The Error Probability of Generalized Perfect Codes

Gonzalo Vazquez-Vilar
Universidad Carlos III de Madrid
Gregorio Marañón Health Research Institute
gvazquez@ieee.org

Albert Guillén i Fàbregas
ICREA & Universitat Pompeu Fabra
University of Cambridge
guillen@ieee.org

Sergio Verdú
Princeton University
verdu@princeton.edu

***Abstract*—We introduce a definition of perfect and quasi-perfect codes for symmetric channels parametrized by an auxiliary output distribution. This new definition generalizes previous definitions and encompasses maximum distance separable codes. The error probability of these codes, whenever they exist, is shown to attain the meta-converse lower bound.**

I. Introduction

Polyanskiy *et al.* proved that the error probability of a certain auxiliary binary hypothesis test can be used to lower bound the error probability of channel coding in the finite blocklength regime [1, Th. 27]. In particular, [1, Th. 27] establishes the following lower bound on the error probability of a code $\mathcal{C}$ with cardinality M used over a channel $P_{Y|X}$,

$$P_{\mathrm{e}}(\mathcal{C}) \geq \inf_{P_X} \sup_{Q} \left\{ \alpha_{\frac{1}{M}} \bigl(P_X \times P_{Y|X}, P_X \times Q\bigr) \right\}, \qquad (1)$$

where $\alpha_\beta\bigl(P_0, P_1\bigr)$ denotes the Neyman-Pearson performance of a binary hypothesis test discriminating between distributions P_0 and P_1 (see definition in (2)). This bound is usually referred to as the meta-converse bound, since several previous converse bounds in the literature can be derived from it via relaxation.

Particularized for n-uses of a binary symmetric channel (BSC), the meta-converse bound recovers the sphere-packing bound for BSC channels [2, Eq. (5.8.19)] (see [1, Sec. III.H]). This bound is known to be tight for perfect or quasi-perfect binary codes. A binary code is said to be *perfect* if non-overlapping Hamming spheres of radius t centered on the codewords exactly fill out the space. More generally, a *quasi-perfect* code is defined as a code in which Hamming spheres of radius t centered on the codewords do not overlap, while Hamming spheres of radius $t{+}1$ cover the space, possibly with overlaps. Since quasi-perfect codes attain the lower bound (1), they achieve the minimum error probability in a BSC among all the codes with the same blocklength and rate [2, Sec. 5.8].

In [3], Hamada studied a generalization of perfect and quasi-perfect codes beyond Hamming distance. Using a variation of the Fano metric, Hamada derived a lower bound to the channel coding error probability for a class of symmetric channels. This bound is achievable by perfect and quasi-perfect codes (defined with respect to the new metric), whenever they exist.

In this work, we generalize the definition of perfect and quasi-perfect codes via an auxiliary output distribution. We show that, for symmetric channels, these codes attain equality in (1) achieving the minimum error probability among all the codes with the same blocklength and rate. Our definition is more general than Hamada's [3], e.g., it subsumes maximum distance separable (MDS) codes which are generalized quasi-perfect with respect to the erasure channel.

II. Binary Hypothesis Testing

Consider a binary hypothesis test discriminating between distributions P_0 and P_1 defined over some discrete alphabet $\mathcal{Z}$. Let $T(z) \in [0,1]$ denote the probability of the test deciding P_0 given an observation z, $0 \leq T(z) \leq 1$. The minimum error probability under P_0 among all tests T with error probability upper bounded by β under P_1 is

$$\alpha_\beta\bigl(P_0, P_1\bigr) \triangleq \inf_{T: \sum_z T(z) P_1(z) \leq \beta} \left\{ 1 - \sum_z T(z) P_0(z) \right\}. \qquad (2)$$

Neyman and Pearson provided in [4] an explicit form for the test achieving this trade-off, which yields the following alternative characterization for $\alpha_\beta\bigl(P_0, P_1\bigr)$.

Lemma 1: For a binary hypothesis test between P_0 and P_1,

$$\alpha_\beta\bigl(P_0, P_1\bigr) = \max_{\gamma \geq 0} \left\{ \mathbb{P}_0\left[\tfrac{P_0(Z)}{P_1(Z)} \leq \gamma\right] + \gamma \mathbb{P}_1\left[\tfrac{P_0(Z)}{P_1(Z)} > \gamma\right] - \gamma\beta \right\}. \qquad (3)$$

where $\mathbb{P}_i[\cdot]$ is computed with respect to $Z \sim P_i$, $i = 0, 1$.

