Quantum Error Detection II: Bounds

A. Ashikhmin * A. Barg † E. Knill ‡ S. Litsyn §

Abstract

In Part I of this paper we formulated the problem of error detection with quantum codes on the depolarizing channel and gave an expression for the probability of undetected error via the weight enumerators of the code. In this part we show that there exist quantum codes whose probability of undetected error falls exponentially with the length of the code and derive bounds on this exponent. The lower (existence) bound is proved for stabilizer codes by a counting argument for classical self-orthogonal quaternary codes. Upper bounds are proved by linear programming. First we formulate two linear programming problems that are convenient for the analysis of specific short codes. Next we give a relaxed formulation of the problem in terms of optimization on the cone of polynomials in the Krawtchouk basis. We present two general solutions of the problem. Together they give an upper bound on the exponent of undetected error. The upper and lower asymptotic bounds coincide for a certain interval of code rates close to 1.

Index Terms — Probability of undetected error, quantum code, self-orthogonal codes, polynomial method.

1 Introduction

In part I of this paper we defined the undetected error event for transmission with quantum codes over a completely depolarized channel and explained a way to compute its probability via the weight enumerators of the code. This part is independent of part I once we agree on the definitions of a quantum code, the channel and error correction, and the weight enumerators. The main results of part I can be summarized as follows. Let Q be an $((n, K))$ quantum code,

*Present address: Bell Laboratories, Lucent Technologies, 600 Mountain Avenue 2C-179, Murray Hill, NJ 07974, aea@research.bell-labs.com. Previous address: Los Alamos National Laboratory, Los Alamos, NM 87545.

†Bell Laboratories, Lucent Technologies, 600 Mountain Avenue 2C-375, Murray Hill, NJ 07974. E-mail: abarg@research.bell-labs.com

‡Los Alamos National Laboratory Group CIC-3, Mail Stop P990, Los Alamos, NM 87545. E-mail: knill@c3serve.c3.lanl.gov

§Department of Electrical Engineering-Systems, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: litsyn@eng.tau.ac.il

i.e., a linear K-dimensional subspace of $\mathcal{H}_n := (\mathbb{C}^2)^{\otimes n}$. Let

$$\begin{aligned} B(x,y) &= \sum_{i=0}^{n} B_i x^{n-i} y^i, \\ B^\perp(x,y) &= \sum_{i=0}^{n} B_i^\perp x^{n-i} y^i \end{aligned}$$

be the (Shor-Laflamme) weight polynomials of Q, where the weight distributions B_i, $0 \le i \le n$, and $B_i^\perp, 0 \le i \le n$, are given by (I.9) and (I.10), respectively[1]. Then the probability of undetected error equals

$$P_{ue}(Q,p) = \sum_{i=1}^{n} (B_i^\perp - B_i) \Big(\frac{p}{3}\Big)^i (1-p)^{n-i}, \tag{1}$$

where p is the probability for a factor τ_i in the error operator $E = \tau_1 \otimes \cdots \otimes \tau_n$ to be nonidentical (this probability does not depend on i). In some situations this formula should contain an additional constant factor (see Theorem I.6) which throughout this part will be omitted.

The goal of this part of the paper is to derive bounds on $P_{ue}(Q,p)$ for the best possible code Q with given parameters. More specifically, in Part I we defined the quantity

$$P_{ue}(n,K,p) = \min_{\substack{Q \in \mathcal{H}_n \\ \dim(Q)=K}} P_{ue}(Q,p).$$

We will derive upper and lower bounds on $P_{ue}(n,K,p)$. Just as in the classical case, this probability falls exponentially with n; therefore, let us also introduce the exponent of undetected error

$$E(n,R_Q,p) = -\frac{1}{n} \log_2 P_{ue}(\lceil 2^{R_Q n} \rceil, n, p),$$

where $R_Q = \frac{\log_2 K}{n}$ is the code rate. Speaking of asymptotics, we are interested in upper bounds

$$\overline{E}(R_Q,p) = \limsup_{n\to\infty} E(n,R_Q,p)$$

and lower bounds

$$\underline{E}(R_Q,p) = \liminf_{n\to\infty} E(n,R_Q,p)$$

(this corresponds to lower and upper bounds on the probability $P_{ue}(n,K,p)$, respectively). Let $E(R_Q,p)$ be the common limit of these functions, provided that it exists.

Throughout the paper

$$\begin{aligned} T_q(x,y) &= x \log_q(q-1) - x \log_q y - (1-x) \log_q (1-y) \\ H_q(x) &= T_q(x,x). \end{aligned}$$

In the classical case, error detection has been extensively studied. The probability of undetected error in the classical case is defined in (I.1); its exponent $E^{(cl)}(R,p)$ is defined similarly

[1] References (I.9), Theorem I.6, and so on point to equations, theorems, etc., in the first part of this paper.

to the above. Best known lower bounds on $E^{(cl)}(R,p)$ (upper bounds on the probability) were derived in [9] building upon the Varshamov-Gilbert-type existence arguments. We consider the binary case only. Let $R_{vg}(x) := 1 - H_2(x)$ be the Varshamov-Gilbert function and $\delta_{vg}(x)$ its inverse. Also let $\bar{R}(x)$ be the function giving the best known upper bound on codes [12] and $\bar{\delta}(x)$ its inverse. It is easy to prove that there exist binary linear codes with $A_i \leq n\binom{n}{i}2^{k-n}$, where A_i is the number of vectors of weight i in the code. Substituting this in (I.1), we obtain the lower bound of [9]:

$$E^{(cl)}(R,p) \geq \begin{cases} T_2(\delta_{vg}(R),p) & 0 \leq R \leq 1 - H(p) \\ 1 - R & 1 - H_2(p) \leq R \leq 1. \end{cases}$$

Upper bounds require more involved arguments [3], [11]. The results have the form

$$E^{(cl)}(R,p) \leq \begin{cases} 1 - R - H_2(\bar{\delta}(R),p) + T_2(\bar{\delta}(R),p) & 0 \leq R \leq \bar{R}(p) \\ 1 - R & \bar{R}(p) \leq R \leq 1. \end{cases}$$

In this paper we derive analogous results for the quantum case. In Section 2 we prove the existence of quantum stabilizer codes with bounded above weight enumerators; a substitution in (1) yields lower bounds on $E(R,p)$. In this part we rely upon the results of [13] on quaternary self-orthogonal codes. Then we move on to lower bounds on $P_{ue}(n,K,p)$. In this part we employ the linear programming technique. In Section 3 we formulate a linear programming (LP) problem with the objective function $P_{ue}(n,K,p)$. Though in examples this problem gives good lower bounds (which can be found by solving it on a computer), analytically it is difficult to deal with. Therefore, in the second part of this section we propose a relaxation of the problem which enables us to derive general bounds. This part of the paper is based on an application of the LP approach in the quantum case [4] in conjunction with the methods of [3], [11]. The results include two upper bounds on $E(R,p)$. These bounds show that for the rate R_Q in a certain neighborhood of 1, dependent on p, the exponent $E(R,p)$ is known exactly. For lower rates, the bounds are in general location, i.e., there exists a segment in which each of them is better than the other. This part of the paper is technically the most involved. We chose to formulate the results for arbitrary q (instead of concentrating on $q = 4$), the reasons being that once we look at $q > 2$, it does not make much of a difference whether it is 4 or anything else, and that this is helpful in studying error detection of nonbinary classical codes on which we plan to report elsewhere. Moreover, the theory of quantum codes generalizes to larger q [8], [14], though the presentation is somewhat less systematic and the results more scattered than for binary quantum codes.

Some further remarks on notation. Throughout the paper $F = \mathbb{F}_4 = \{0, 1, \omega, \omega^2\}$. The Krawtchouk polynomial is given by

$$K_k(q;x) = \sum_{\ell=0}^{n} (-1)^\ell \binom{x}{\ell}\binom{n-x}{k-\ell}(q-1)^{k-\ell}$$

(the implicit parameter n – the length of the code – is usually clear from the context). Properties of $K_k(q;x)$ used throughout the paper are summarized in the appendix. As remarked above, by R_Q we denote the rate of the quantum code. We also use two associated numbers $|C| =$

$\sum_{i=0}^{n} B_i$ and $|C^\perp| = \sum_{i=0}^{n} B_i^\perp$; in the case of stabilizer codes they are equal to the size of the two underlying classical codes, C and $C^\perp$ (see Part I). Likewise, let $R = (1/n)\log_4 |C|$, $R^\perp = (1/n)\log_4 |C^\perp|$. The rate R_Q and these 2 quantities are connected by the following relations:

$$R_Q = 2R^\perp - 1 = 1 - 2R, \quad R + R^\perp = 1, \quad 0 \le R \le \frac{1}{2} \le R^\perp \le 1. \tag{2}$$

2 Upper (existence) bounds on $P_{ue}(n, K, p)$

In this section we show that there exist quantum codes for which probability of undetected error falls exponentially. for all rates $0 \le R_Q < 1$, and bound this exponent below More specifically, we prove the following theorem.

