us to obtain a JSCC with $D = \delta$. For ρ repetition scheme, we need to transmit k blocks from the set $\{00\ldots0, 11\ldots1\} \in \mathbb{F}_2^\rho$. Given the received vector $y = (y_1, y_2, \ldots, y_k) \in \mathbb{F}_2^{\rho k}$, we employ a randomized decoder which maps the ρ-bit long block y_i to message bit x_i such that

$$\Pr(x_i = 0) = \frac{\rho - \text{wt}(y_i)}{\rho}.$$

Assuming that there are ℓ_i errors in the ith received block, the decoded makes an error in decoding the ith message bit with probability $\frac{\ell_i}{\rho}$. Therefore, the randomized decoder leads to the expected Hamming distortion of

$$\sum_{i=1}^{k} \frac{\ell_i}{\rho} \leq \frac{\delta n}{\rho} = \delta k.$$

We can further argue that with high probability the randomized decoder leads to the Hamming distortion of at most $\sim \delta$ (by Chernoff bound). Recently, Young and Polyanskiy (personal communication) have shown that for $\frac{1}{4} \leq \delta \leq \frac{1}{2}$, a linear encoding scheme can not achieve $D < \delta$. This implies that the repetition scheme with the randomized decoder described above is optimal for the range $\frac{1}{4} \leq \delta \leq \frac{1}{2}$.

C. Project and transmit scheme

For $\rho = \frac{n}{k} \leq 1$, as a JSSC we can transmit any $n = \rho k$ coordinates of a k-length source sequence. This operation of projecting source sequence to ρk coordinates leads to at most $(1 - \rho)k$ Hamming distortion during the reconstruction process. Note that we receive at least $(1-\delta)\rho k$ symbols of the source sequence that remain intact even after the introduction of at most $\delta n = \delta \rho k$ adversarial errors. This simple scheme ensures that the source sequence is recovered with at most $Dk = (1 - \rho)k + \delta\rho k$ Hamming distortion.

Using the separation scheme, we get $D \leq h_B^{-1}(1 - \rho(1 - h_B(2\delta)))$. This is worse than what we obtain from the projection scheme when ρ is close to 1 (and δ is large enough).

V. GEOMETRY OF (D, δ)-JSCC

In this section we state two necessary geometric conditions on a codebook associated with a (D, δ)-JSSC. Given a code $\mathcal{C} \subseteq \mathbb{F}_2^k$, the *distance distribution* of $\mathcal{C}$ is defined as follows. For $i = 0, 1, 2, \ldots, n$,

$$A_i \triangleq \frac{1}{|\mathcal{C}|} |\{(c_1, c_2) \in \mathcal{C}^2 : d_H(c_1, c_2) = i\}|. \quad (9)$$

As can be seen, $A_0 = 1$. For a code $\mathcal{C}$ to be able to be the image of a (D, δ)-JSCC with a suitable encoding map, it must satisfy the following conditions.

Proposition 4: For any (k, n, D, δ)-JSCC $\mathcal{C}$, we must have

$$\sum_{w=0}^{2\delta n} A_w \leq \sum_{i=0}^{2Dk} \binom{k}{i}, \quad (10)$$

$$\frac{\sum_{i=0}^{2\delta n} A_i \binom{i}{\lceil i/2 \rceil} \binom{n-i}{\delta n - \lceil i/2 \rceil}}{\binom{n}{\delta n}} \leq \sum_{j=0}^{Dk} \binom{k}{j}. \quad (11)$$

Proof (sketch): Let $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ be an encoding map for (D, δ)-JSCC which maps $\mathbb{F}_2^k$ to the sequences in $\mathcal{C}$. Now, the proof of (10) follows from the fact that two source sequences that are more than $2Dk$ Hamming distance apart can not be mapped by f to the codewords that have Hamming distance at most $2\delta n$ between them.

