On the bit error rate of repeated error-correcting codes

Weihao Gao and Yury Polyanskiy

***Abstract*—Classically, error-correcting codes are studied with respect to performance metrics such as minimum distance (combinatorial) or probability of bit/block error over a given stochastic channel. In this paper, a different metric is considered. It is assumed that the block code is used to repeatedly encode user data. The resulting stream is subject to adversarial noise of given power, and the decoder is required to reproduce the data with minimal possible bit-error rate. This setup may be viewed as a combinatorial joint source-channel coding.**

Two basic results are shown for the achievable noise-distortion tradeoff: the optimal performance for decoders that are informed of the noise power, and global bounds for decoders operating in complete oblivion (with respect to noise level). General results are applied to the Hamming $[7,4,3]$ code, for which it is demonstrated (among other things) that no oblivious decoder exist that attains optimality for all noise levels simultaneously.

I. INTRODUCTION

Suppose a very large chunk of data is encoded via a fixed error-correcting block code, whose block length is significantly smaller than the total data volume. The data in turn is affected by a noise of high level, thus not permitting correcting errors perfectly. What is the best achievable tradeoff between the noise level and the (post-decoding) bit-error rate?

Such situation may arise, for example, in the forensic analysis of a severely damaged optical, magnetic or flash drive. We note that there are two different scenarios depending on whether the noise level $\delta \in [0,1]$ (the fraction of bits flipped) is known to the decoder or not. The second case presents an additional challenge as apriori it is not clear whether a given error-correcting code admits a *universal* decoder that is simultaneosly optimal for all noise levels (in the sense of minimizing the bit-error rate).

In this paper we characterize tradeoffs for both cases. The general theory is applied to the example of the Hamming $[7,4,3]$ code uncovering the following basic effects:

1) Known converse bound (r_0^{**} in [1]) is not tight.
2) No single decoder is (even asymptotically) optimal for all δ. In particular, there does not exist a decoder achieving r_0^{**} at all points.
3) For the (practical case of) small δ, the optimal decoder is not the minimum distance one.

W.G. is with the Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, 100084 China. e-mail: gwh10@mails.tsinghua.edu.cn Y.P. is with the Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139 USA. e-mail: yp@mit.edu.

This material is based upon work supported by the National Science Foundation under Grant No CCF-13-18620.

We emphasize that the last observation suggests that conventional decoders of block codes should not be used in the cases of significant defect densities.

We proceed to discussing the basic framework and some known results.

A. General Setting of Joint-Source Channel Coding

The aforementioned problem may be alternatively formalized as a combinatorial (or adversarial) joint-source channel coding (JSCC) as proposed in [1]. The gist of it for the binary source and symmetric channel (BSSC) can be summarized by the following

Definition 1: Consider a pair of maps $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ (encoder) and $g : \mathbb{F}_2^n \to \mathbb{F}_2^k$ (decoder). The distortion-noise tradeoff is the non-decreasing right-continuous function

$$D(f,g,\delta) \triangleq \max_{x\in\mathbb{F}_2^k} \max_{e:|e|\le\delta n} \frac{1}{k}|x + g(f(x)+e)| \qquad \delta\in[0,1]$$

where $|\cdot|$ denotes the Hamming weight. The tradeoff for the optimal decoder is denoted as

$$D(f,\delta) \triangleq \min_g D(f,g,\delta) \qquad \delta\in[0,1]$$

A pair (f,g) is called a (k,n,D,δ)-JSCC if $D(f,g,\delta)\le D$.

Note that the definition $D(f,\delta)$ characterizes the smallest distortion attainable for a given encoder, provided the decoder knows δ and can adapt to it. Shortly, we will also address the case when δ is unknown to the decoder (see the concept of asymptotic decoding curve below).

In this paper we focus on a particular special case of encoders obtained via repetition of a single "small code", cf. [1]. Formally, fix an arbitrary encoder given by the mapping $f : \mathbb{F}_2^u \to \mathbb{F}_2^v$ (a small code). If there are at most t errors in the block of length v, $t \in [0,v]$ the performance of the optimal decoder (knowing t) is given by the non-decreasing right-continuous function

$$r_0(t) \triangleq \max_{y\in\mathbb{F}_2^v} \mathrm{rad}(f^{-1}B_v(y,t)), \qquad (1)$$

where

$$B_n(x,\alpha) \triangleq \{x' \in \mathbb{F}_2^n : |x'-x| \le \alpha\}$$

is a Hamming ball of (possibly non-integral) radius α and

$$\mathrm{rad}(S) = \min_{x\in\mathbb{F}_2^n} \max_{y\in S} |y-x|$$

is the radius of the smallest Hamming ball enclosing the set S. Consider also an arbitrary decoder $g : \mathbb{F}_2^v \to \mathbb{F}_2^u$ and its performance curve:

$$r_g(t) \triangleq \max_{|e|\le t} \max_{x\in\mathbb{F}_2^u} |g(f(x)+e)+x|. \qquad (2)$$