Theorem 1

$$E(R_Q, p) \ge \begin{cases} T_4\Big(H_4^{-1}\big(\frac{1}{2}(1-R_Q)\big), p\Big) & 0 \le R_Q \le 2(1-H_4(p)) - 1; \\ \frac{1}{2}(1-R_Q) & 2(1-H_4(p)) - 1 \le R_Q \le 1. \end{cases}$$

To prove this theorem, we restrict our attention to quantum stabilizer codes. In analogy with the classical case, we show that there exist sequences of codes Q of growing length n and size $K = 2^{k_Q}$ with weight distribution

$$B_i^\perp \le n^2 \binom{n}{i} 3^i 2^{k_Q - n}.$$

(in fact, n^2 can be easily replaced by n).

Below $F := \mathbb{F}_4$. Since the weight distributions $B_i, B_i^\perp$ correspond to classical quaternary code, we prove this estimate by considering families of quaternary self-orthogonal codes. Let $C \subset F^n$. Throughout the section we denote by $C^\perp$ a linear code dual to C with respect to the Hermitian inner product $(a, b) = a\bar{b}$, where $\bar{b} = b^2$ for any $b \in F$. Let

$$\mathcal{S}_{n,k} = \{C \subset F^n \mid C \text{ even linear code}, \dim C = k\}$$
$$\mathcal{S}_{n,k}^\perp = \{C \subset F^n \mid C^\perp \in \mathcal{S}_{n,k}\},$$

where $k \le n/2$ by (2).

Remark. By the definition, a stabilizer code Q is constructed from a pair $C \subset C^\perp$ of codes over F, where the orthogonality is understood with respect to the trace inner product $\langle u, v\rangle = \mathrm{tr}_{\mathbb{F}_4/\mathbb{F}_2}(u\bar{v})$. However, as is easily checked [13], the equality $(u, v) = 0$ for even codes implies $\langle u, v\rangle = 0$.

Below we use the following results from [13].

Lemma 2 *Let $C \subset F^n$ be an even linear code and $C^\perp$ be its dual. Then the number of even-weight code vectors in $C^\perp$ equals $\frac{1}{2}(4^{n-k} + (-2)^n)$.*

Lemma 3 *Let $C \subset F^n$ be an even linear code and $C^\perp$ be its dual. If $\mathbf{a} \in C^\perp$ has even (odd) weight then the coset $C + \mathbf{a}$ is formed by vectors of even (odd) weight.*

Existence of codes with bounded distance distribution will follow from the following lemma, based on Lemmas 2-3.

Lemma 4 *Let $\mathbf{v} \in F^n$ be any nonzero even-weight vector. The number of codes from $\mathcal{S}$ containing $\mathbf{v}$ does not depend on $\mathbf{v}$.*

PROOF. Let us count the number of linear $[n,k]$ codes containing $\mathbf{v}$. Let $C_1 = \langle \mathbf{v} \rangle$ be the $[n,1]$ code and $\mathbf{w}$ an even-weight vector distinct from $\mathbf{v}$ such that $(\mathbf{w},\mathbf{v}) = 0$. By Lemma 3 all the cosets $C_1 + \alpha\mathbf{w}, \alpha \in F$, are even; so by adjoining $\mathbf{w}$ we obtain an even $[n,2]$ code C_2. By Lemma 2 this can be done in

$$\frac{1}{4}\Big[\frac{1}{2}\Big(4^{n-1} + (-2)^n\Big) - 1\Big]$$

ways independent of $\mathbf{v}$. Similarly C_2 can be extended to an even $[n,3]$ code in

$$\frac{1}{16}\Big[\frac{1}{2}\Big(4^{n-2} + (-2)^n\Big) - 1\Big]$$

ways. Continuing in this manner, we obtain all even $[n,k]$ codes that contain $\mathbf{v}$. It is obvious that their number does not depend on a particular choice of $\mathbf{v}$. □

Denote by $B_i(C)$ the number of codewords of weight i of $C \in \mathcal{S}_{n,k}$. Let

$$\tilde{B}(x,y) = \frac{1}{|\mathcal{S}_{n,k}|} \sum_{C \subset \mathcal{S}_{n,k}} B_i(C) x^{n-i} y^i$$

be the average weight enumerator of the family $\mathcal{S}_{n,k}$. Similarly let $\tilde{B}^\perp(x,y)$ be the average enumerator for $\mathcal{S}_{n,k}^\perp$.

Theorem 5

$$\begin{aligned} \tilde{B}(x,y) &= (1-\alpha)x^n + \frac{\alpha}{2}\Big((x+3y)^n + (x-3y)^n\Big), \\ \tilde{B}^\perp(x,y) &= \frac{1}{4^k}\Big((1-\alpha)(x+3y)^n + \frac{\alpha}{2}((4x)^n + (-2x+6y)^n)\Big), \end{aligned}$$

where

$$\alpha = \frac{4^k - 1}{\frac{1}{2}(4^n + (-2)^n) - 1}.$$

PROOF. Let $N = |\mathcal{S}_{n,k}|$. By Lemma 4 every nonzero even vector $\mathbf{v}$ is contained in one and the same number, say L, of codes from $\mathcal{S}_{n,k}$.

$$\begin{aligned}
\tilde{B}(x,y) &= \frac{1}{N}\sum_{C\in\mathcal{S}_{n,k}}\sum_{i=0}^{n} B_i(C)x^{n-i}y^i \\
&= \frac{1}{N}\sum_{i=0}^{n} x^{n-i}y^i \sum_{\substack{\mathbf{v}\in F^n \\ \mathrm{wt}(\mathbf{v})=i}} \sum_{\substack{C\in\mathcal{S}_{n,k} \\ \mathbf{v}\in C}} 1 \\
&= x^n + \frac{1}{N}\sum_{\substack{i\geq 2 \\ i \text{ even}}} x^{n-i}y^i \binom{n}{i} 3^i \sum_{\substack{C\in\mathcal{S}_{n,k} \\ \mathbf{v}\in C}} 1 \\
&= x^n\left(1-\frac{L}{N}\right) + \frac{1}{2}\frac{L}{N}\Big((x+3y)^n + (x-3y)^n\Big).
\end{aligned}$$

From the boundary condition $\tilde{B}(1,1) = 4^k$ we find that $\frac{L}{N} = \alpha$.

Let $B_C(x,y)$ be the weight polynomial of a code C. Then

$$\begin{aligned}
\tilde{B}^{\perp}(x,y) &= \frac{1}{N}\sum_{C^{\perp}\in\mathcal{S}_{n,k}^{\perp}} B_{C^{\perp}}(x,y) \\
&= \frac{1}{N}\sum_{C\in\mathcal{S}_{n,k}} \frac{1}{|C|} B_C(x+3y, x-y) \\
&= \frac{1}{|C|}\tilde{B}(x+3y, x-y).
\end{aligned}$$

□

This theorem implies that there exists an $[[n,k]]$ quantum stabilizer code, say Q, such that

$$P_{ue}(Q,p) \leq \tilde{B}^{\perp}\left(1-p, \frac{p}{3}\right) - \tilde{B}\left(1-p, \frac{p}{3}\right).$$

Substituting the $\tilde{B}$ and $\tilde{B}^{\perp}$ found above then proves the second part of Theorem 1, but not its first part. To derive the complete theorem, as usual, we must first prove existence of sequences of codes with large minimum distance. By the Chebyshev inequality one can easily see that

$$\Pr\big\{B_i(C) \leq n^2\tilde{B}_i \text{ for all } i = 1, \ldots, n\big\} \geq 1 - \frac{1}{n}.$$

Hence from Theorem 5 it follows that there exists a code $C \in \mathcal{S}_{n,k}$ such that

$$B_i(C^{\perp}) \leq \frac{n^2}{4^k} 3^i \binom{n}{i}\Big((1-\alpha) + (-1)^{n-i}2^n\frac{\alpha}{2}\Big), \quad i = 1, \ldots, n. \tag{3}$$

It is easy to see that

$$B_i(C^{\perp}) \lesssim \binom{n}{i} 3^i |C^{\perp}| 4^{-n} < \exp_4 n\big(H_4\big(\frac{i}{n}\big) + R^{\perp} - 1\big)$$

i.e., $B_i(C^\perp)$ asymptotically is at most binomial. If

$$\frac{i}{n} < H_4^{-1}\left(1 - R^\perp\right) = \delta_{vg}(R^\perp),$$

where $\delta_{vg}(R^\perp)$ is the asymptotic Varshamov-Gilbert bound for rate $R^\perp$, the exponent of $B_i(C^\perp)$ is negative and therefore $B_i(C^\perp) < 1$. Since $C^\perp$ is a linear code, all $B_i(C^\perp)$ are integer. Hence for sufficiently large n

$$B_i(C^\perp) = 0,\ 1 \le i \le (\delta_{vg}(R^\perp) - \varepsilon)n,\ \text{for any } \varepsilon > 0. \tag{4}$$