For (11), we note that the number of codewords from $\mathcal{C}$ within any Hamming ball in $\mathbb{F}_2^n$ of radius δn shoul be at most size of an Hamming ball in $\mathbb{F}_2^k$ of radius Dk. Otherwise the optimal decoder can not guarantee the distortion of at most Dk. We next compute the expected number of codewords in a Hamming ball of radius δn in $\mathbb{F}_2^n$ that is at δn Hamming distance from a codeword of the underlying JSSC. It turns out to be the following quantity.

$$N(\mathcal{C}, \delta) = \frac{\sum_{c \in \mathcal{C}} \sum_{x \in S(c, \delta n)} \sum_{\tilde{c} \in \mathcal{C}} \mathbb{1}_{\{\tilde{c} \in B(x, \delta n)\}}}{|\mathcal{C}| \binom{n}{\delta n}}, \quad (12)$$

where $S(c, \delta n)$ denotes the Hamming sphere of radius δn which is centered at c. Note that (details omitted),

$$N(\mathcal{C}, \delta) = \frac{\sum_{i=0}^{2\delta n} A_i \sum_{a=\lceil i/2 \rceil}^{\min\{\delta n, i\}} \binom{i}{a} \binom{n-i}{\delta n - a}}{\binom{n}{\delta n}}. \quad (13)$$

Hence, there is a Hamming ball in $\mathbb{F}_2^n$ of radius δn (centered at distance δn from a codeword in $\mathcal{C}$) such that it contains $N(\mathcal{C}, \delta)$ codewords in $\mathcal{C}$. Now, (11) follows from bounding this quantity by $\sum_{j=0}^{Dk} \binom{k}{j}$. ∎

Remark 1: It turns out that (10), if used in the Delsarte linear programming (LP) bounds of codes [10], leads to the coding converse (details omitted). By analyzing the stronger inequality (11) and using the LP bounds, we stand to get improvement over the coding theoretic converse.

VI. JSCC FOR ADVERSARIAL ERASURES

In the rest of this paper, we consider the setting with adversarial erasures as opposed to errors.

A. JSCC for Hamming distortion and adversarial erasures

Our objective is to devise a joint source-channel coding scheme such that source sequences can reconstructed within a certain Hamming distortion in the presence of adversarial erasures. In particular, an encoding map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ and an associated decoding map $g : \{\mathbb{F}_2 \cup \{\star\}\}^n \to \mathbb{F}_2^k$ constitutes a (D, ϵ)-joint source channel code against erasures (JSSCE) if, for every $x \in \mathbb{F}_2^k$, we have,

$$d_H(x, g(f(x) * e)) \leq Dk, \quad \forall e \in \{1, \star\}^n \text{ s.t. } w_E(e) \leq \epsilon n.$$

Here, $\star$ denotes an erasure and the operation $*$ between two vectors is an elementwise operation which is defined as follows. For $a \in \mathbb{F}_2$, we have $a * \star = \star$ and $a * 1 = a$. Moreover, $w_E(e)$ counts the number of coordinates in the vector e that have $\star$ as their value.

1) Converse results for JSCCE with Hamming distortion: Here, we employ the ideas similar to those used in [4] to obtain the information theoretic and coding theoretic converse results for the underlying setting with adversarial erasures and Hamming distortion metric.

First, we describe an optimal decoder for a given encoding map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ in the setting with Hamming distortion and adversarial erasures. Given the encoding map f and a received sequence $z \in \{\mathbb{F}_2 \cup \{\star\}\}^n$, we define the set $\mathcal{M}(f, z)$ as the set of source sequences mapped to those codewords, by the encoding function f, that agree with the received sequence z.

$$\mathcal{M}(f, z) = \{x : f(x) \overset{\star}{=} z\} \subseteq \mathbb{F}_2^k. \quad (14)$$

Here, $f(x) \overset{\star}{=} z$ denotes that $f(x)$ and z agree on the unerased coordinates. Next, a decoding map $g : \{\mathbb{F}_2 \cup \{\star\}\}^n \to \mathbb{F}_2^k$ is defined as follows.

$$g(z) = \text{argmin}_{y \in \mathbb{F}_2^k} \max_{x \in \mathcal{M}(f, z)} d_H(x, y). $$