Proof: Proof is omitted due to space constraints. ■

III. Generalized Perfect Codes

An equiprobable message $v \in \{1, \ldots, M\}$ is to be transmitted over a channel with transition probability $P_{Y|X}$, $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ with $\mathcal{X}$ and $\mathcal{Y}$ being the one-shot input/output discrete alphabets. A channel code is the set of codewords $\mathcal{C} = \{x_1, \ldots, x_M\}$ assigned to each of the messages. Under maximum likelihood (ML) decoding, the error probability for the code $\mathcal{C}$ is given by

$$P_{\mathrm{e}}(\mathcal{C}) = 1 - \frac{1}{M} \sum_y \max_{x \in \mathcal{C}} P_{Y|X}(y|x). \qquad (4)$$

For any $\tau \geq 0$ and any distribution Q defined over $\mathcal{Y}$, we define $\mathcal{S}_x(\tau, Q)$ to be the set of outputs y with likelihood given input x at least $\tau Q(y)$, i.e.,

$$\mathcal{S}_x(\tau, Q) \triangleq \left\{ y \in \mathcal{Y} \;\middle|\; \tfrac{P_{Y|X}(y|x)}{Q(y)} \geq \tau \right\}. \qquad (5)$$

This work has been funded in part by the European Research Council (ERC) under grants 714161 and 725411, by the Spanish Ministry of Economy and Competitiveness under Grants TEC2016-78434-C3 and IJCI-2015-27020, by the National Science Foundation under Grant CCF-1513915 and by the Center for Science of Information, an NSF Science and Technology Center under Grant CCF-0939370.

We define the interior and the outer shell of $\mathcal{S}_x(\tau,Q)$ as

$$\mathcal{S}_{\mathrm{i},x}(\tau,Q) \triangleq \left\{ y\in\mathcal{Y} \,\middle|\, \tfrac{P_{Y|X}(y|x)}{Q(y)} > \tau \right\}, \qquad (6)$$

$$\mathcal{S}_{\mathrm{o},x}(\tau,Q) \triangleq \left\{ y\in\mathcal{Y} \,\middle|\, \tfrac{P_{Y|X}(y|x)}{Q(y)} = \tau \right\}. \qquad (7)$$

We refer to $\mathcal{S}_x(\tau,Q)$ as a sphere of radius τ centered on x, although in general $\mathcal{X}\neq\mathcal{Y}$ and $\frac{P_{Y|X}(y|x)}{Q(y)}$ is not a distance measure. This metric is equivalent to Fano metric [5, Eq. (9.10)]. For channels such as the BSC, $\log P_{Y|X}(y|x)$ is an affine function of the Hamming distance between x and y and, hence, $\mathcal{S}_x(\tau,Q)$ becomes a sphere with respect to that distance when Q is the equiprobable distribution.

Definition 1: Let $F_x(\tau) \triangleq \mathbb{P}\bigl[P_{Y|X}(Y|x)\geq\tau\bigr]$, where the $Y\sim P_{Y|X=x}$ and $\tau\in[0,1]$. A channel $P_{Y|X}$ is *symmetric* if $F_x(\tau)$ does not depend on the input x,

$$F_x(\tau) = F(\tau), \quad \forall x\in\mathcal{X}, \quad \tau\in[0,1]. \qquad (8)$$

This definition implies that the rows of the channel transition matrix (with inputs as rows and outputs as columns), $P_{Y|X}(\cdot|x)$, are permutations of each other.

We consider the set of output distributions Q such that the tilted channel $\tilde{P}_{Y|X}(y|x) \propto \frac{P_{Y|X}(y|x)}{Q(y)}$ remains symmetric. More precisely, we define

$$\mathcal{Q} \triangleq \Bigl\{ Q \;\bigm|\; F_x(\tau,Q) = F(\tau,Q),\ \forall x\in\mathcal{X},\ \tau\in[0,1] \Bigr\}, \qquad (9)$$

where $F_x(\tau,Q) \triangleq \mathbb{P}\bigl[Y\in\mathcal{S}_x(\tau,Q)\bigr]$ with $Y\sim P_{Y|X=x}$.

For symmetric channels $P_{Y|X}$, the set $\mathcal{Q}$ is non-empty as it always includes the equiprobable distribution. For example, consider a single use of the binary erasure channel (BEC) with erasure symbol e. In this case, any distribution with $Q(0)=Q(1)=\xi$, $Q(\mathrm{e})=1-2\xi$, does not alter the symmetry of the original channel, and therefore it is included in $\mathcal{Q}$.

Definition 2: A code $\mathcal{C}$ is *generalized perfect* for $P_{Y|X}$, if there exists $\gamma\geq 0$ and $Q\in\mathcal{Q}$ such that the codeword-centered sets $\bigl\{\mathcal{S}_x(\gamma,Q)\bigr\}_{x\in\mathcal{C}}$ are disjoint and

$$\bigcup_{x\in\mathcal{C}} \mathcal{S}_x(\gamma,Q) = \mathcal{Y}. \qquad (10)$$

A code is *generalized quasi-perfect* if there exists $\gamma\geq 0$ and $Q\in\mathcal{Q}$ such that (10) is satisfied and the codeword-centered sets $\bigl\{\mathcal{S}_{\mathrm{i},x}(\gamma,Q)\bigr\}_{x\in\mathcal{C}}$ are disjoint.

Theorem 1: Let $P_{Y|X}$ be a symmetric channel and $\mathcal{C}$ be generalized quasi-perfect code with respect to $P_{Y|X}$. Then, $\mathcal{C}$ attains the minimum error probability among all codes with M codewords, and it is given by

$$P_{\mathrm{e}}(\mathcal{C}) = \inf_{P_X}\max_{Q\in\mathcal{Q}} \Bigl\{ \alpha_{\frac{1}{M}}\bigl(P_X\times P_{Y|X}, P_X\times Q\bigr) \Bigr\} \qquad (11)$$

$$= \max_{Q\in\mathcal{Q}} \Bigl\{ \alpha_{\frac{1}{M}}\bigl(P_{Y|X=x}, Q\bigr) \Bigr\}, \text{ for all } x\in\mathcal{X}. \qquad (12)$$

Conversely, any code for which (11)-(12) hold is generalized quasi-perfect.