Clearly

$$r_g(t) \geq r_0(t)$$

From a given code f we may construct a longer code $f^{\oplus L}$ by repetition to obtain an $\mathbb{F}_2^k \to \mathbb{F}_2^n$ code as follows, where $Lu = k, Lv = n$:

$$f^{\oplus L}(x_1, \ldots, x_L) \triangleq (f(x_1), \ldots, f(x_L)) .$$

This yields a sequence of codes with *bandwidth expansion factor* $\rho = \frac{n}{k} = \frac{v}{u}$. We want to find out the achieved distortion $D(\delta)$ as a function of the maximum crossover portion δ of the adversarial channel.

Theorem 1 ([1]): The asymptotic distortion achievable by the repetition construction satisfies

$$\liminf_{L\to\infty} D(f^{\oplus L}, \delta) \geq \frac{1}{u} r_0^{**}(\delta v) . \tag{3}$$

A block-by-block decoder g achieves

$$\lim_{L\to\infty} D(f^{\oplus L}, g^{\oplus L}, \delta) = \frac{1}{u} r_g^{**}(\delta v) , \tag{4}$$

where r_0^{**} and r_g^{**} are upper concave envelopes of r_0 and r_g respectively.

Below we extend and refine these prior results. Namely, in Section II we show how to compute the limit in (3) exactly (correcting a previous version in [2]). In Section III we present upper and lower bounds for the case of δ not known at the decoder. Finally, in Section IV we demonstrate our findings on the example of the (repetition of the) Hamming $[7, 4, 3]$ code.

II. Decoder knows δ

A. Optimal performance curve: correction to [2]

Asymptotic performance of a repetition construction is given by:

Theorem 2: Fix a small code $f : \mathbb{F}_2^u \to \mathbb{F}_2^v$ and consider the repetition construction. The limit

$$D(f^{\oplus\infty}, \delta) \triangleq \lim_{L\to\infty} D(f^{\oplus L}, \delta) \tag{5}$$

exists and is a non-negative concave continuous function of $\delta \in [0, 1]$ given by

$$D(f^{\oplus\infty}, \delta) = \frac{1}{u} \max_{P_Y} \min_{P_{\hat{S}|Y}} \max_{P_{S|Y,\hat{S}}:(*)} \mathbb{E}\left[|S - \hat{S}|\right] , \tag{6}$$

where P_Y ranges over all distributions on $\mathbb{F}_2^v$, $P_{\hat{S}|Y}$ ranges over all Markov kernels $\mathbb{F}_2^v \to \mathbb{F}_2^u$ and $P_{S|Y,\hat{S}}$ ranges over Markov kernels $\mathbb{F}_2^v \times \mathbb{F}_2^u \to \mathbb{F}_2^u$ satisfying

$$(*) \qquad \mathbb{E}\left[|f(S) - Y|\right] \leq \delta v$$

with expectations computed over

$$P_{S,Y,\hat{S}}(s, y, \hat{s}) = P_Y(y) P_{\hat{S}|Y}(\hat{s}|y) P_{S|Y,\hat{S}}(s|y, \hat{s})$$

Proof: The key step is the formula for the optimal decoder [1, Section IV.D]:

$$D(f^{\oplus L}, \delta) = \frac{1}{uL} \max_{y\in\mathbb{F}_2^{vL}} \mathrm{rad}\left((f^{\oplus L})^{-1} B_{vL}(y, \delta v L)\right) \tag{7}$$

Note that once existence of the limit is proven, concavity follows immediately. Indeed for any $L_1 + L_2 = L$, integers s_i and $y_i \in \mathbb{F}_2^{vL_i}$ with $i = 1, 2$ we have

$$B_{vL}(y_1 \oplus y_2, s_1 + s_2) \supset B_{vL_1}(y_1, s_1) \oplus B_{vL_2}(y_2, s_2) .$$

Applying $(f^{\oplus L})^{-1}$ and taking rad we get from additivity of the radius [3, Section II]:

$$\begin{aligned} D(f^{\oplus(L_1+L_2)}, \frac{s_1 + s_2}{L_1 + L_2}) &\geq \frac{L_1}{L_1 + L_2} D(f^{\oplus L_1}, \frac{s_1}{L_1}) \\ &\quad + \frac{L_2}{L_1 + L_2} D(f^{\oplus L_2}, \frac{s_2}{L_2}) . \end{aligned}$$

Since s_i are arbitrary by taking the limit $L \to \infty$ of both sides, concavity of $D(f^{\oplus\infty}, \delta)$ follows. Concavity in turn implies continuity.