Given this optimal decoder, in order to show information theoretic converse, we argue that there exists a sequence z with at most ϵn erasures which agrees on the unerased coordinates with the codewords associated with at least $2^k / 2^{n(1-\epsilon)}$ source sequences. Therefore, for the optimal decoder g to work, size of a Hamming ball of radius Dk in $\mathbb{F}_2^k$, *i.e.*, $2^{k h_B(D)}$, must be at least $\frac{2^k}{2^{n(1-\epsilon)}}$. Hence, for a (D, ϵ)-JSCCE with Hamming distortion, we have

$$\frac{k}{n} \leq \frac{1 - \epsilon}{1 - h_B(D)}. \quad (15)$$

For coding theoretic converse, we note that the encoding function $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ of a (D, δ)-JSSC must map any two source sequences that at least $2Dk+1$ Hamming distance apart in $\mathbb{F}_2^k$ to two codewords that are at least $\epsilon n + 1$ Hamming distance apart in $\mathbb{F}_2^n$. Thus,

$$A(k, 2Dk + 1) \leq A(n, \epsilon n + 1), \quad (16)$$

which in turn implies,

$$\frac{k}{n} \leq \frac{R_{MRRW}(\epsilon)}{1 - h_B(2D)}. \quad (17)$$

2) Repetition scheme for JSSCE with Hamming distortion:
Consider a ρ-repetition scheme as a JSSCE, where each bit of
the source sequence is repeated ρ times. Therefore, the length
of the codewords $n = \rho k$. For this setting, we incur distortion
in a bit of the source sequence only if adversary erases all ρ
copies of the bit in the codeword. Thus, for the ρ-repetition
scheme we have $Dk \le \frac{n\epsilon}{\rho} = \frac{\rho k \epsilon}{\rho} = \epsilon k$. Compare this
with the setting with adversarial errors; where the ρ-repetition
scheme gives us [4] $D \le \frac{2\rho\delta}{\rho+1}$. Thus in the case with errors,
increasing the repetition factor ρ leads to higher distortion. On
the other hand, the incurred distortion is independent of the
repetition factor ρ in the setting with erasures.

Remark 2: Note that, the simple randomized decoding
will result in $D = \frac{\epsilon}{2}$ with high probability for any repetition
scheme.

B. JSSC for erasure distortion and adversarial erasures

We now study the setting where our goal is to reconstruct the
source squence up to Dk *erasure distortion* from the received
sequence with at most ϵn erasures. Note that under erasure
distortion metric, we have $d(0,0) = d(1,1) = 0$; $d(1,\star) =
d(0,\star) = 1$; and $d(0,1) = d(1,0) = \infty$.

1) Converse results for JSSCE with erasure distortion: The
following result presents converse for the underlying setting.

Theorem 5: For a (D, ϵ)-JSSCE with erasure distortion
metric, we must have

$$\frac{k}{n} \le \frac{1-\epsilon}{1-D}. \tag{18}$$

and

$$\frac{k}{n} \le \frac{R_{MRRW}(\epsilon)}{1-h_B(D)}. \tag{19}$$

Proof: Again, we define an optimal decoder for the
setting. Given an encoding map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ and a received
sequence $z \in \{\mathbb{F}_2 \cup \{\star\}\}^n$, let $\mathcal{M}(f, z)$ be the set as defined in
(14). The optimal decoder $g : \{\mathbb{F}_2 \cup \{\star\}\}^n \to \{\mathbb{F}_2 \cup \{\star\}\}^k$ for
the current setting is defined as follows.

$$g(z) = \operatorname{argmin}_{y \in \{\mathbb{F}_2 \cup \{\star\}\}^k} \{w_E(y) : y \stackrel{\star}{=} m \ \forall \ m \in \mathcal{M}(f, z)\}.$$

For information theoretic converse, we know that for an
encoding map f, there exists a received sequence z with at
most ϵn erasures such that $|\mathcal{M}(f, z)| \ge \frac{2^k}{2^{n(1-\epsilon)}}$. Therefore,
given z as the received sequence, the optimal decoder g
will output a sequence with at most Dk erasures only if
$\frac{2^k}{2^{n(1-\epsilon)}} \le |\mathcal{M}(f, z)| \le 2^{Dk}$.