Remark. From Theorem 5 it follows that $B_i(C)$ are also asymptotically at most binomial, i.e.

$$B_i(C) \lesssim \binom{n}{i} 3^i |C| 4^{-n}.$$

PROOF (of Theorem 1). Let Q be an $[[n,k]]$ quantum code such that its enumerator $B_i^\perp$ satisfies (3) and (4). Then

$$\begin{aligned} P_{ue}(Q,p) &\le B^\perp\left(\frac{1-p.p}{3}\right) - B\left(1-p, \frac{p}{3}\right) \\ &\le B^\perp\left(1-p, \frac{p}{3}\right) - (1-p)^n \\ &= \sum_{i=\lfloor n\delta_{vg}(R^\perp)\rfloor}^{n} \frac{n^2}{4^k}\binom{n}{i} 3^i \left((1-\alpha) + (-1)^{n-i}\frac{\alpha}{2}\right)(1-p)^{n-i}\left(\frac{p}{3}\right)^i. \end{aligned}$$

The exponent of the summation term equals

$$-n\Big[T_4\Big(\frac{i}{n}, p\Big) - H_4\Big(\frac{i}{n}\Big) + (1-R_Q)/2\Big], \tag{5}$$

where we have omitted the $o(1)$ terms. The expression in the brackets attains its maximum for $i/n = p$. Hence as long as $(i/n) \ge \delta_{vg}(R^\perp)$, the sum asymptotically is dominated by the term with $i = \lfloor np \rfloor$. Thus as long as $1 - H_4(p) \le R^\perp = \frac{1}{2}(1 + R_Q)$, we have

$$E(R_Q, p) \le \frac{1}{2}(1 - R_Q),$$

i.e., the second part of the theorem. Otherwise, the maximum moves outside the summation range, so the largest term asymptotically is the first one in the sum, i.e., the one corresponding to $i = \lfloor n\delta_{vg}(R^\perp)\rfloor$. This gives the first part of the theorem. □

3 A linear program for quantum undetected error

The approach leading to best known lower bounds on the probability of undetected error in the classical case has been the linear programming one [3], [11]. In this section we develop a similar technique for the quantum case. First we formulate two theorems that enable one to obtain

good lower estimates on $P_{ue}(n,K,p)$ for finite n. Then we formulate a relaxed LP problem which is not as good for finite n but lends itself to asymptotic analysis.

For the reasons outlined in the introduction we will study a general alphabet of size r. An r-ary quantum code Q is a K-dimensional linear subspace of $\mathbb{C}^{r^n}$. Without going into details we say that one can associate with Q two weight distributions, B_i and $B_i^\perp$, $0 \le i \le n$, connected by the q-ary MacWilliams identities, $q = r^2$. Furthermore,

$$K = \frac{1}{q^n} \sum_{i=0}^{n} B_i^\perp.$$

As above, we use the notation

$$|C^\perp| = \sum_{i=0}^{n} B_i^\perp, \quad |C| = \sum_{i=0}^{n} B_i,$$
$$R = (1/n)\log_q |C|, \quad R^\perp = (1/n)\log_q |C^\perp|;$$

these numbers and the rate R_Q are again related through (2). We have for the probability of undetected error

$$P_{ue}(Q,p) = \sum_{j=1}^{n} (B_j^\perp - B_j)\Big(\frac{p}{q-1}\Big)^j (1-p)^{n-j}. \tag{6}$$

Our first result is given by the following theorem.

Theorem 6 *Let Q be an $((n, r^{nR_Q}))$ quantum code. Let $q = r^2$, $R^\perp = (1+R_Q)/2$. Let $Z(x) = \sum_{i=0}^n z_i K_i(q;x)$ and $Y(x) = \sum_{i=0}^n y_i K_i(q,x)$ be polynomials such that*

$$\begin{aligned} &Z(j) - Y(j) + y_0 + y_j q^{nR^\perp} \\ &\qquad \le q^{nR^\perp}\Big(\frac{p}{q-1}\Big)^j (1-p)^{n-j} - \Big(\frac{q-1-qp}{q-1}\Big)^j, 1 \le j \le n, \\ &z_0, y_0 \lessgtr 0;\ z_j \ge 0,\ y_j \ge 0, 1 \le j \le n. \end{aligned} \tag{7}$$

Then

$$P_{ue}(n, r^{nR_Q}, p) \ge q^{-nR^\perp}\Big(z_0 q^{nR^\perp} - Z(0) + Y(0) - y_0\Big) + (1-p)^n - q^{-nR^\perp}.$$

Proof. Using the MacWilliams identities we can rewrite (6) as follows:

$$\begin{aligned} P_{ue}(Q,p) \;=\; & \frac{1}{|C^\perp|}\sum_{j=1}^{n} B_j^\perp \Big[|C^\perp| v\Big(\frac{p}{q-1}\Big)^j (1-p)^{n-j} \\ & - \sum_{t=1}^{n} K_t(j)\Big(\frac{p}{q-1}\Big)^t (1-p)^{n-t}\Big] - \frac{1}{|C^\perp|}\sum_{j=1}^{n}\binom{n}{j} p^j (1-p)^{n-j}. \end{aligned}$$

The middle term here is calculated using the generating function (36):

$$\sum_{t=0}^{n} K_t(j)\Big(\frac{p}{q-1}\Big)^t (1-p)^{n-t} = \Big(\frac{q-1-qp}{q-1}\Big)^j.$$

Hence

$$\begin{aligned} P_{ue}(Q,p) &= \frac{1}{|C^\perp|}\sum_{j=1}^{n} B_j^\perp\Big[|C^\perp|\Big(\frac{p}{q-1}\Big)^j(1-p)^{n-j} - \Big(\frac{q-1-qp}{q-1}\Big)^j + (1-p)^n\Big] \\ &\quad -\frac{1}{|C^\perp|}\sum_{j=1}^{n}\binom{n}{j}p^j(1-p)^{n-j} \\ &= \frac{1}{|C^\perp|}\sum_{j=1}^{n} B_j^\perp\Big[|C^\perp|\Big(\frac{p}{q-1}\Big)^j(1-p)^{n-j} - \Big(\frac{q-1-qp}{q-1}\Big)^j\Big] \\ &\quad +(1-p)^n - \frac{1}{|C^\perp|}. \end{aligned} \tag{8}$$

Thus, $P_{ue}(Q,p)$ is a linear form of the coefficients $B_i^\perp$ which we have to minimize. We can formulate the following LP problem:

$$\min\Big\{\sum_{j=1}^{n} B_j^\perp\Big[|C^\perp|\Big(\frac{p}{q-1}\Big)^j(1-p)^{n-j} - \Big(\frac{q-1-qp}{q-1}\Big)^j\Big]\Big\} \tag{9}$$

subject to the restrictions

$$\begin{aligned} B_j^\perp &\geq 0, \quad 1\leq j\leq n \\ \sum_{j=1}^{n} B_j^\perp &= |C^\perp|-1 \\ \sum_{j=1}^{n} B_j^\perp K_i(q;j) &\geq -\binom{n}{i}(q-1)^i, \quad 1\leq i\leq n \\ B_j^\perp|C^\perp| - \sum_{i=1}^{n} B_i^\perp K_j(q;i) &\geq \binom{n}{j}(q-1)^j, \quad 1\leq j\leq n. \end{aligned}$$

The last inequality follows from Theorem I.2(i).

Now the theorem follows by the LP duality. Indeed, the dual program has $2n+1$ variables $(z_0, z_1, \ldots, z_n)$ and $(y_1, y_2, \ldots, y_n)$ of which z_0 can take on any value and all the other variables are nonnegative. The dual objective function has the form

$$\max\Big\{z_0(|C^\perp|-1) - \sum_{i=1}^{n}(z_i-y_i)\binom{n}{i}(q-1)^i\Big\}$$

subject to restrictions

$$\begin{aligned} z_0 &+ \sum_{i=1}^{n}(z_i-y_i)K_i(q;j) + y_j|C^\perp| \\ &\leq |C^\perp|\Big(\frac{p}{q-1}\Big)^j(1-p)^{n-j} - \Big(\frac{q-1-qp}{q-1}\Big)^j \quad 1\leq j\leq n. \end{aligned}$$

Any feasible solution of the this problem gives a lower estimate of $P_{ue}(n,K,p)$. Let us introduce the polynomials $Y(x)=\sum_{i=0}^{n} y_i K_i(q;x)$ and $Z(x)=\sum_{i=0} z_i K_i(q;x)$. Since $K_0(q;x)\equiv 1$, this implies our claim. $\square$

Sometimes it is convenient to rewrite the linear problem via the enumerator B_j instead of $B_j^{\perp}$. The proof of the following theorem is similar to the above.