We complete the proof by showing existence of the limit and formula (6). To that end, first we expanding the definition of radius in (7). Second, we represent vectors in $\mathbb{F}_2^{vL}$ as $\mathbb{F}_2^v$-valued vectors of length L, and similarly for $\mathbb{F}_2^{uL}$. Then, the expression entirely equivalent to (7) is the following:

$$D(f^{\oplus L}, \delta) = \frac{1}{u} \max_{P_Y} \min_{P_{\hat{S}|Y}} \max_{P_{S|Y,\hat{S}}:(*)} \mathbb{E}\left[|S - \hat{S}|\right] , \tag{8}$$

with optimizations satisfying the same constraints as in (6) with the following additions:

1) $P_Y(b) \in \frac{1}{L}\mathbb{Z}$ for every $b \in \mathbb{F}_2^v$
2) $P_{\hat{S}|Y}(a|b) \in \frac{1}{LP_Y(b)}\mathbb{Z}$ for every $a \in \mathbb{F}_2^u$
3) $P_{S|Y,\hat{S}}(a'|b, a) \in \frac{1}{LP_Y(b)P_{\hat{S}|Y}(a|b)}\mathbb{Z}$ for every $a' \in \mathbb{F}_2^u$

Note that since expectations appearing in constraint $(*)$ and (6) are continuous functions of P_Y, $P_{\hat{S}|Y}$ and $P_{S|Y,\hat{S}}$ we may additionally impose constraint $P_{\hat{S}Y}(a, b) \geq \frac{1}{\sqrt{L}}$. This guarantees that in the integrality constraint 3 the denominator is a large integer for each (a, b). Consequently, arbitrary kernel $P_{S|Y,\hat{S}}$ can be approximated with precision of order $\frac{1}{\sqrt{L}}$ by kernels satisfying constraint 3. Hence, in the limit as $L \to \infty$ the inner maximization in (8) can be performed without verifying integrality condition 3. Similar argument applies to $P_{\hat{S}|Y}$ and P_Y. Overall, this is a standard exercise in approximating joint distributions by L-types, see [4, Chapter 1]. ∎

III. Decoder does not know δ

A. Asymptotic decoder curves

Definition 2: A non-decreasing right-continuous function $r : [0, v] \to [0, u]$ is called an asymptotic decoder curve (a.d.c.) for a given small code $f : \mathbb{F}_2^u \to \mathbb{F}_2^v$ if there exists a sequence of integers, L_j, of decoders $g_j : \mathbb{F}_2^{vL_j} \to \mathbb{F}_2^{uL_j}$ such that

$$D\left(f^{\oplus L_j}, g_j, \frac{t}{v}\right) \to \frac{1}{u} r(t) \tag{9}$$

for all $t \in [0, v]$ points of continuity of r. An a.d.c. r is called minimal if for any other a.d.c. r' there is an $s \in [0, v]$ such that $r(s) < r'(s)$.

Note that the LHS of (9) is a sequence of non-decreasing, right-continuous functions. Thus by Helly's theorem [5, Chapter 7] given any sequence of decoders $g_j : \mathbb{F}_2^{L_j v} \to \mathbb{F}_2^{L_j u}$ there always exists at least one limiting a.d.c. The set of all a.d.c.'s describe a totality of performance curves achievable (for large

L) by decoders oblivious to the actual value of adversarial noise δ. Recall that (Theorem 2) the optimal performance for the decoder that can adapt to δ is given by $D(f^{\oplus\infty}, \delta)$. It turns out (unsurprisingly) that $D(f^{\oplus\infty}, \delta)$ is just a lower bound of all of the a.d.c.'s:

Proposition 3: For every $t \in [0, v]$ we have

$$D\left(f^{\oplus\infty}, \frac{t}{v}\right) = \frac{1}{u} \min r(t-)\,, \tag{10}$$

where minimum is over the set of all a.d.c.'s.

Proof: For convenience, denote

$$r^*(t) \triangleq u \cdot D\left(f^{\oplus\infty}, \frac{t}{v}\right)\,.$$

Consider arbitrary a.d.c. r and a sequence

$$r_j(t) \triangleq D\left(f^{\oplus L_j}, g_j, \frac{t}{v}\right) \to \frac{1}{u} r(t)\,.$$

Then, by the general properties of convergence of distributions we have (for each t):

$$r(t-) \le \liminf_{j\to\infty} r_j(t-) \le \limsup_{j\to\infty} r_j(t) \le r(t)\,.$$