The optimizing Q in Theorem 1 coincides with that in Definition 2. Setting Q to be the equiprobable distribution, we recover [3, Th. 3] in the same generality. Allowing different $Q\in\mathcal{Q}$ yields a larger family of quasi-perfect codes which encompasses MDS codes for erasure channels, as shown next.

A. Proof of Theorem 1

We provide two auxiliary lemmas, whose proof is given in the appendix, that will be used in the proof of Theorem 1.

For discrete alphabets $\mathcal{X},\mathcal{Y}$ we define the countable set

$$\mathcal{L}_Q \triangleq \Bigl\{ \tau\in\mathbb{R} \,\Bigm|\, \exists x\in\mathcal{X}, \exists y\in\mathcal{Y}, \tfrac{P_{Y|X}(y|x)}{Q(y)} = \tau \Bigr\}. \qquad (13)$$

Lemma 2: Let $P_{Y|X}$ be a symmetric channel according to Definition 1 and $Q\in\mathcal{Q}$ defined in (9). Then, the probability measure of the sets $\mathcal{S}_x(\tau,Q)$, $\mathcal{S}_{\mathrm{i},x}(\tau,Q)$ and $\mathcal{S}_{\mathrm{o},x}(\tau,Q)$ with respect to Q is independent of $x\in\mathcal{X}$ for any $\tau\geq 0$.

Then, for symmetric channels, we define the probability measures $\mathsf{Q}_{\mathrm{i}}(\tau) \triangleq \mathbb{Q}\bigl[\mathcal{S}_{\mathrm{i},x}(\tau,Q)\bigr]$ and $\mathsf{Q}_{\mathrm{o}}(\tau) \triangleq \mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau,Q)\bigr]$. The next result is a refinement of [5, Eqs. (9.15)-(9.16)].

Lemma 3: Let $P_{Y|X}$ be a symmetric channel. Then, $\mathcal{C}$ is generalized quasi-perfect if and only if

$$P_{\mathrm{e}}(\mathcal{C}) = \sum_{\tau\in\mathcal{L}_Q,\,\tau\leq\gamma} \tau\mathsf{Q}_{\mathrm{o}}(\tau) - \gamma\Bigl(\tfrac{1}{M} - \mathsf{Q}_{\mathrm{i}}(\gamma)\Bigr), \qquad (14)$$

for some $\gamma\geq 0$ and $Q\in\mathcal{Q}$.

Let us consider the hypothesis test in (1). We apply Lemma 1 with $P_0 \leftarrow P_X\times P_{Y|X}$ and $P_1\leftarrow P_X\times Q$. Using the definition of the set $\mathcal{S}_{\mathrm{i},x}(\cdot)$ and $\mathsf{Q}_{\mathrm{i}}(\cdot)$ in Lemma 2 yields

$$\alpha_{\frac{1}{M}}\bigl(P_X\times P_{Y|X}, P_X\times Q\bigr)$$

$$= \max_{\gamma\geq 0}\Biggl\{ \sum_x \sum_{y\notin\mathcal{S}_{\mathrm{i},x}(\gamma,Q)} P_X(x)P_{Y|X}(y|x) + \gamma\mathsf{Q}_{\mathrm{i}}(\gamma) - \tfrac{\gamma}{M} \Biggr\}. \qquad (15)$$

For any $y\in\mathcal{S}_{\mathrm{o},x}(\tau,Q), \tau\in\mathcal{L}_Q$, it holds that $\frac{P_{Y|X}(y|x)}{Q(y)} = \tau$. Then,

$$\sum_{y\notin\mathcal{S}_{\mathrm{i},x}(\gamma,Q)} P_{Y|X}(y|x) = \sum_{\substack{\tau\in\mathcal{L}_Q,\,\tau\leq\gamma,\\ y\in\mathcal{S}_{\mathrm{o},x}(\tau,Q)}} \frac{P_{Y|X}(y|x)}{Q(y)}Q(y) \qquad (16)$$

$$= \sum_{\tau\in\mathcal{L}_Q,\,\tau\leq\gamma}\ \sum_{y\in\mathcal{S}_{\mathrm{o},x}(\tau,Q)} \tau Q(y) \qquad (17)$$

$$= \sum_{\tau\in\mathcal{L}_Q,\,\tau\leq\gamma} \tau\mathsf{Q}_{\mathrm{o}}(\tau), \qquad (18)$$

which does not depend on x (see Lemma 2).

Then, (15) becomes

$$\alpha_{\frac{1}{M}}\bigl(P_X\times P_{Y|X}, P_X\times Q\bigr)$$

$$= \max_{\gamma\geq 0}\Biggl\{ \sum_{\tau\in\mathcal{L}_Q,\,\tau\leq\gamma} \tau\mathsf{Q}_{\mathrm{o}}(\tau) + \gamma\mathsf{Q}_{\mathrm{i}}(\gamma) - \tfrac{\gamma}{M} \Biggr\}. \qquad (19)$$

According to (1), the right-hand side of (19) is a lower bound to $P_{\mathrm{e}}(\mathcal{C})$. According to Lemma 3, the term in braces in (19) is precisely the error probability of a generalized quasi-perfect code. Then, whenever this code exists the lower bound (19) is achievable and (11) holds with equality. Moreover, since this bound does not depend on P_X for symmetric channels and $Q\in\mathcal{Q}$, then (12) follows.