Theorem 7 *Suppose $Q,n,r,R_Q,q,Z(x),Y(x)$ have the same meaning as in Theorem 6. Let $R=(1-R_Q)/2$. Suppose that*

$$\begin{aligned} &Z(j)+Y(j)-y_0-y_j q^{nR} \\ &\qquad \le \Big(\frac{q-1-pq}{q-1}\Big)^j - q^{nR}\Big(\frac{p}{q-1}\Big)^j (1-p)^{n-j}, \quad 1\le j\le n, \qquad (10)\\ &z_0, y_0 \lessgtr 0,\ z_j\ge 0,\ y_j\ge 0, \quad 1\le j\le n. \end{aligned}$$

Then

$$P_{ue}(R_q,p)\ge q^{-nR}\Big(z_0 q^{nR}-Z(0)-Y(0)+y_0\Big)-(1-p)^n+q^{-nR}.$$

Though the LP problems of Theorems 6 and 7 enable one to find bounds for short codes with the use of computer, they are not easy to analyze in general. The reason for this is that the sign of the quantities on the right-hand side of (7) or (10) alternates. This significantly complicates checking feasibility of a putative solution. For this reason below we take on a different approach which, though it does not yield optimal solutions for the LP problem, gives rise to good asymptotic upper bounds on $E(R,p)$.

Theorem 8 *Let Q be an r-ary quantum code with weight enumerators B_i and $B_i^{\perp}$. Let $h(i), 1\le i\le n$, be a real-valued function and*

$$Z(x)=\sum_{i=0}^{n} z_i K_i(q;x), \quad (q=r^2)$$

be a polynomial that for $1\le i\le n$ satisfies the conditions

$$\text{(i)}\quad Z(i)\le h(i); \qquad (11)$$

$$\text{(ii)}\quad z_i|C^{\perp}|-Z(i)\ge 0. \qquad (12)$$

Then

$$\sum_{i=0}^{n}(B_i^{\perp}-B_i)h(i)\ge z_0|C^{\perp}|-Z(0). \qquad (13)$$

PROOF. The proof will follow from the following chain of relations:

$$\begin{aligned}\sum_{i=0}^{n}(B_i^{\perp}-B_i)h(i) &\geq \sum_{i=0}^{n}(B_i^{\perp}-B_i)Z(i)\\ &\stackrel{\text{(a)}}{=} \sum_{i=0}^{n}\Big(\frac{1}{|C|}\sum_{j=0}^{n}B_jK_i(q;j)-B_i\Big)Z(i)\\ &= \sum_{j=0}^{n}B_j\frac{1}{|C|}\sum_{i=0}^{n}K_i(q;j)Z(i)-\sum_{i=0}^{n}B_iZ(i)\\ &\stackrel{\text{(b)}}{=} \sum_{j=0}^{n}B_j\frac{r^n}{|C|}z_j-\sum_{j=0}^{n}B_jZ(j)=\sum_{j=0}^{n}B_j(|C^{\perp}|z_j-Z(j))\\ &\geq |C^{\perp}|z_0-Z(0),\end{aligned}$$

where the first inequality follows by (i) and the obvious $B_0 = B_0^{\perp} = 1$; step (a) is implied by the MacWilliams identities, in (b) we use (38), and the final inequality follows by (ii) and the fact that $B_j \geq 0, B_0 = 1$. □

If (12) is replaced by the condition

$$Z(i) - |C|z_i \geq 0,\ i = 1, \ldots, n, \tag{14}$$

then by a similar argument one can prove that

$$\sum_{i=0}^{n}(B_i^{\perp}-B_i)h(i) \geq Z(0) - |C|z_0.$$

We wish to stress the difference between the conditions on $Z(x)$ in this theorem and in the classical (non-quantum) case [3]. The standard condition $z_j \geq 0, 1 \leq j \leq n$, is not needed to prove (13); it is replaced by related though not equivalent conditions (12), (14). In the situation when one bounds above the size of the quantum code, the corresponding inequality is $|C| \leq \max_{1\leq j\leq n}(Z(j)/z_j)$; see [4] for details.

In the next section we use Theorem 8 to derive asymptotic bounds on $E(R,p)$.

4 Lower bounds on $P_{ue}(n, K, p)$

In this section we prove two lower bounds on the probability of undetected error that are valid for any quantum code of given length and size. The bounds are derived by a suitable choice of polynomials in Theorem 8. The results are similar in spirit to [3], [4], [11]. In this section $\gamma = q - 1$,

$$h(x) = (p/\gamma)^x(1-p)^{n-x}, \tag{15}$$

and as usual $0 \leq p < \gamma/q$.

4.1 An Aaltonen-MRRW-type bound

In this part we rely on the technique in [12][2], extended to arbitrary q in [1]. Let

$$\tau_0(z) := \frac{\gamma}{q} - \frac{\gamma-1}{q}z - \frac{2}{q}\sqrt{\gamma z(1-z)} \quad (0 \le z \le (\gamma/q)). \tag{16}$$

By [12], [1] $R_q^{lp}(\delta) := H_q(\tau_0(\delta))$ is the maximal asymptotically attainable rate of a q-ary (classical) code with relative distance δ. This is proved by studying Delsarte's linear programming problem with the polynomial

$$f_t(x) = \frac{(K_t(q;x) + K_{t+1}(q;x))^2}{a-x},$$

where $t = \lfloor n\tau \rfloor$, $\tau = \tau_0(\delta)$, and a is the smallest zero of $K_t(q;x) + K_{t+1}(q;x)$. Conversely, the function $\delta_q^{lp}(R) := \tau_0(H_q^{-1}(R))$ gives an asymptotic upper bound on the minimum distance of q-ary codes of rate R. (An interesting remark is that the function $\tau_0(\cdot)$ is involutive; this is intimately related to the self-duality of the Hamming scheme [6]).

We begin with an appropriate modification (rescaling) of the polynomial $f(x)$. Let

$$Z_t(x) = \frac{h(x^*)}{K_t^2(q;x^*)} f_t(x), \tag{17}$$

where $x^* = \xi^* n$ is an integer parameter, $0 < x^* < n$, and $t = \lfloor n\tau_0(\xi^*) \rfloor$. This choice is motivated by the following argument. The polynomial has to satisfy the inequality (11); any reasonable choice of $Z(x)$ suggests that it be equal to $h(x)$ at least at one point $x = x^*$. This point is left a free parameter, chosen later.

The program (17) gives rise to the following bound. For the reasons revealed in [4] and outlined in footnote 3 below, in the quantum case the bound is valid for all but very low rates. Below $\alpha_1(q)$ is a certain small positive number dependent only on q. It can be computed for any q; for instance, $\alpha_1(4) \approx 0.0028$. We could not find a closed-form expression for it.

Theorem 9 *Let $\alpha_1(q) \le R_Q \le 1$, $R = (1/2)(1-R_Q)$. Then*

$$E(R_Q, p) \le \begin{cases} R - H_q(\delta_q^{lp}(1-R)) + T_q(\delta_q^{lp}(1-R), p) & \alpha_1(q) \le R_Q \le 2R_q^{lp}(p) - 1; \\ R & 2R_q^{lp}(p) - 1 \le R_Q \le 1. \end{cases} \tag{18}$$

Proof. We first prove the bound (18) and then prove feasibility of the program (17). By (42) we obtain

$$Z_t(0) = \frac{\gamma^{2t} h(x^*)}{a K_t^2(q;x^*)} \binom{n}{t}^2 \Big(\frac{t+1+\gamma(n-t)}{t+1}\Big)^2 \tag{19}$$

Further, by [1],

$$z_0 = \frac{q\gamma^t h(x^*)}{(t+1)K_t^2(q;x^*)} \binom{n}{t}. \tag{20}$$

[2]The abbreviation in the title is derived from its authors' names.

We would like to substitute these values in (13). Recall the notation $R^\perp = (1/n)\log_q |C^\perp| = (1/2)(1+R_Q)$. Note that whenever

$$\frac{\log_q \binom{n}{t}^2 \gamma^{2t}}{\log_q \left(\binom{n}{t}\gamma^t |C^\perp|\right)} \to 0,$$

i.e., $H_q(\tau_0(\xi^*)) + R^\perp > 2H_q(\tau_0(\xi^*))$ or

$$R^\perp \geq H_q(\tau_0(\xi^*))$$

we have $Z_t(0) = o\Big(z_0|C^\perp|\Big)$. The restriction $R^\perp > H_q(\tau_0(\xi^*))$ by our choice of τ is equivalent to $R^\perp \geq R_q^{lp}(\xi^*)$, which is always the case whenever $\xi^* \geq \delta_q^{lp}(R^\perp)$. In this case the main term of the estimate (13) is given by the exponent of $z_0|C^\perp|$. Differentiating $(1/n)\log_q z_0$ on ξ^*, we obtain

$$\log_q \frac{p(1-\xi^*)}{(1-p)\xi^*}.$$

The zero of this expression is $\xi^* = p$, and $d(\log_q z_0)/d\xi^*$ it is negative for $\xi^* \in (p,1)$. Thus, the optimal choice of ξ^* is $\xi^* = p$ if this value is not less than $\delta_q^{lp}(R^\perp)$ and $\xi^* = \delta_q^{lp}(R^\perp)$ otherwise. By (20), (41), and the fact that the exponent of $h(x)$ is given by $-T_q(\xi, p)$, we obtain

$$\frac{1}{n}\log_q z_0 = H_q(\tau) - T_q(\xi^*, p) - 1 - H_q(\tau) + H_q(\xi^*) \tag{21}$$

Substituting this in (13), we obtain (18).