But by (5) we have

$$\liminf_{j\to\infty} r_j(t) \ge r^*(t)\,.$$

and therefore

$$r(t) \ge r^*(t) \qquad \forall t \in [0, v] \tag{11}$$

Since r^* is continuous in t (Theorem 2) we can strengthen (11) to

$$r(t-) \ge r^*(t) \tag{12}$$

and therefore

$$r^*(t) \le \inf_{r-\text{ a.d.c.}} r(t-) \tag{13}$$

Next, consider a sequence of decoders g_L, $L \to \infty$ which attain $r^*(t_0)$ for some fixed t_0. Denote

$$r_L(t) \triangleq u \cdot D\left(f^{\oplus L}, g_L, \frac{t}{v}\right)$$

then we have

$$r_L(t_0) \to r^*(t_0)\,. \tag{14}$$

By Helly's theorem there exists a subsequence L_j and some non-decreasing right-continuous function $r : [0, v] \to [0, u]$ such that $r_{L_j}(t) \to r(t)$ for every point of continuity of t. Thus r is an a.d.c. with g_{L_j} as a limiting sequence of decoders. Again by convergence of distributions we have

$$r(t_0-) \le \liminf_{j\to\infty} r_{L_j}(t_0-)$$

Then from $r_{L_j}(t_0-) \le r_{L_j}(t_0)$, (12) and (14) we obtain

$$r^*(t) \le r(t_0-) \le r^*(t)$$

implying that $r(t_0-) = r^*(t)$ and thus the bound in (13) is tight. ■

Examples of a.d.c.'s can be obtained via the following result:

Proposition 4: Given $k \ge 1$ decoders $g_1, \ldots, g_k : \mathbb{F}_2^v \to \mathbb{F}_2^u$, their envelopes $r^{**}_{g_1}, \ldots, r^{**}_{g_k}$ and positive weights λ_j such that $\sum_{j=1}^k \lambda_j = 1$, the following is a continuous concave a.d.c.:

$$r(t) = \max \sum_{j=1}^k \lambda_j r^{**}_{g_j}(\tau_j)\,, \tag{15}$$

where maximum is over all $\tau_j \in [0, v]$ such that $\sum_{j=1}^k \lambda_j \tau_j \le t$.

Proof: The idea is to use each decoder g_j for λ_j-portion of blocks. Let us denote such a decoder by

$$g_L \triangleq \bigoplus_{j=1}^k g_j^{\oplus \lambda_j L}\,.$$

The statement of the Proposition is then equivalent to: *The function of $t \in [0, v]$ given by (15) is continuous and concave; furthermore the following holds for all $t \in [0, v]$:*

$$\lim_{L\to\infty} D\left(f^{\oplus L}, g_L, \frac{t}{v}\right) = \frac{1}{u} r(t) \tag{16}$$

Consider any $\theta t_1 + (1-\theta) t_2 = t$ for $\theta \in [0, 1]$. Let $\{\tau_j^{(1)}\}_{j=1}^k$ and $\{\tau_j^{(2)}\}_{j=1}^k$ be the coefficients achieving $r(t_1)$ and $r(t_2)$ in (15) respectively. Then by taking $\tau_j = \theta \tau_j^{(1)} + (1-\theta)\tau_j^{(2)}$ and using the concavity of r_g^{**}, we obtain the concavity of $r(t)$. Concavity then implies continuity immediately.

Next, we show

$$\limsup_{L\to\infty} D\left(f^{\oplus L}, g_L, \frac{t}{v}\right) \le \frac{1}{u} r(t)\,, \tag{17}$$

Suppose the adversary flips $\tau_j \lambda_j L$ bits in the j-th block. By (4), the decoder commits at most $r_g^{**}(\tau_j)\lambda_j L$ bits of error in the j-th block. In total, the number bits of error is $\sum_{j=1}^k r_g^{**}(\tau_j)\lambda_j L$, with number of flipped bits by the adversary $\sum_{j=1}^k \tau_j \lambda_j L \le (t/v) v L = tL$. By optimizing τ_j, we obtain (17).

The proof concludes by demonstrating

$$\liminf_{L\to\infty} D\left(f^{\oplus L}, g_L, \frac{t}{v}\right) \ge \frac{1}{u} r(t) \tag{18}$$

Let $\{\tau_j\}_{j=1}^k$ be those coefficients achieving (15), then for each block j, there exists a source realization and adversary noise vector e_j with $|e_j| \le \tau_j \lambda_j L$ such that the decoder commits at least $r_g^{**}(\tau_j)\lambda_j L$ bits of errors by (4). Take the summation over the k blocks, there exists a source realization and adversary noise vector $e = e_1 || \ldots || e_k$ where $|e| \le tL$ such that the decoder commits at least $\sum_{j=1}^k r_g^{**}(\tau_j)\lambda_j L$ bits of error. So (18) holds.

■

B. Converse bounds on a.d.c.'s

Our goal now is to develop a tool for demonstrating that an a.d.c. cannot be very small for all t. Our result is a certain global (i.e. over a range of t's) condition on $r(t)$, as opposed to pointwise lower bound of Proposition 3. We start with some preliminary definitions and remarks.