Let now $Q\in\mathcal{Q}$ achieve (12), and let fix $\gamma\geq 0$ to be the maximizer in (19). We conclude from Lemma 3 that the term in braces in (19) is the error probability of a code $\mathcal{C}$ if and only if $\mathcal{C}$ is generalized quasi-perfect.

IV. Symmetric Erasure/Error Channels

Consider a family of symmetric erasure channels $P_{Y|X}$ with discrete input alphabet $\mathcal{X}$, $|\mathcal{X}|=q$, and output alphabet $\mathcal{Y}=\mathcal{X}\cup\{\mathrm{e}\}$ where e corresponds to the erasure symbol. The transition probability of the channels is

$$P_{Y|X}(y|x)=\begin{cases}1-\delta-\epsilon, & y=x,\\ \delta, & y=\mathrm{e},\\ \frac{\epsilon}{q-1}, & \text{otherwise.}\end{cases} \tag{20}$$

This channel corresponds to a q-ary symmetric channel with q-ary inputs whose outputs are either the unchanged input symbol, with probability $1-\delta-\epsilon$, the erasure symbol with probability δ, or any of the other $q-1$ input symbols, with probability $\frac{\epsilon}{q-1}$. This channel includes as particular cases the BSC and the BEC when $q=2$, $\delta=0$ and $\epsilon=0$, respectively.

We consider n uses of this channel. Let $\boldsymbol{x}=(x_1,\ldots,x_n)$ and $\boldsymbol{y}=(y_1,\ldots,y_n)$ denote the channel input and output, respectively. For a given pair of $\boldsymbol{x}$ and $\boldsymbol{y}$, we define the number of erasures and the number of flip-errors, respectively, as

$$e_{\boldsymbol{y}}\triangleq\sum_i \mathbb{1}[y_i=\mathrm{e}], \tag{21}$$

$$d_{\boldsymbol{x},\boldsymbol{y}}\triangleq\sum_i \mathbb{1}[x_i\neq y_i]-e_{\boldsymbol{y}}. \tag{22}$$

The n-dimensional channel transition probability is given by

$$P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})=\delta^{e_{\boldsymbol{y}}}\left(\tfrac{\epsilon}{q-1}\right)^{d_{\boldsymbol{x},\boldsymbol{y}}}(1-\delta-\epsilon)^{n-e_{\boldsymbol{y}}-d_{\boldsymbol{x},\boldsymbol{y}}}. \tag{23}$$

We assume that $\frac{\epsilon}{q-1}<1-\delta-\epsilon$. Otherwise, observing the transmitted symbol at the output of the channel is less likely than observing any of the other $q-1$ symbols. Particularized to the BSC (with $q=2$, $\delta=0$), this assumption reduces to the crossover probability being $\epsilon<\frac{1}{2}$.

We define the auxiliary distribution

$$Q_{\boldsymbol{Y}}(\boldsymbol{y})\triangleq\tfrac{1}{c}\delta^{e_{\boldsymbol{y}}}\left(\tfrac{\epsilon}{q-1}\right)^{D(e_{\boldsymbol{y}})}(1-\delta-\epsilon)^{n-e_{\boldsymbol{y}}-D(e_{\boldsymbol{y}})}, \tag{24}$$

where c is a normalizing constant, and $D(e)\geq 0$ is an arbitrary function of the number of erasures, which can be optimized over. For a binary input channel, a good choice is given by

$$D(e)=\max\left(0,\left\lfloor\tfrac{\lceil n-\log_2 M\rceil-e+1}{2}\right\rfloor\right). \tag{25}$$

Since $Q_{\boldsymbol{Y}}(\boldsymbol{y})$ only depends on $\boldsymbol{y}$ via the number of erasures $e_{\boldsymbol{y}}$ it does not affect the symmetry of the vector channel $P_{\boldsymbol{Y}|\boldsymbol{X}}$ and thus $Q_{\boldsymbol{Y}}\in\mathcal{Q}$. Theorem 1 is applied to this channel and auxiliary distribution $Q=Q_{\boldsymbol{Y}}$ to obtain the following result.

Corollary 1: For the channel with transition matrix in (23), the error probability of any code $\mathcal{C}$ with cardinality M satisfies

$$P_{\mathrm{e}}(\mathcal{C})\geq\sum_{e=0}^{n}\sum_{d=0}^{n-e}\binom{n}{e}\binom{n-e}{d}(q-1)^d\delta^e(1-\delta-\epsilon)^{n-e}\times\left(\varphi^{\max(d,D(e))}-\tfrac{\varphi^{D(e)}}{M}\right), \tag{26}$$

where $\varphi\triangleq\frac{\epsilon}{q-1}(1-\delta-\epsilon)^{-1}$ and $D(e)\geq 0$ is any function of the number of erasures e. Moreover, if $\mathcal{C}$ is a generalized quasi-perfect code for the channel $P_{\boldsymbol{Y}|\boldsymbol{X}}$ with parameters $\gamma=c$ and $Q=Q_{\boldsymbol{Y}}$ as defined in (24), then (26) holds with equality.