Let us prove that polynomial (17) is admissible with respect to the restrictions (11)-(12). The proof will be broken into two cases,

(a) $\xi \in (0, \xi^*]$ and

(b) $\xi \in (\xi^*, 1)$.

We begin with the first case and (11). We are only going to prove that it holds asymptotically, i.e., to prove the inequality

$$\frac{1}{n}\log_q Z_t(\xi n) \leq \frac{1}{n}\log_q h(\xi n). \tag{22}$$

Here we employ a method used in the corresponding part of [3]. Namely, by our choice of τ, the smallest zero (39) of K_t tends to ξ^*; hence in the interval considered the exponent of K_t is given by (40). Then we can write

$$\begin{aligned}\frac{1}{n}\log_q Z(\xi n) &= -T_q(\xi^*, p) \\ &- 2\int_{\xi^*}^{\xi} \log_q \frac{(1-y)\gamma + y - q\tau + \sqrt{((1-y)\gamma + y - q\tau)^2 - 4\gamma y(1-y)}}{2\gamma(1-y)} dy.\end{aligned}$$

with $o(1)$ terms omitted. Let $\psi(\tau, \xi) := 2\int_0^\xi \ldots dy + T_q(\xi, p)$; then we have

$$\frac{1}{n}\log_q Z(\xi n) + T_q(\xi, p) = \psi(\tau, \xi) - \psi(\tau, \xi^*).$$

Since $(1/n)\log_q Z_t(x^*) = (1/n)\log_q h(x^*)$, all we need to prove is that

$$\psi'_\xi(\tau,\xi) = 2\log_q \frac{(1-\xi)\gamma+\xi-q\tau+\sqrt{((1-\xi)\gamma+\xi-q\tau)^2-4\gamma\xi(1-\xi)}}{2\gamma(1-\xi)} - \log_q \frac{p}{(1-p)\gamma} > 0, \quad 0 \le \xi < \xi^*.$$

First note that $\psi'_\xi(\tau,\xi)$ is a monotone decreasing function of $\tau = \tau_0(\xi^*)$ and $\tau_0(z)$ is a monotone decreasing function of z (16). Hence if we prove that ψ'_ξ is positive for $\xi^* = p$ this will also imply that it is positive for $\xi^* = \delta_q^{lp}(R^\perp) > p$. Therefore, let $\xi^* = p$. In [3, Appendix B] a similar function was proved to be positive. The proof proceeds as follows: consider the difference

$$g(\tau,\xi) = \frac{\left[(1-\xi)\gamma+\xi-q\tau+\sqrt{((1-\xi)\gamma+\xi-q\tau)^2-4\gamma\xi(1-\xi)}\,\right]^2}{4\gamma^2(1-\xi)^2} - \frac{p}{(1-p)\gamma},$$

$\xi \in [0,\xi^*), \tau = \tau_0(p)$. The required inequality $\psi'_\xi(\tau_0(p),\xi) > 0$ is implied by $g(\tau_0(p),\xi) \ge 0$; the latter follows by the fact that $g(\tau_0(p),p) = 0$ and that upon substituting τ and simplifying we obtain a fraction whose denominator is positive and the derivative of the numerator on ξ is negative in the whole segment $\xi \in [0,\xi^*]$.

Now let us prove (12) in case (a). According to [4] the function $\frac{1}{n}[\log Z_t(\xi n) - \log z_{\xi n}]$, $\xi \in (0,\xi^*)$, achieves its maximum at $\xi = 0$ for $1 \ge R_Q \ge \alpha_1(q)$[3]. If $p \le \delta_{LP1}(R^\perp)$ then as said above, we put $\xi^* = \delta_{LP1}(R^\perp)$. This means that

$$\log_q \frac{Z_t(0)}{z_0} \to H_q\Big(\frac{t}{n}\Big) = R_q^{lp}(\xi^*) = R^\perp,$$

and so $z_0|C^\perp| - Z_t(0) \to 0$. Therefore, for any integer s, $0 < s \le x^*$, we have $z_s|C^\perp| - Z_t(s) \ge 0$.

Finally if $p = \xi^* > \delta_q^{lp}(R^\perp)$ then $(1/n)\log_q(Z_t(0)/z_0)$ converges to a number less than $R^\perp$. Hence for sufficiently large n and any integer $s \in (0,x^*]$ we have $z_s|C^\perp| - Z_t(s) > 0$. This takes care of case (a).

To verify feasibility in case (b), i.e., to prove (11)-(12) for $x^* < x \le n$, we recall that a is the smallest zero of $K_t(q;x) + K_{t+1}(q;x)$. Let y_s be the smallest zero of $K_s(q;x)$. Then by the well-known properties of Krawtchouk polynomials we obtain that $y_{t+1} < a < y_t$; so by (39), $a \searrow y_{t+1}$ as $n \to \infty$. Then we have, for large n and all integer x, $x^* < x \le n$,

$$Z_t(x) \le 0 \le \frac{p^x}{\gamma^x}(1-p)^{n-x};$$

hence (11). To prove (12), observe that if $x > x^*$ then $Z_t(x) \le 0$ and so for any integer $s \ge x$ we have $z_s|C^\perp| - Z_t(s) \ge 0$. This exhausts case (b) and completes the proof. □

[3] This is the reason for the bound to fail for very low rates both in [4] and here: the maximum shifts away from 0 and the analysis becomes unmanageable.

4.2 Hamming-type bound

In the R-δ problem for nonbinary codes the bound [1] is not the best one known. It can be improved in several ways, in particular, in the frame of the polynomial method a better result is given in [2]. However, the technique in [2] does not readily carry over to the present situation. Another, somewhat simpler bound that improves upon [1] is the Hamming one which is better for a certain segment of rates close to 1. Therefore, in this subsection we derive a Hamming-type bound on $E(R_Q, p)$. This bound is valid for low error probabilities: $p \in [0, p_{cr}]$, where the critical value p_{cr} depends on q (it is 0.19 for $q = 2$ and 0.30 for $p = 4$). This improves Theorem 9 for some values of R_Q dependent on q and further extends the segment in which the exponent $E(R_Q, p)$ is known exactly.

We begin with the polynomial

$$F_e(x) = \sum_{i=0}^{n} f_i K_i(q; x),$$

where $f_i = K_e(q; i)^2$. This polynomial is used in the proof of the Hamming bound on the size of the code with a given minimum distance [5]. Our first goal is to show how to modify it for use in our problem.

Delsarte [6, p.13] proved that

$$K_i(q; x) K_j(q; x) = \sum_{k=0}^{n} p_{ij}^k K_k(q; x).$$

where p_{ij}^k are the intersection numbers of the q-ary Hamming scheme. By a straightforward generalization of the binary case [12, (A.19)] we have

$$p_{ij}^k = \sum_{s=0}^{n-k} \binom{k}{2k+2s-i-j} \binom{n-k}{s} \binom{2k+2s-i-j}{k+s-j} (q-2)^{i+j-2s-k} (q-1)^s$$

Therefore,

$$f_i = \sum_{k=0}^{n} p_{ii}^k K_k(q; i).$$

Substituting in $F_e(x)$, we obtain the following:

$$\begin{aligned}
F_e(x) &= \sum_{j=0}^{n}\sum_{k=0}^{n}\sum_{s=0}^{n-k}\binom{k}{2k+2s-2e}\binom{n-k}{s}\binom{2k+2s-2e}{k+s-e} \\
&\qquad\qquad \cdot (q-2)^{2e-2s-k}(q-1)^s K_k(q;j)K_j(q;x) \\
&= \sum_{k=0}^{n}\sum_{s=0}^{n-k}\binom{k}{2k+2s-2e}\binom{n-k}{s}\binom{2k+2s-2e}{k+s-e} \\
&\qquad\qquad \cdot (q-2)^{2e-2s-k}(q-1)^s \sum_{j=0}^{n} K_k(q;j)K_j(q;x) \\
&= q^n \sum_{s=\max\{0,e-x\}}^{e-x/2}\binom{x}{2x+2s-2e}\binom{n-x}{s}\binom{2x+2s-2e}{x+s-e}(q-2)^{2e-2s-x}(q-1)^s,
\end{aligned}$$

where in the last step we made use of (37).