Definition 3: Function $x \mapsto \ell(x)$ is called a feasible distance profile (FDP) if $\exists x_0$ s.t. $\ell(x) \ge |x - x_0|$ for all x.

The next proposition is our main tool to derive global constraints on a.d.c.'s. It's meaning is that functions r_g corresponding to arbitrary decoder (see (2)) have rather special structure, intertwined with the geometry of the Hamming space:

Proposition 5: For any JSCC $f:\mathbb{F}_2^k\to\mathbb{F}_2^n$ and $g:\mathbb{F}_2^n\to\mathbb{F}_2^k$, and for any $y\in\mathbb{F}_2^n$ the map

$$x\mapsto r_g(|f(x)-y|)$$

– is an FDP.

Proof: Just take $x_0=g(y)$ in the definition of the FDP. ■

Definition 4: For each $x_0\in\mathbb{F}_2^u$, define:

$$\rho_{y,x_0}(s)=\max_{x:|f(x)-y|\le s}|x-x_0|\,.$$

Value of ρ is taken to be $-\infty$ if the constraint set is empty.

Proposition 6: For any JSCC $f:\mathbb{F}_2^k\to\mathbb{F}_2^n$ and $g:\mathbb{F}_2^n\to\mathbb{F}_2^k$, and for any $y\in\mathbb{F}_2^n$ there exists $x_0\in\mathbb{F}_2^k$ such that

$$\forall s\in[0,v]:\quad r_g(s)\ge\rho_{y,x_0}(s)\,,$$

where

$$\rho_{y,x_0}(s)\triangleq\max_{x:|f(x)-y|\le s}|x-x_0| \tag{19}$$

($\rho=-\infty$ when the constraint set is empty).

Proof: Since $r_g(|f(x)-y|)$ is an FDP, by definition there exists x_0 such that

$$r_g(|f(x)-y|)\ge|x-x_0|\qquad\forall x\in\mathbb{F}_2^u$$

Taking max over all $x\in f^{-1}B_n(y,s)$ we obtain the result. ■

Finally, we are ready to prove our main converse bound for the a.d.c.'s:

Theorem 7: Fix code $f:\mathbb{F}_2^u\to\mathbb{F}_2^v$. Then every asymptotic decoder curve of f satisfies the following: for every $y\in\mathbb{F}_2^v$, there exists a probability distribution Λ on $\mathbb{F}_2^u$ such that for all

$$\forall s\in[0,u]:\quad r(s)\ge\rho_{y,\Lambda}(s)\,,$$

where

$$\rho_{y,\Lambda}(s)\triangleq\max_{s_x:\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x\le s}\sum_{x\in\mathbb{F}_2^u}\Lambda(x)\rho_{y,x}(s_x)$$

and $\rho_{y,x}(\cdot)$ is defined in (19).

Proof: For every y, it suffices to prove that for each L_j and associated decoder g_j in (9), there exists a distribution Λ_j on $\mathbb{F}_2^u$ such that:

$$uD(f^{\oplus L_j},g_j,\frac{s}{v})\ge\rho_{y,\Lambda_j}(s) \tag{20}$$

for every $s\in[0,1]$. Then by the compactness of the set of all distributions on $\mathbb{F}_2^u$, there exists a subsequence $\{L_{n_i}\}$ of $\{L_j\}$ such that $\lim_{i\to\infty}\Lambda_{n_i}=\Lambda$ exists, hence $\lim_{i\to\infty}\rho_{y,\Lambda_{n_i}}(s)$ also exists. Then by replacing j by n_i and let i goes to infinity in (20), we obtain $r(s)\ge\rho_{y,\Lambda}(s)$.

Now for fixed block length L_j, expand the LHS of (20) as:

$$\frac{1}{L_j}\max_{x\in\mathbb{F}_2^{uL_j}}\max_{e:|e|\le sL_j}|x+g_j(f^{\oplus L_j}(x)+e)| \tag{21}$$

Restrict $x\in(f^{\oplus L_j})^{-1}B(y^{L_j},sL_j)$, (21) is lower bounded by:

$$\frac{1}{L_j}\max_{x:|f^{\oplus L_j}(x)-y^{L_j}|\le sL_j}|x+g_j(y^{L_j})| \tag{22}$$

Assume the decoder g_j decodes y^{L_j} to $(\hat{x}_1,\ldots,\hat{x}_{L_j})$. Then (22) can be further expressed as:

$$\frac{1}{L_j}\max_{x_i\in\mathbb{F}_2^u:\sum_{i=1}^{L_j}|f(x_i)-y|\le sL_j}\sum_{i=1}^{L_j}|\hat{x}_i-x_i| \tag{23}$$