A. MDS codes

Let $d_{\min}$ denote the minimum Hamming distance between any pair of codewords in $\mathcal{C}$. The Singleton bound establishes the maximum number of codewords M in a q-ary block code $\mathcal{C}$ of length n and minimum distance $d_{\min}$,

$$\log_q M\leq n-d_{\min}+1. \tag{27}$$

Codes achieving the Singleton bound with equality are termed MDS codes. Examples of MDS codes include those that have only two complementary codewords (having thus minimum distance equal to the blocklength), codes that use the whole input alphabet ($d_{\min}=1$), codes with a single parity symbol ($d_{\min}=2$) and their dual codes. These are often called trivial MDS codes. In the case of binary alphabets, only trivial MDS codes exist. For non-binary alphabets, Reed-Solomon codes are an example of non-trivial MDS codes.

MDS codes are indeed generalized quasi-perfect codes with respect to the q-ary erasure channel, given by $P_{\boldsymbol{Y}|\boldsymbol{X}}$ in (23) when $\epsilon=0$. Letting $\epsilon\to 0$, $\lim_{\epsilon\to 0}\epsilon^A=0$ for any $A>0$ and $\lim_{\epsilon\to 0}\epsilon^A=1$ for $A=0$. Then, for any function $D(e)\geq 0$ such that $D(e)=0$ iff $e>n-\log_q M$, (24) becomes

$$Q_{\boldsymbol{Y}}(\boldsymbol{y})=\begin{cases}0, & e_{\boldsymbol{y}}\leq n-\log_q M,\\ \frac{1}{c}\delta^{e_{\boldsymbol{y}}}(1-\delta)^{n-e_{\boldsymbol{y}}}, & e_{\boldsymbol{y}}>n-\log_q M.\end{cases} \tag{28}$$

Consider a generalized quasi-perfect code according to Definition 2 with parameters $Q=Q_{\boldsymbol{Y}}$ and $\gamma=c$ as defined in (28). For the sets $\mathcal{S}_x(\cdot)$ we use the convention that, when $Q_{\boldsymbol{Y}}(\boldsymbol{y})=0$,

$$\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}{Q_{\boldsymbol{Y}}(\boldsymbol{y})}=\begin{cases}0, & \text{if } P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})=0,\\ \infty, & \text{if } P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})>0.\end{cases} \tag{29}$$

The spheres induced by this code are such that their interior $\mathcal{S}_{\mathrm{i},\boldsymbol{x}}(c,Q_{\boldsymbol{Y}})$ is the set of the output sequences $\boldsymbol{y}$ that are compatible with the input $\boldsymbol{x}$ with a number of erasures $e_{\boldsymbol{y}}\leq n-\log_q M$. Since the codeword-centered interiors do not overlap, the minimum distance of the code is at least $\lfloor n-\log_q M\rfloor+1$. Since the codeword centered shells $\mathcal{S}_{\mathrm{o},\boldsymbol{x}}(c,Q_{\boldsymbol{Y}})$ overlap at some point, $d_{\min}$ is exactly

$$d_{\min}=\lfloor n-\log_q M\rfloor+1. \tag{30}$$

When $\log_q M$ is an integer, this expression coincides with the Singleton bound (27). As a result, MDS codes are also quasi-perfect. By letting $\epsilon\to 0$ in Corollary 1 with $D(e)$ satisfying $D(e)=0$ iff $e>n-\log_q M$, we obtain the following result.

Corollary 2: The error probability of any code $\mathcal{C}$ with cardinality M used over a q-ary erasure channel satisfies

$$P_{\mathrm{e}}(\mathcal{C})\geq\sum_{e=\lfloor n-\log_q M\rfloor+1}^{n}\binom{n}{e}\delta^e(1-\delta)^{n-e}\left(1-\tfrac{q^{n-e}}{M}\right), \tag{31}$$

where equality holds if $\mathcal{C}$ is a generalized quasi-perfect code with parameters $\gamma=c$ and $Q=Q_{\boldsymbol{Y}}$ in (28).

The bound in (31) coincides with the converse bound [1, Th. 38]. As observed in [1], this lower bound is tight when $\mathcal{C}$ is an MDS code. Here this result is recovered via the definition of generalized quasi-perfect codes.

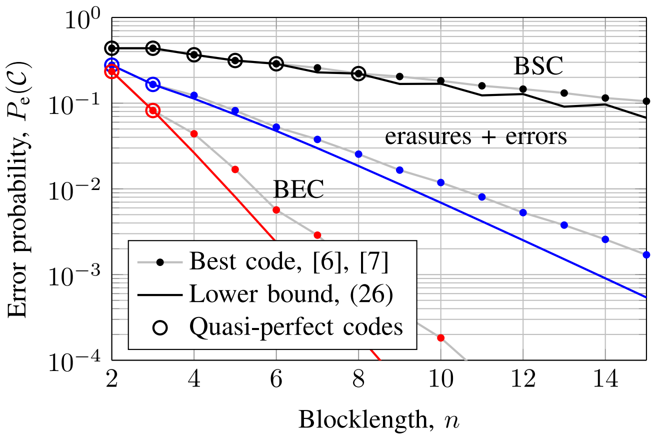

Figure 1. Error probability for n uses of the channel (23), with $q = 2$, $M = 4$ and BSC: $(\epsilon, \delta) = (0.25, 0)$, erasures and errors: $(\epsilon, \delta) = (0.05, 0.2)$, and BEC: $(\epsilon, \delta) = (0, 0.25)$.