Let us analyze the asymptotics of $F_e(x)$. Letting $x=\xi n$, $s=\sigma n$, and $e=\tau n$, we can write the exponent of the summation term as follows:

$$\begin{aligned}
&\frac{1}{n}\log_q\Big[\binom{x}{2x+2s-2e}\binom{n-x}{s}\binom{2x+2s-2e}{x+s-e}(q-2)^{2e-2s-x}(q-1)^s\Big] \\
&\qquad = \Big[\xi H_2\Big(\frac{2\xi+2\sigma-2\tau}{\xi}\Big)+2\xi+2\sigma-2\tau\Big]\log_q 2 + (1-\xi)H_q\Big(\frac{\sigma}{1-\xi}\Big) \\
&\qquad\qquad + (2\tau-2\sigma-\xi)\log_q(q-2) + O\Big(\frac{1}{n}\Big). \qquad (24)
\end{aligned}$$

Computing the derivative of the last expression on σ and equating it to 0, we arrive at the following condition:

$$\frac{(2\sigma+\xi-2\tau)^2(1-\xi-\sigma)(q-1)}{\sigma(2\sigma+2\xi-2\tau)^2(q-2)^2} = \frac{1}{4}. \qquad (25)$$

It is not difficult to check (see [4, Appendix]) that this equation has only one root in the interval $\max\{0,\tau-\xi\} < \sigma < \tau-\xi/2$. Denote this root by σ_0. The main term in $F_e(x)$ asymptotically corresponds to the value $s=\lfloor\sigma_0 n\rfloor$. Thus, defining

$$\begin{aligned}
\phi(\tau,\sigma,\xi) &:= 1+(1-\xi)H_q\Big(\frac{\sigma}{1-\xi}\Big)+\Big[\xi H_2\Big(\frac{2\xi+2\sigma-2\tau}{\xi}\Big)+2\xi+2\sigma-2\tau\Big]\log_q 2 \\
&\qquad + (2\tau-2\sigma-\xi)\log_q(q-2), \qquad (26)
\end{aligned}$$

we observe that

$$\frac{1}{n}\log_q F_e(x) = \phi(\tau,\sigma_0,\xi) + O\Big(\frac{1}{n}\Big).$$

The analysis is complicated by the fact that σ_0 itself is a function of τ and ξ.

Our general plan is, starting with $F_e(x)$, to construct a polynomial $Z(x)$ so that $Z(x)$ be equal to $h(x)$ at one point and less than $h(x)$ at all other integer points of the interval, thus

guaranteeing feasibility. Together with (25) this gives two conditions on the 2 parameters, σ_0 and τ, both functions of ξ. It remains to make a suitable choice for ξ; this we simply guess prompted by an analogy in the binary case. This is the actual sequence of steps that we perform to derive the bound. Calculations, though elementary, are fairly involved, and we will not write them out in full. Instead, we perform a similar analysis in the binary case; this can be done explicitly within reasonable space and fixes ideas for the general result.

THE BINARY CASE. We have the following simple expression for $F_e(x)$:

$$\begin{aligned} F_e(x) &= 2^n \binom{x}{x/2}\binom{n-x}{e-x/2} \\ &= \sum_{i=0}^{n} K_i(2,x) \sum_{k=0}^{n} \binom{n-k}{e-k/2}\binom{k}{k/2} K_k(2;i). \end{aligned} \tag{27}$$

Note that $F_e(x) = 0$ when $x > 2e$. The exponents of $F_e(x)$ and $h(x)$ are

$$\begin{aligned} \phi(\tau,\xi) := \frac{1}{n}\log_2 F_e(x) &= 1 + \xi + (1-\xi)H_2\Big(\frac{2\tau-\xi}{2-2\xi}\Big) + o(1), \\ \frac{1}{n}\log_2 h(x) &= -T_2(\xi,p). \end{aligned} \tag{28}$$

Now let $x^* = \xi^* n$ be a point at which these exponents have equal slopes. Let us first convince ourselves that such a point exists and is unique. Indeed, the polynomial $\phi'_\xi(\tau,\xi) + (T_2(\xi,p))'_\xi$ is quadratic in ξ; its zeros are

$$1 \pm \frac{(1-p)(1-2\tau)}{\sqrt{1-2p}}.$$

Of them the one with the + sign is greater than one; the other one is always between 0 and 1.

The equation $\big[\phi'_\xi(\tau,\xi) = -(T_2(\xi,p))'_\xi\big]_{\xi=\xi^*}$ defines τ as a function of ξ^*. Namely,

$$\tau = \tau(p) = \frac{1}{2}\Big(1 - \frac{\sqrt{1-2p}}{1-p}(1-\xi^*)\Big). \tag{29}$$

Now we can define the polynomial $Z(x)$ by rescaling $F_e(x)$ as follows. Let

$$Z(x) = \frac{h(x^*)}{F_e(x^*)} F_e(x),$$

where $e = \lfloor \tau n \rfloor$. Note that we have ensured that $Z(x)$ equals $h(x)$ at x^* and that their exponents are tangent; it will be seen below that for large n, $Z(x) < h(x)$ at all the other integer points of the interval. It remains to choose the value of ξ^*. This point is taken to maximize z_0, i.e., the estimate (13). Namely, the logarithm of z_0 equals

$$\frac{1}{n}\log_2 z_0 = 2H_2(\tau) - T_2(\xi^*,p) - \phi(\tau,\xi^*) + o(1). \tag{30}$$

Substituting τ from (29) and taking the derivative on ξ^*, we find that the optimal choice is $\xi^* = p$. Note that substituting $\xi^* = p$ and τ from (29) in (30), we find $(1/n)\log_2 z_0 \to -1$.

Let us examine feasibility of $Z(x)$. As a preliminary remark, note that we are allowed to put $\xi^* = p$ as long as p is greater than the Hamming distance and we have to take ξ^* equal to this distance otherwise. Indeed, substituting $x = 0$ in (27), we obtain

$$\frac{1}{n}\log_2 F_{\tau n}(0) \sim 1 + H_2(\tau)$$
$$\frac{1}{n}\log_2 f_0 = \frac{1}{n}\log K^2_{\lfloor \tau n \rfloor}(2;0) \sim 2H_2(\tau),$$

the latter by (42). From this and the definition of $Z(x)$ it follows that whenever

$$H_2(\tau) \geq 1 - R^\perp, \tag{31}$$

we have $Z(0) = o(z_0|C^\perp|)$, needed for the estimate (13) to be nontrivial. So we can choose $\xi^* = p$, $\tau = \tau(p)$ if (31) holds and we choose $\tau = H_2^{-1}(1 - R^\perp)$ (the Hamming distance for the rate $R^\perp$) and ξ^* the root of $H_2(\tau(\xi)) = 1 - R^\perp$ otherwise.

Note that by (29) τ grows on ξ^*. Thus, (11) will follow in both cases if we prove that it holds for all $(1/2)(1 - \sqrt{1-2p}) \leq \tau \leq (1/2)$. As above, we are only going to prove that (11) holds asymptotically, i.e., that

$$\log_2 Z(x) - \log_2 h(x) \leq 0, \quad 0 \leq \xi \leq 2\tau, \ \frac{1}{2}(1 - \sqrt{1-2p}) \leq \tau \leq \frac{1}{2}. \tag{32}$$

We begin with the case $\xi^* = p$. Note that $(1/n)\log_2 h(x)$ is a straight line; its derivative is $\log_2(p/(1-p)) < 0$. Inequality (32) will follow from the following set of conditions:

(i) $\phi'_\xi(\tau, \xi) = \log_2(p/(1-p))$ has a unique zero for $\xi \in [0, 2e]$;

(ii) $0 > \phi'_\xi(\tau, \xi)|_{\xi=0} > \log_2(p/(1-p))$;

(iii) $\phi'_\xi(\tau, \xi)|_{\xi=2e} < \log_2(p/(1-p)) < 0$,

where $\tau = (1/2)(1 - \sqrt{1-2p})$.

Condition (i) was established above. Conditions (i)-(iii) imply that $\phi_\xi(\tau, \xi) < -T_2(\xi, p)$ for $\xi \in [0, p)$ and $\phi_\xi(\tau, \xi) < -T_2(\xi, p)$ for $\xi \in [0, p)$. Indeed, suppose that $\phi(\tau, \xi) > \log_2(p/(1-p)$ in the neighborhood of ξ^*, say on the left of it. This implies that in the small neighborhood of the point of tangency the derivative ϕ'_ξ is smaller that $\log_2(p/(1-p))$; however by (ii) it is greater than that for $\xi = 0$, hence there is another point between 0 and p at which ϕ and $-T_2$ have equal slopes, but this violates (i). Supposing that ϕ and $-T_2$ intersect at some point between 0 and ξ^*, we again find a similar contradiction. The second part of the claim follows by the same argument. Thus to establish (32) it suffices to prove (ii)-(iii).