Eq. (19) follows from the arguments similar to those em-
ployed before. ∎

Remark 3 (Optimal scheme for $\rho \le 1$): For $\rho \le 1$, the
project and transmit scheme presented in Sec. IV-C can be
used to achieve the erasure distortion $D = 1 - \rho + \rho\epsilon$ in the
presence of ϵn adversarial erasures. Note that this scheme is
optimal as it attains the information theoretic bound in (18).

Remark 4: The repetition scheme as described in Sec-
tion VI-A2 also works for this setting ensuring that $D = \epsilon$.

Fig. 2 highlights the gap between the performances of the
speration scheme and the project and transmit scheme as
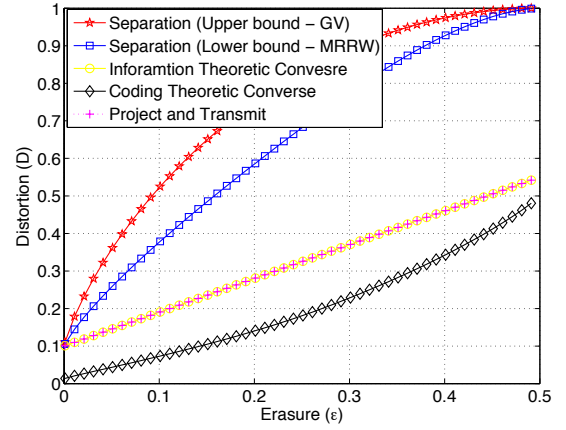(D, ϵ)-JSSCE with erasure distortion.

Fig. 2. Performance of the project and transmit scheme (coinciding with
information theoretic converse) for a (D, ϵ)-JSCCE with erasure distortion at
$\rho = 0.9$.

VII. CONCLUSION: JSCC WITH DISTANCE

It is also interesting to analyze a (D, δ)-JSCC code $\mathcal{C}$
whose minimum distance is τn. In practice, this type of code
is useful, as they perform perfect error-correction up to a
limit and then up to another limit decodes within a bounded
distortion. For these codes, the distance distribution follows,
$A_1 = A_2 = \cdots = A_{\tau n - 1} = 0$ as well as satisfies (10). It is
possible to obtain LP bounds for these codes with the extra
constraint of (10).

REFERENCES

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity
criterion," *IRE Nat. Conv. Rec.*, vol. 4, pp. 142–163, 1959.
[2] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for
Discrete Memoryless Systems.* Academic Press, Inc., 1982.
[3] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting
codes.* North-Holland, 1978.
[4] Y. Kochman, A. Mazumdar, and Y. Polyanskiy, "The adversarial joint
source-channel problem," in *Proc. of IEEE International Symposium on
Information Theory (ISIT)*, pp. 2112–2116, July 2012.
[5] Y. Kochman, A. Mazumdar, and Y. Polyanskiy, "Results on combinato-
rial joint source-channel coding," in *Proc. of IEEE Information Theory
Workshop (ITW)*, pp. 10–14, Sept 2012.
[6] A. Mazumdar, Y. Polyanskiy, and B. Saha, "On chebyshev radius of a set
in hamming space and the closest string problem," in *Information Theory
Proceedings (ISIT), 2013 IEEE International Symposium on*, pp. 1401–
1405, IEEE, 2013.
[7] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, "New upper
bounds on the rate of a code via the delsarte-macwilliams inequalities,"
Information Theory, IEEE Transactions on, vol. 23, no. 2, pp. 157–166,
1977.
[8] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting
Codes.* North-Holland, 1977.
[9] R. Graham and N. Sloane, "Lower bounds for constant weight codes,"
Information Theory, IEEE Transactions on, vol. 26, no. 1, pp. 37–43,
1980.
[10] P. Delsarte and V. I. Levenshtein, "Association schemes and coding
theory," *Information Theory, IEEE Transactions on*, vol. 44, no. 6,
pp. 2477–2504, 1998.