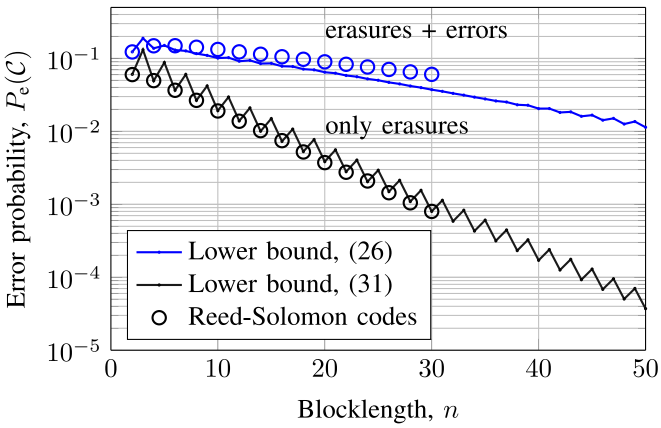

Figure 2. Error probability for n uses of the channel (23) with $q = 32$, fixed transmission rate $R = \frac{1}{n}\log_q M = \frac{1}{2}$, and erasures and errors: $(\epsilon, \delta) = (0.05, 0.25)$, only erasures: $(\epsilon, \delta) = (0, 0.25)$.

B. Examples

Consider the transmission of $M = 4$ codewords over a length-n binary input channel (23) for three sets of parameters: BSC with $(\epsilon, \delta) = (0.25, 0)$, channel with erasures and errors with $(\epsilon, \delta) = (0.05, 0.2)$ and BEC with $(\epsilon, \delta) = (0, 0.25)$. Figure 1 depicts the exact error probability $P_e(\mathcal{C})$ of the best code compared with the lower bound (26) for $D(e)$ given in (25). The optimum codes for the BSC and BEC are from [6] and [7], respectively. For the channel with combined erasures and errors we also use the code for the BEC, since it offers a better performance at the points where they differ. Figure 1 shows that the bound (26) for the BSC coincides with the code error probability at the points where quasi-perfect codes exist with respect to the Hamming distance ($n = 2, 3, 4, 5, 6, 8$). For the BEC, the bound (26) (which coincides with (31)) provides the exact error probability at the points where (trivial) MDS codes exist ($n = 2, 3$), which are generalized quasi-perfect. For the combined errors-erasures channel, the codes need be generalized quasi-perfect for both the BSC and BEC, which only occurs at $n = 2, 3$.

We now consider the q-ary channel (20) with $q = 32$, and fixed transmission rate $R = \frac{1}{n}\log_q M = \frac{1}{2}$. Fig. 2 depicts the lower bound (26) (optimized over a family of functions $D(e)$) for a scenario combining erasures and errors with $(\epsilon, \delta) = (0.05, 0.25)$, and the lower bound (31) for erasures only with $(\epsilon, \delta) = (0, 0.25)$. For even blocklengths, we have simulated the performance of a Reed-Solomon code in both scenarios with 10^6 Monte Carlo realizations. Reed-Solomon codes are defined for blocklengths $n \le q - 1$ and they are generalized quasi-perfect for the q-ary erasure channel. Therefore, they attain the lower bound (31) with equality in the erasure-only case. While their performance with errors and erasures is not far from the lower bound (26), a gap exists in this case.

Appendix

A. Proof of Lemma 2

We prove that the term $\mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau, Q)\bigr]$ does not depend on x. Then, the independence of the other two terms follows since

$$\mathbb{Q}\bigl[\mathcal{S}_{x}(\tau, Q)\bigr] = \sum_{\tau' \in \mathcal{L}_Q,\, \tau' \ge \tau} \mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau', Q)\bigr], \qquad (32)$$

$$\mathbb{Q}\bigl[\mathcal{S}_{\mathrm{i},x}(\tau, Q)\bigr] = \sum_{\tau' \in \mathcal{L}_Q,\, \tau' > \tau} \mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau', Q)\bigr], \qquad (33)$$

where $\mathcal{L}_Q$ is defined in (13). To show that $\mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau, Q)\bigr]$ is independent of x, we write

$$\mathbb{Q}\bigl[\mathcal{S}_{\mathrm{o},x}(\tau, Q)\bigr] = \sum\nolimits_y Q(y) 1\!\!1\bigl[P_{Y|X}(y|x) = \tau Q(y)\bigr] \qquad (34)$$

$$= \frac{1}{\tau} \sum\nolimits_y P_{Y|X}(y|x) 1\!\!1\bigl[P_{Y|X}(y|x) = \tau Q(y)\bigr]. \qquad (35)$$

According to the definition of $\mathcal{Q}$ in (9), for any $Q \in \mathcal{Q}$,

$$F_x(\tau, Q) = \sum\nolimits_y P_{Y|X}(y|x) 1\!\!1\bigl[P_{Y|X}(y|x) \ge \tau Q(y)\bigr] \qquad (36)$$

does not depend on the specific value of x for any $\tau \ge 0$. Then, noting that the summation in (35) is given by $\lim_{\delta\to 0}\left(F_x(\tau, Q) - F_x(\tau + \delta, Q)\right)$, the result follows.