We have

$$\phi'_\xi(\tau, \xi) = \frac{1}{2}\log_2 \frac{(2\tau - \xi)(2 - 2\tau - \xi)}{(1-\xi)^2};$$

so

$$\phi'_\xi((1/2)(1 - \sqrt{1-2p}, 0) = \frac{1}{2}\log_2 2p > \log_2 \frac{p}{1-p} \quad (0 \leq p < \frac{1}{2}).$$

This proves (ii). Condition (iii) is equally elementary; we omit the easy check. This establishes (11) for $\xi^* = p$.

Now suppose that $\xi^* > p$. Condition (i) was proved above for any τ. To prove (ii) and (iii) we only have to show that $\phi'_\xi(\tau, 0)$ grows and $\phi'_\xi(\tau, 2\tau)$ falls on τ. Observe that $\phi'_\xi(\tau, 0) = \frac{1}{2}\log_2 4\tau(1-\tau)$ indeed grows as long as $\tau < 1/2$, which is true, and $\phi'_\xi(\tau, \xi)$ falls indefinitely as $\xi \to 2\tau$. This proves (11), or rather (32), for $q = 2$.

To prove (12), we choose the following tactics. We already know by (31) that (12) holds for $\xi \to 0$. Hence it suffices to prove that the expression $\log_2 Z(\lfloor \xi n \rfloor)/z_{\lfloor \xi n \rfloor}$ achieves its maximum at $\xi = 0$. Numerical computations show that this condition holds if $0 \le \tau \le \tau_1 \approx 0.1069$.

Finally, by the definition, $Z(x) = 0$ for $x \ge \lceil 2\tau n \rceil$. Hence for these values of x (12) holds trivially. Otherwise by the preceding paragraph, (12) is true at least as long as i is less than the smallest root of $K_{\lfloor \tau n \rfloor}$ since otherwise the coefficients $z_i = (K_{\lfloor \tau n \rfloor}(i))^2$ of $Z(x)$ can be very small. The smallest zero is given by (16); so the discussed constraint is satisfied in particular if $\tau \le \tau_2$, where $\tau_2 = 0.1$ is a root of $2\tau = (1/2) - \sqrt{\tau(1-\tau)}$. Thus a sufficient condition for (12) to hold true is that $\tau \le \tau_{cr} = \min\{\tau_1, \tau_2\} = 0.1$. Note also that $\tau(p)$ is monotone increasing in p. Hence we substitute $\xi^* = p$ in (29) and denote by p_{cr} the root of $\tau(p) = \tau_{cr}$, $p_{cr} = 0.18$.

In summary, we obtain the following theorem.

Theorem 10 *Let $R = (1/2)(1 - R_Q)$. Then for any $0 \le p \le p_{cr}$*

$$E(R_Q, p) \le \begin{cases} -1 - R + T_2(\xi, p) + \phi(H_2^{-1}(R), \xi) & 1 - 2H_2(\tau_{cr}) \le R_Q \le 1 - 2H_2(\tau(p)); \\ R & 1 - 2H_2(\tau(p)) \le R_Q \le 1, \end{cases}$$

where ξ is a root of $H_2(\tau(\xi)) = R$, and $\phi(\cdot)$ is given by (28).

This completes the argument in the binary case.

THE GENERAL CASE. The analysis is similar but significantly more complicated since apart from τ and ξ we have a third parameter, σ. In this part we are more sketchy than above. The polynomial $Z(x)$ is again taken in the form

$$Z(x) = \frac{h(x^*)}{F_e(x^*)} F_e(x),$$

where this time $h(x)$ is as in (15) and $x^* = \xi^* n$ is a parameter.

By (26)

$$(1/n)\log_q Z(x) = \phi(\tau, \sigma_0, \xi) - T_q(\xi^*, p) - \phi(\tau, \sigma_0, \xi^*).$$

We proceed exactly as above. Namely, from the equation $\phi'_\xi(\tau, \sigma_0, \xi^*) = (T_q(\xi^*, p))'_\xi$ we find τ as a function of ξ^* and σ_0. This gives

$$\tau = \tau(\xi^*, \sigma_0) := \frac{\xi^*(\xi^* + \sigma_0 - 1)(q - 1 - p) - p\sigma_0(q-2)}{p(\xi^* - 1)(q-2)} + \frac{\sqrt{(q-1)(1-p)\xi^{*2}\left[(1-\xi^*)^2(q-1-p) - \sigma_0(1-\xi^*)(2q-2-qp) + \sigma_0^2(q-1)(1-p)\right]}}{p(\xi^* - 1)(q-2)}.$$

Next, we substitute this value of τ in (25) and find σ_0 as a function of ξ^*. This gives

$$\sigma_0 := \sigma_0(\xi^*) = \frac{2q-2-qp-2(1-\xi^*)\sqrt{(q-1)(q-1-qp)}}{q^2(1-p)}. \tag{33}$$

To complete the definition of the parameters we have to chose ξ^*. As above, we take $\xi^* = p$ as long as this does not violate the feasibility condition (12)[4]. Substituting $x = 0$ in (23), we obtain

$$\frac{1}{n}\log_q F_{\tau n}(0) \sim 1 + H_q(\tau)$$
$$\frac{1}{n}\log_q f_0 = \frac{1}{n}\log K^2_{\lfloor \tau n \rfloor}(q;0) \sim 2H_q(\tau),$$

the latter by (42). From this and the definition of $Z(x)$ it follows that whenever

$$H_q(\tau) \geq 1 - R^\perp \tag{34}$$

we have $Z(0) = o(z_0|C^\perp|)$.

So the best possible choice is $\xi^* = p$, $\tau = \tau(p,\sigma_0)$ if (34) holds and $\tau = H_q^{-1}(1-R^\perp)$ (the Hamming distance for the rate $R^\perp$) and ξ^* the root of $H_q(\tau(\xi,\sigma_0)) = 1 - R^\perp$ otherwise. Computations with Maple show that in the first case $\frac{1}{n}\log_q z_0 \to -1$, exactly as in the binary case above. We did not find a closed-form expression for the second case.

Similarly to the binary case τ should satisfy the inequality

$$2\tau \leq \frac{\gamma}{q} - \frac{\gamma-1}{q}\tau - \frac{2}{q}\sqrt{\gamma\tau(1-\tau)} \quad \text{(cf. (16))},$$

or

$$\tau \leq \tau_2 := \frac{(3\sqrt{q}-2\sqrt{2})(q-1)}{\sqrt{q}(9q-8)}.$$

Also similarly to the binary case we have to choose τ such that

$$\arg\max_{\xi} \left\{ \log_2 \frac{Z(\lfloor \xi n \rfloor)}{z_{\lfloor \xi n \rfloor}} \right\} = 0. \tag{35}$$

For $\tau = 0$ this maximum is obviously achieved at $\xi = 0$ (note that $F_0(x) = q^n\delta_{x0}$ and $f_{\lfloor x \rfloor} = 1$). Define $\tau_1, 0 \leq \tau_1 \leq \tau_2$, as follows:

$$\tau_1 = \inf(\tau : \ (35) \text{ does not hold}).$$

Let $\tau_{cr} = \tau_1$ if τ_1 is well-defined and $\tau_{cr} = \tau_2$ otherwise. The function again $\tau(p)$ is increasing in p. Let p_{cr} be the root of $\tau(p,\sigma_0(p)) = \tau_{cr}$. Now we are ready to formulate the theorem.

Theorem 11 *Let $R = (1/2)(1-R_Q)$. Then for any $0 \leq p \leq p_{cr}$*

$$E(R_Q,p) \leq \begin{cases} -1-R+ & T_q(\xi^*,p) + \phi(H_q^{-1}(R),\sigma_0(\xi^*),\xi^*), \\ & 1-2H_q(\tau_{cr}) \leq R_Q \leq 1-2H_q(\tau(p,\sigma_0)); \\ R, & 1-2H_q(\tau(p,\sigma_0)) \leq R_Q \leq 1, \end{cases}$$

where ξ^* is a root of $H_q(\tau(\xi,\sigma_0(\xi))) = R$, σ_0 is defined in (33), and $\phi(\cdot)$ is given by (26).

[4]Though we do not prove this, this choice of ξ^* is optimal with respect to the bound (13).

Remark For $q = 4$, $p_{cr} = 0.301$, and numerical computations show that (35) holds true in the entire interval $\tau \in [0, \tau_2]$. Therefore in this case $\tau_{cr} = \tau_2$.

Fore reference purposes we composed a short table of values of the bounds for $q = 4$, $p = 0.1$.

R_Q	Existence	A-MRRW	Hamming
0	0.5260	0.6270	–
0.1	0.4637	0.5458	–
0.2	0.4054	0.4685	0.4774
0.3	0.3509	0.3952	0.3951
0.4	0.3	0.3262	0.3216
0.5	0.25	0.2618	0.2567
0.6	0.2	0.2028	0.2003
0.7	0.15	0.15	0.15
0.8	0.1	0.1	0.1
0.9	0.05	0.05	0.05
1	0	0	0

These bounds are also plotted in Fig. 1. It can be seen that the Hamming bound is the best of the two upper bounds for large rates. Unlike the classical case, the upper bounds do not approach the lower bound as the rate R_Q becomes small. However this is due rather to the way of measuring the rate of quantum codes than to an imperfection of the method. Indeed, roughly speaking, the case $R_Q = 0$ corresponds to classical codes of rate $R = 1/2$ (cf. (2)). The function $E(R_Q, p)$ is known exactly at least for $2R_4^{lp}(p) - 1 \le R_Q \le 1$. In fact, by Theorem 11 the left end of this interval is provably smaller than this value; however, it is difficult to make any exact statements other than just plotting the bounds.