Now take

$$\Lambda(\hat{x})=\frac{\text{number of appearance of }\hat{x}\text{ in }g_j(y^{L_j})}{L_j}$$

By the definition of $\rho_{y,x}(\cdot)$ in (19), (23) can be expressed as:

$$\max_{\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x\le s}\sum_{\hat{x}\in\mathbb{F}_2^u}\Lambda(\hat{x})\rho_{y,\hat{x}}(s_x)$$

which is just the definition of $\rho_{y,\Lambda}(s)$. ■

C. Alternative interpretation of Theorem 7

For any $y\in\mathbb{F}_2^v$ and any function $h:\mathbb{F}_2^u\to\mathbb{F}_2^u$, we define a set $S_{y,h}$ as

$$S_{y,h}=\{(s,t)|s=|f(h(x))-y|,t=|x-h(x)|,\forall x\in\mathbb{F}_2^u\}$$

Proposition 8: If an asymptotic decoder curve $r(\cdot)$ passes though the convex closure of $S_{y,h}$ for any y and h, there exists a distribution Λ on $\mathbb{F}_2^u$ such that for all $s\in[0,1]$

$$r(s)\ge\rho_{y,\Lambda}(s)$$

Conversely, if there exists y and distribution Λ such that $r(s)\ge\rho_{y,\Lambda}(s)$ for all s, then r passes through $S_{y,h}$ for some h.

Proof: If there exists a distribution Λ on $\mathbb{F}_2^u$ such that $r(s)\ge\rho_{y,\Lambda}(s)$ for all s and s_x's satisfying $s=\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x$. we have:

$$r(\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x)\ge\rho_{y,\Lambda}(\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x)$$
$$\ge\sum_{x\in\mathbb{F}_2^u}\Lambda(x)\rho_x(s_x)=\sum_{x\in\mathbb{F}_2^u}\Lambda(x)\max_{x_0:|f(x_0)-y|\le s_x}|x-x_0|$$

Take $s_x=|f(h(x))-y|$ and $x_0=h(x)$, we obtain

$$r(\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x)\ge\sum_{x\in\mathbb{F}_2^u}\Lambda(x)|x-h(x)|$$

Then the node

$$(\sum_{x\in\mathbb{F}_2^u}\Lambda(x)|f(h(x))-y|,\sum_{x\in\mathbb{F}_2^u}\Lambda(x)|x-h(x)|)$$

is inside the region $S_{y,h}$. So r must pass through $S_{y,h}$.

Conversely, given y, if for any distribution Λ on $\mathbb{F}_2^u$, there exists s such that $r(s)<\rho_{y,\Lambda}(s)$, that means there exists a set of integers s_x such that

$$r(\sum_{x\in\mathbb{F}_2^u}\Lambda(x)s_x)<\sum_{x\in\mathbb{F}_2^u}\Lambda(x)\rho_{y,x}(s_x)$$

TABLE I

INPUT-OUTPUT BER CURVES r_0, r_{g_1}, r_{g_2} AND THEIR ENVELOPES

s	0	1	2	3	4	5	6	7
$r_0(s)$	0	0	2	3	3	3	4	4
$r_0^{**}(s)$	0	1	2	3	$3\frac{1}{3}$	$3\frac{2}{3}$	4	4
$r_{g_1}(s)$	0	0	3	3	3	3	4	4
$r_{g_1}^{**}(s)$	0	1.5	3	$3\frac{1}{4}$	$3\frac{1}{2}$	$3\frac{3}{4}$	4	4
$r_{g_2}(s)$	0	1	2	3	3	4	4	4
$r_{g_2}^{**}(s)$	0	1	2	3	$3\frac{1}{2}$	4	4	4

Since there exists some function h such that $|f(h(x))-y| = s_x$ for any x. Then by 19, we have:

$$r(\sum_{x\in\mathbb{F}_2^u} \Lambda(x)|f(h(x)-y)|) < \sum_{x\in\mathbb{F}_2^u} \Lambda(x)|x-h(x)|$$

So r do not pass through $S_{y,h}$. ■

IV. EXAMPLE: HAMMING $[7,4,3]$ CODE

In conclusion, we particularize our results to the usual Hamming $[7,4,3]$ code. Note that up until now, the only codes which we considered were the $[2m+1,1,2m+1]$ repetition codes, cf. [1]. There, computation of $D(f^{\oplus\infty},\delta)$ was done by finding a decoder with $r_g^{**} = r_0^{**}$ and Theorem 1. In this section we show:

1) For Hamming $[7,4,3]$ there does not exist decoder with $r_g^{**} = r_0^{**}$.
2) Evaluation of $D(f^{\oplus\infty},\delta)$ is nevertheless possible via Theorem 2.
3) Results of Section III show that there does not exist a decoder that is simultaneously optimal for all δ (i.e. the minimum in (10) is attained by different a.d.c.'s depending on the adversary noise).