B. Proof of Lemma 3

Let $\mathcal{C} = \{x_1, \ldots, x_M\}$ be an arbitrary code and let $Q \in \mathcal{Q}$. We define $\eta \ge 0$ be the largest value such that $\bigcup_{x\in\mathcal{C}} \mathcal{S}_x(\eta, Q) = \mathcal{Y}$. Similarly, we define $\nu \ge 0$ as the smallest value such that the codeword centered sets $\bigl\{\mathcal{S}_{\mathrm{i},x}(\nu, Q)\bigr\}_{x\in\mathcal{C}}$ are disjoint. We shall respectively refer to η and ν as the *covering* and *packing radius* of the code $\mathcal{C}$ with respect to Q.

We consider a deterministic ML decoder with disjoint decoding regions $\{\mathcal{D}_1, \ldots, \mathcal{D}_M\}$. This set defines a partition of the output space and the error probability (4) becomes

$$P_e(\mathcal{C}) = 1 - \frac{1}{M} \sum_{m=1}^{M} \sum_{y\in\mathcal{D}_m} P_{Y|X}(y|x_m). \qquad (37)$$

For an observed y, the codeword $x \in \mathcal{C}$ that maximizes the metric $P_{Y|X}(y|x)$ coincides with the one maximizing the metric $q(x, y) = \frac{P_{Y|X}(y|x)}{Q(y)}$. Then, using the definition of the covering and packing radius, it follows that

$$\mathcal{S}_{\mathrm{i},x_m}(\nu, Q) \subseteq \mathcal{D}_m \subseteq \mathcal{S}_{x_m}(\eta, Q), \qquad (38)$$

for $1 \leq m \leq M$. As a result, $\mathcal{D}_m$ can be decomposed as

$$\mathcal{D}_m = \mathcal{S}_{\mathrm{i},x_m}(\nu, Q) \bigcup_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \big(\mathcal{D}_m \cap \mathcal{S}_{\mathrm{o},x_m}(\tau, Q)\big), \tag{39}$$

and (37) becomes

$$P_{\mathrm{e}}(\mathcal{C}) = 1 - \frac{1}{M}\sum_{m=1}^{M}\Bigg(\sum_{y\in\mathcal{S}_{\mathrm{i},x_m}(\nu,Q)} P_{Y|X}(y|x_m) + \sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \sum_{y\in\{\mathcal{D}_m\cap\mathcal{S}_{\mathrm{o},x_m}(\tau,Q)\}} P_{Y|X}(y|x_m)\Bigg). \tag{40}$$

Using that $\frac{P_{Y|X}(y|x)}{Q(y)} = \tau$ for any $y \in \mathcal{S}_{\mathrm{o},x}(\tau, Q)$, we write

$$\sum_{y\in\mathcal{S}_{\mathrm{i},x}(\nu)} P_{Y|X}(y|x) = \sum_{y\in\mathcal{S}_{\mathrm{i},x}(\nu,Q)} \frac{P_{Y|X}(y|x)}{Q(y)} Q(y) \tag{41}$$

$$= \sum_{\tau\in\mathcal{L}_Q,\tau>\nu} \sum_{y\in\mathcal{S}_{\mathrm{o},x}(\tau,Q)} \tau Q(y) \tag{42}$$

$$= \sum_{\tau\in\mathcal{L}_Q,\tau>\nu} \tau \mathsf{Q}_{\mathrm{o}}(\tau), \tag{43}$$

where in (43) we used Lemma 2 and $\mathsf{Q}_{\mathrm{o}}(\tau) = \mathbb{Q}\big[\mathcal{S}_{\mathrm{o},x}(\tau, Q)\big]$. Similarly,

$$\sum_{y\in\{\mathcal{D}_m\cap\mathcal{S}_{\mathrm{o},x}(\tau,Q)\}} P_{Y|X}(y|x) = \sum_{y\in\{\mathcal{D}_m\cap\mathcal{S}_{\mathrm{o},x}(\tau,Q)\}} \tau Q(y) \tag{44}$$

$$= \tau \mathsf{Q}_{\mathrm{o},m}(\tau). \tag{45}$$

where we defined $\mathsf{Q}_{\mathrm{o},m}(\tau) \triangleq \mathbb{Q}\big[\mathcal{D}_m \cap \mathcal{S}_{\mathrm{o},x_m}(\tau, Q)\big]$.