Acknowledgment. We are grateful to an anonymous referee for helpful suggestions regarding Theorem 1.

5 Appendix

Let $K_i(q; x)$ be the q-ary Krawtchouk polynomial, $\gamma = q - 1$. Here we list its properties used in the paper.

The following 3 basic facts are well known:

$$\sum_{i=0}^{n} K_i(q; x) z^i = (1 + \gamma z)^{n-x} (1 - z)^x \quad (x \text{ integer}); \tag{36}$$

$$\sum_{i=0}^{n} K_r(q; i) K_i(q; s) = q^n \delta_{rs}; \tag{37}$$

$$f(x) = \sum_{i=0}^{t} f_i K_i(q; x) \quad \Leftrightarrow \quad f_i = q^{-n} \sum_{j=0}^{n} f(j) K_j(q; i), \tag{38}$$

where in (38) $f(x)$ is any polynomial with $\deg f \le n$.

Let y_s be the smallest zero of K_s. For $s = \sigma n$, $n \to \infty$ we have [1],

$$\frac{y_s}{n} = \tau(\sigma) + o(1), \tag{39}$$

where the function $\tau(\cdot)$ is defined in (16). Further, by [7] for $n \to \infty$ and $\xi \in [0, \tau(\sigma)]$

$$\frac{1}{n} \log K_s(q; \xi n) \sim H_q(\sigma)$$
$$+ \int_0^{\xi} \log \frac{(1-y)\gamma + y - q\sigma + \sqrt{((1-y)\gamma + y - q\sigma)^2 - 4\gamma y(1-y)}}{2\gamma(1-y)} dy. \tag{40}$$

In particular, for $\sigma = \tau(\xi)$, i.e., $\xi = \tau(\sigma)$,

$$\frac{1}{n} \log_q K_{\sigma n}(q; \xi n) = \frac{1 + H_q(\sigma) - H_q(\xi)}{2} + o(1). \tag{41}$$

Finally, from the definition of K_s we find

$$K_s(q; 0) = \binom{n}{s} \gamma^s. \tag{42}$$

References

[1] M. Aaltonen, "Linear programming bounds for tree codes," *IEEE Trans. Info. Theory*, vol. 25, no. 1, pp. 85–90, 1979.

[2] ———, "A new upper bound on nonbinary block codes," *Discrete Math.*, vol. 83, no. 2-3, pp.139–160, 1990.

[3] A. Ashikhmin and A. Barg, "Binomial moments of the distance distribution: Bounds and applications," *IEEE Trans. Info. Theory*, vol 45, no. 2, pp. 438-452, 1999.

[4] A. Ashikhmin and S. Litsyn, "Upper bounds of the size of quantum codes," *IEEE Trans. Info. Theory*, vol 45, no. 4, pp.1205-1215, 1999.

[5] P. Delsarte, "Bounds for unrestricted codes, by linear programming," *Philips Res. Repts,* **27** (1972), 272–289.

[6] ———, *An Algebraic Approach to the Association Schemes of Coding Theory*, Philips Research Reports Supplements, No. 10, 1973.

[7] G. Kalai and N. Linial, "On the distance distribution of codes," *IEEE Trans. Inform. Theory*, vol. 42, pp.1467–1472, 1995.

[8] E. Knill, Non-binary unitary error bases and quantum codes, Los Alamos National Laboratory Report LAUR-96-2717 (1996).

[9] V. I. Levenshtein, "Bounds on the probability of undetected error," *Problemy Peredachi Informatsii*, vol. 13, no. 1, pp. 3–18, 1977.

[10] ———, "Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 1303–1321, 1995.

[11] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Info. Theory*, vol. 45, no. 2, pp. 385–398, 1999.

[12] R. J. McEliece, E. R. Rodemich,H. C. Rumsey, Jr. and L. R. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams Inequalities,"*IEEE Trans. Inform. Theory*, vol. 23, pp.157–166, 1977.

[13] F. J. MacWilliams, A. M. Odlyzko, and N. J. A. Sloane, "Self-dual codes over $GF(4)$," *J. of Combin. Theory*, Ser. A. vol. 25, pp. 288–318, 1978.

[14] E. Rains, Nonbinary quantum codes, *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1837–1832, 1999.

[15] P.W. Shor and R. Laflamme, "Quantum analog of the MacWilliams identities in classical coding theory," *Phys. Rev. Lett.*, vol. 78, pp. 1600-1602, 1997.

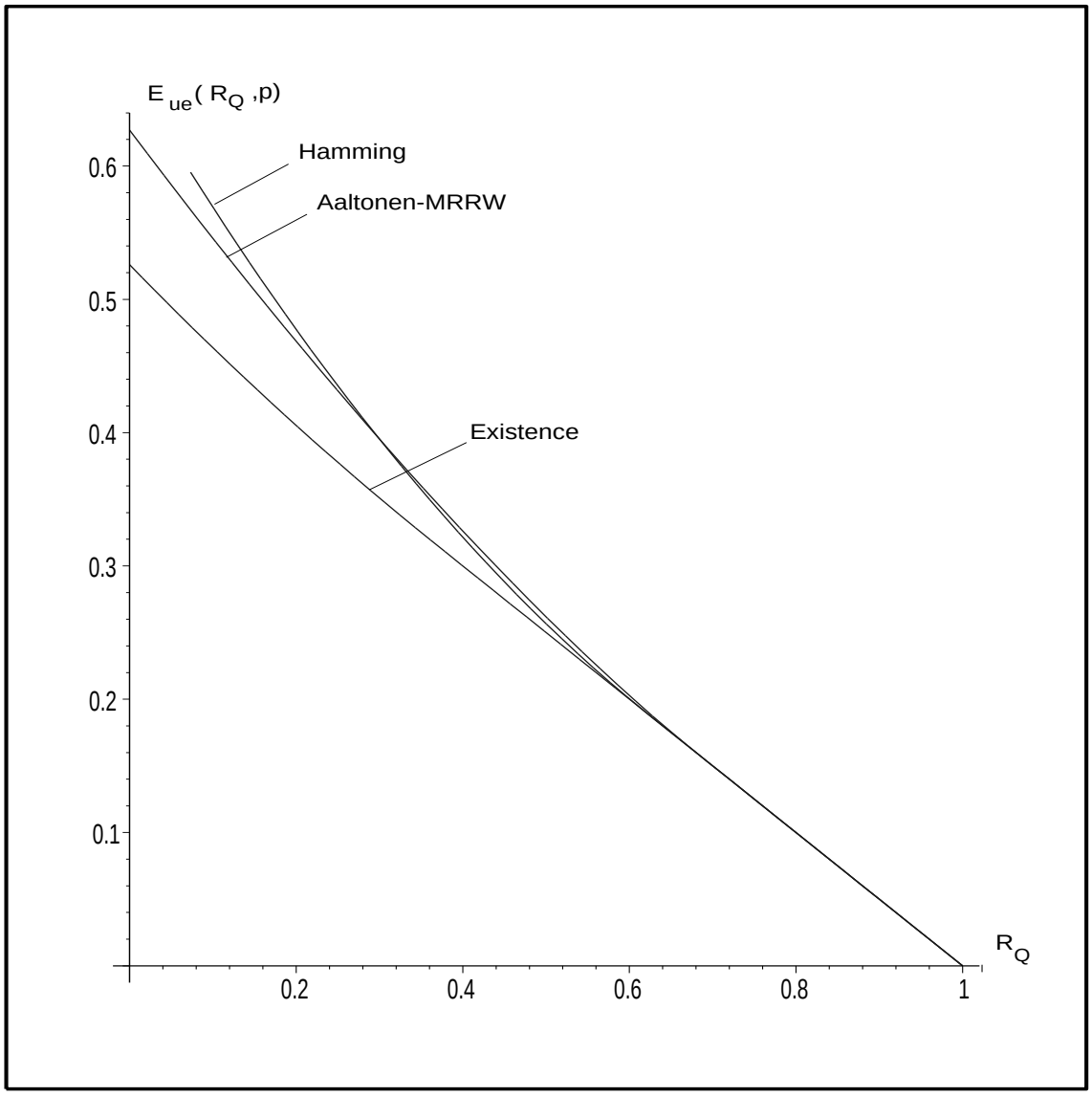


Figure 1: Bounds on $E(R_Q, p)$: Existence bound, Theorem 1; Aaltonen-MRRW-type bound, Theorem 9; Hamming-type bound, Theorem 11; $p = 0.1$.