A. *Two decoders for Hamming* $[7,4,3]$

For $[7,4,3]$- Hamming code $f(\vec{x}) = \vec{x}G$ where

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \tag{24}$$

The quantity r_0 in (1) and its envelope are given in Table I.

Consider two decoders:

- The minimum distance decoder g_1: firstly compute the parity $b = \vec{y}H$ where

$$H^T = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

  And if $b \neq 000$, g_1 corrects the error on the i-th bit where the i-th row of H is just b.
- Alternative decoder g_2: upon receiving the input $\vec{y}$, take $\vec{x}$ as the first four bits of $\vec{y}$. Then compute $|\vec{y}-\vec{x}G|_H$. If the Hamming distance is 3, then it flips the last bit of $\vec{x}$ and output it. Otherwise, directly output $\vec{x}$. Then if the first four bits of the codeword are all flipped, the decoder will detect and correct the error, so $r_{g_2}(4) = 3$. While if more than 4 bits are modified, g_2 cannot detect the error.

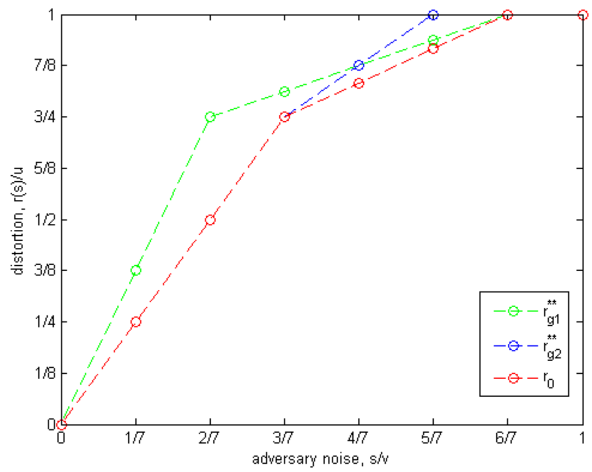

Fig. 1. Comparison of r_0^{**}, $r_{g_1}^{**}$, $r_{g_2}^{**}$. Note that according to (25) it is asymptotically optimal to use g_2 for $\delta \le \frac{4}{7}$ and g_1 for $\delta \ge \frac{4}{7}$. Consequently, the bound r_0^{**} (Theorem 1) is not tight for $\delta \in (\frac{3}{7}, \frac{6}{7})$.

The quantity r_g in (2) for decoder g_1 and g_2, as well as their envelopes, are given in Table I.

We notice that there exist some s such that $r_g^{**}(s) > r_0^{**}(s)$ for both decoders g_1 and g_2. Actually it holds for every deterministic decoder.

Proposition 9: For a $[7,4,3]$ Hamming code (24), there is no deterministic decoder $g : \mathbb{F}_2^7 \to \mathbb{F}_2^4$ achieving $r_g^{**}(s) = r_0^{**}(s)$ for all $s \in [0,7]$ simultaneously. Furthermore, asymptotic performance of the best decoder (with knowledge of δ) is given by

$$D(f^{\oplus\infty},\delta) = \frac{1}{4}\min\left(r_{g_1}^{**}(7\delta), r_{g_2}^{**}(7\delta)\right) \tag{25}$$

Proof: If we want $r_g^{**}(s) = r_0^{**}(s)$ for each possible s, then for each $0 \le s \le 7$

$$\max_{|e|\le s}\max_{x\in\mathbb{F}_2^4} |g(f(x)+e)-x| = r_g(s) \le r_g^{**}(s) = r_0^{**}(s)$$

$$\iff \forall|e| \le s, \forall x \in \mathbb{F}_2^4, |g(f(x)+e)-x| \le r_0^{**}(s)$$

Let $y = f(x)+e$, this is equivalent to

$$\forall y \in \mathbb{F}_2^7, \forall x \in \mathbb{F}_2^4, |g(y)-x| \le r_0^{**}(|f(x)-y|)$$

We notice that for $y = 0000011$ (also some other strings, we just take this for example), it is impossible to find such a $g(y)$ to satisfy this condition for all x.

Indeed, by inspecting Table II we notice that no matter what $g(y)$ is, there exists an x such that $|g(y)-x| = 4$. Notice that there is only $x = 1101$ which allows $|g(y)-x| = 4$. So $g(y)$ could only be 0010. But then $|g(y)-1100| = 3 > 2$. Therefore, we can not find assignment $g(y)$ to satisfy all the conditions. So no decoder g can achieve $r_g^{**}(s) = r_0^{**}(s)$ for all s.