Substituting (43) and (45) in (40), yields

$$P_{\mathrm{e}}(\mathcal{C}) = 1 - \Bigg(\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \tau>\nu}} \tau\mathsf{Q}_{\mathrm{o}}(\tau) + \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \tau\mathsf{Q}_{\mathrm{o},m}(\tau)\Bigg). \tag{46}$$

Since $\{\mathcal{D}_m\}_{m=1}^{M}$ defines a partition of the output space, $\sum_{m=1}^{M} \mathbb{Q}\big[\mathcal{D}_m\big] = 1$. Using (39) and the definitions of $\mathsf{Q}_{\mathrm{i}}(\cdot)$ and $\mathsf{Q}_{\mathrm{o},m}(\cdot)$, we obtain

$$1 = \sum_{m=1}^{M} \mathbb{Q}\big[\mathcal{D}_m\big] = M\mathsf{Q}_{\mathrm{i}}(\nu) + \sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \mathsf{Q}_{\mathrm{o},m}(\tau). \tag{47}$$

Upon rearranging terms, (47) yields

$$\nu\left(\frac{1}{M} - \mathsf{Q}_{\mathrm{i}}(\nu)\right) = \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \nu\mathsf{Q}_{\mathrm{o},m}(\tau) \tag{48}$$

$$\geq \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\nu}} \tau\mathsf{Q}_{\mathrm{o},m}(\tau). \tag{49}$$

Then, using (48)-(49) in (46) we obtain $P_{\mathrm{e}}(\mathcal{C}) \geq \Gamma(\nu)$ where

$$\Gamma(\nu) \triangleq 1 - \Bigg(\sum_{\tau\in\mathcal{L}_Q,\tau>\nu} \tau\mathsf{Q}_{\mathrm{o}}(\tau) + \nu\left(\frac{1}{M} - \mathsf{Q}_{\mathrm{i}}(\nu)\right)\Bigg). \tag{50}$$

For quasi-perfect codes satisfying Definition 2, there exist $Q \in \mathcal{Q}$ and $\gamma = \nu = \eta$ such that covering and packing radius coincide. Then, for this choice of parameters, the inequality (49) becomes equality and $P_{\mathrm{e}}(\mathcal{C}) = \Gamma(\gamma)$. We conclude that, for a generalized quasi-perfect code $\mathcal{C}$, (14) holds for any choice (not necessarily unique) of γ and Q satisfying the conditions in Definition 2.

If $\mathcal{C}$ is not generalized quasi-perfect, $\nu > \eta$ for every $Q \in \mathcal{Q}$ and the inequality (49) is strict. Then, $P_{\mathrm{e}}(\mathcal{C}) > \Gamma(\nu)$. Moreover, for any choice of $\gamma \geq 0$ not necessarily the packing radius, we next show that $P_{\mathrm{e}}(\mathcal{C}) > \Gamma(\gamma)$. To see this, note that for $\gamma > \eta$, both (46) and (48)-(49) still hold substituting η by γ. Then, the discussion above still applies.

Assume now that $\eta \leq \gamma < \nu$. We can rewrite (46) as

$$P_{\mathrm{e}}(\mathcal{C}) = 1 - \Bigg(\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \tau>\gamma}} \tau\mathsf{Q}_{\mathrm{o}}(\tau) + \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\gamma}} \tau\mathsf{Q}_{\mathrm{o},m}(\tau)\Bigg) + \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \gamma<\tau\leq\nu}} \tau\Delta_m(\tau). \tag{51}$$

where $\Delta_m(\tau) \triangleq \mathsf{Q}_{\mathrm{o}}(\tau) - \mathsf{Q}_{\mathrm{o},m}(\tau)$. Similarly, (47) becomes

$$1 = M\mathsf{Q}_{\mathrm{i}}(\gamma) + \sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \eta\leq\tau\leq\gamma}} \mathsf{Q}_{\mathrm{o},m}(\tau) - \sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \gamma<\tau\leq\nu}} \Delta_m(\tau). \tag{52}$$

Following analogous steps as in (48)-(49), via (51) we obtain

$$P_{\mathrm{e}}(\mathcal{C}) \geq \Gamma(\gamma) + \frac{1}{M}\sum_{m=1}^{M}\sum_{\substack{\tau\in\mathcal{L}_Q,\\ \gamma<\tau\leq\nu}} (\tau-\gamma)\Delta_m(\tau) > \Gamma(\gamma), \tag{53}$$

as $\tau - \gamma > 0$ and $\Delta_m(\tau) > 0$ in the sum. The same proof steps hold for $\gamma < \eta$. We then conclude that, if $\mathcal{C}$ is not quasi-perfect, $P_{\mathrm{e}}(\mathcal{C}) > \Gamma(\gamma)$ for any $\gamma \geq 0$, $Q \in \mathcal{Q}$.

Acknowledgment

We thank the anonymous referees for calling our attention to [5, Sec. 9.2] and for several valuable comments.

References

[1] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, Inc., 1968.

[3] M. Hamada, "A sufficient condition for a code to achieve the minimum decoding error probability–generalization of perfect and quasi-perfect codes," *IEICE Trans. on Fund. of Electronics, Comm. and Comp. Sciences*, vol. E83-A, no. 10, pp. 1870–1877, Oct. 2000.

[4] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. R. Soc. Lond. A*, vol. 231, no. 694-706, p. 289, 1933.

[5] R. M. Fano, *Transmission of Information: A statistical Theory of Communication*. Cambridge: MIT Press, 1961.

[6] P.-N. Chen, H.-Y. Lin, and S. Moser, "Optimal ultrasmall block-codes for binary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7346–7378, Nov 2013.

[7] H.-Y. Lin, S. M. Moser, and P.-N. Chen, "Weak flip codes and their optimality on the binary erasure channel," *arXiv:1711.03310v1*, 2017.