Finally, (25) is just a numerical evaluation of Theorem 2. ■

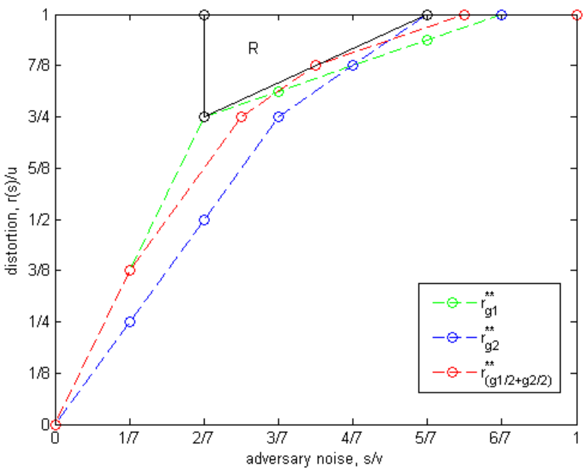

Fig. 2. Comparison of three a.d.c.s: $r^{**}_{g_1}$, $r^{**}_{g_2}$ and the decoder that uses g_1 for 50% of blocks and g_2 for the rest. All of them pass through region R of Proposition 10.

B. Global constraint on a.d.c.'s of the Hamming code

Proposition 10: Any asymptotic decoder curve r for $[7,4,3]$ passes through region R, where R is the convex closure of

$$\{(2,3),(2,4),(5,4)\}$$

Remark: Note that performance of the optimal decoder (with knowledge of δ) does not pass through R, see (25). Thus, the $D(f^{\oplus\infty},\delta)$ is not an a.d.c. and hence no decoder (oblivious to δ) can attain simultaneously all of its points.

Proof: Look at $y = 0000011$, we compute all the ρ_{y,x_0} curves for $x_0 \in \mathbb{F}_2^u$. It turns out that only $\rho_{y,0000}$ and $\rho_{y,0010}$ are minimal curves. Namely for any $x \notin \{0000, 0010\}$, there exists $x_0 \in \{0000, 0010\}$ such that

$$\rho_{y,x}(s) \geq \rho_{y,x_0}(s)$$

for all s. Consequently, for every Λ on $\mathbb{F}_2^u$ there exists Λ' supported on $\{0000, 0010\}$ such that

$$\rho_{y,\Lambda}(s) \geq \rho_{y,\Lambda'}(s) \qquad \forall s\,.$$

By Theorem 7 each a.d.c. is lower bounded by an infimal convolution of the two "minimal" curves $\rho_{y,0000}$ and $\rho_{y,0010}$ shown in Table III.

For any distribution Λ on $\{0000, 0010\}$, consider $s = 5\Lambda(0000) + 2\Lambda(0010)$, we have:

$$\begin{aligned} r(s) &\geq \rho_{y,\Lambda} \\ &\geq \Lambda(0000)\rho_{y,0000}(5) + \Lambda(0010)\rho_{y,0010}(2) \\ &= 4\Lambda(0000) + 3\Lambda(0010) \end{aligned}$$

Since $s \in [2,5]$, this curve should pass trough the region R no matter which distribution Λ is chosen. ■

TABLE II
ILLUSTRATION FOR PROPOSITION 9

x	$f(x)$	$\lvert f(x)-y\rvert$	$r_0^{**}(\lvert f(x)-y\rvert)$
0000	0000000	2	2
0001	0001111	2	2
0010	0010011	1	1
0011	0011100	5	3
0100	0100101	3	3
0101	0101010	3	3
0110	0110110	4	3
0111	0111001	4	3
1000	1000110	3	3
1001	1001001	3	3
1010	1010101	4	3
1011	1011010	4	3
1100	1100011	2	2
1101	1101100	6	4
1110	1110000	5	3
1111	1111111	5	3

TABLE III
TWO MINIMAL CURVES $\rho_{y,0000}$ AND $\rho_{y,0010}$,

s	0	1	2	3	4	5	6	7
$\rho_{y,0000}(s)$	$-\infty$	1	2	2	3	4	4	4
$\rho_{y,0010}(s)$	$-\infty$	0	3	3	3	3	4	4

REFERENCES

[1] Y. Kochman, A. Mazumdar, and Y. Polyanskiy, "The adversarial joint source-channel problem," in *Proc. 2012 IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, Jul. 2012.

[2] ——, "Results on combinatorial joint source-channel coding," in *Proc. 2012 Inf. Theory Workshop (ITW)*, Lausanne, Switzerland, Sep. 2012.

[3] A. Mazumdar, Y. Polyanskiy, and B. Saha, "On Chebyshev radius of a set in Hamming space and the closest string problem," in *Proc. 2013 IEEE Int. Symp. Inf. Theory (ISIT)*, Istanbul, Turkey, Jul. 2013.

[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[5] B. V. Gnedenko, *The Theory of Probability And the Elements of Statistics*. AMS Bookstore, 2005, vol. 132.