[6] A. El Gamal and T. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, pp. 1466–1483, 1980.

[7] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum channel," *Probl. Inform. Transm.*, vol. 9, no. 3, pp. 177–183, 1973.

[8] ——, "Problems in the mathematical theory of quantum communication channels," *Rep. Math. Phys.*, vol. 12, no. 2, pp. 273–278, 1977.

[9] ——, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inform. Theory*, vol. 44, pp. 269–273, Jan. 1998.

[10] M. Huang, Y. Zhang, and G. Hou, "Classical capacity of a quantum multiple-access channel," *Phys. Rev. A*, vol. 62, 2000.

[11] L. B. Levitin, "Quantum generalization of conditional entropy and information," in *Quantum Computing and Quantum Communications (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1999, vol. 1509.

[12] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton Univ. Press, 1955.

[13] M. Ohya and D. Petz, *Quantum Entropy and Its Use*. Berlin, Germany: Springer-Verlag, 1993.

[14] M. Ohya, "Some aspects of quantum information theory and their application to irreversible processes," *Rep. Math. Phys.*, vol. 27, no. 2, pp. 19–47, 1989.

[15] B. Schumacher, "Quantum coding," *Phys. Rev. A*, vol. 51, no. 4, pp. 2738–2747, 1995.

[16] B. Schumacher and W. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131–138, 1997.

[17] C. E. Shannon, "Two-way communication channels," in *Proc. Fourth Berkeley Symp. Probability and Statistics*, J. Neyman, Ed: Berkeley, 1961, pp. 611–644. Reprinted in *Claude Elwood Shannon Colleceted Papers*, N. J. A. Sloane and A. D. Wyner, Eds. New York: IEEE Press, pp. 351–384, 1993.

[18] A. Winter, "Coding theorems of quantum information theory," Ph.D. dissertation, Univ. Bielefeld, http://archiv.ub.uni-bielefeld.de/disshabi/mathe.htm. Also as e-print http://arXiv.org/abs/quant-ph/9907077, Germany, 1999.

[19] ——, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2481–2485, Nov. 1999.

Nonbinary Quantum Stabilizer Codes

Alexei Ashikhmin, *Member, IEEE,* and Emanuel Knill

***Abstract*—We define and show how to construct nonbinary quantum stabilizer codes. Our approach is based on nonbinary error bases. It generalizes the relationship between self-orthogonal codes over $\boldsymbol{F}_4$ and binary quantum codes to one between self-orthogonal codes over $\boldsymbol{F}_{q^2}$ and q-ary quantum codes for any prime power q.**

***Index Terms*—Nonbinary quantum codes, quantum stabilizer codes, self-orthogonal codes.**

I. Introduction

Probably the most important class of binary quantum codes are quantum stabilizer codes. They play a role similar to the linear codes in classical coding theory. Quantum stabilizer codes have simple encoding algorithms, can be analyzed using classical coding theory, and yield methods for fault tolerant quantum computation. The first examples of quantum codes found by Shor [20], and Steane [24], [25] were quantum stabilizer codes. General quantum stabilizer codes were introduced by Gottesman [10] and Calderbank *et al.* [6]. Later, Calderbank *et al.* [7] gave the now standard connection between quantum stabilizer codes and classical self-orthogonal codes, which was used to construct a number of new good quantum codes.

While the theory of binary quantum stabilizer codes is now well developed, nonbinary codes have been relatively ignored. A connection between classical codes over $\boldsymbol{Z}_n$ and quantum codes is given in [13], [14]. The connection is based on a stabilizer construction derived from so-called nice error bases. Rains [18] suggested a number of constructions of p-ary (p prime) quantum stabilizer codes generalizing the $\boldsymbol{F}_4$ constructions for binary quantum codes. This left open the problem of establishing a similar relationship between classical codes over $\boldsymbol{F}_{p^m}$ and quantum codes. It is known that if C is a $\boldsymbol{Z}_{p^m}$-linear code then there exists a linear code D over $\boldsymbol{F}_{p^m}$ of the same length and size and with the same or larger minimum distance. Since the minimum distance of a quantum stabilizer code is strongly related to the minimum distance of the corresponding classical code, it is natural to try to construct nonbinary quantum codes on the basis of classical codes over $\boldsymbol{F}_{p^m}$.

In the present correspondence, we give a general method for constructing p^m-ary quantum codes from classical self-orthogonal codes over $\boldsymbol{F}_{p^m}$ and $\boldsymbol{F}_{p^{2m}}$. The notion of self-orthogonality is induced by an inner product of vectors over a finite field. Self-orthogonality has to be defined so that one can easily construct and analyze the classical codes. The notion of self-orthogonality used in our construction arises naturally from the error bases of [13], [14] and can be identified with that arising from a field-theoretically defined simplectic form. Good self-orthogonal codes with respect to this form have already been found by Bierbrauer and Edel [5], who generalized the properties of classical codes useful for quantum codes without giving a procedure for

Manuscript received August 22, 2000; revised June 6, 2001. The work of E. Knill was supported by the NSA and DOE. The material in this correspondence was presented in part at AMS 2001 Spring Eastern Section Meeting, Hoboken, NJ, April 28–29, 2001.

A. Ashikhmin is with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: aea@research.bell-labs.com).

E. Knill is with Los Alamos National Laboratory, Group CCS-3, Los Alamos, NM 87545 USA (e-mail: knill@lanl.gov).

Communicated by P. W. Shor, Associate Editor for Quantum Information Theory.

Publisher Item Identifier S 0018-9448(01)08956-8.

0018–9448/01$10.00 © 2001 IEEE

obtaining corresponding quantum codes. Together with our construction, the codes of [5] can be used to obtain good quantum codes.

Our construction connects self-orthogonal codes over $\boldsymbol{F}_{p^m}$ of length $2n$ to p^m-ary quantum codes of length n. Based on this construction, we show how to associate p^m-ary quantum codes to classical self-orthogonal codes over $\boldsymbol{F}_{p^{2m}}$ of the same length. This enables the application of powerful methods from classical coding theory for constructing and analyzing nonbinary quantum codes. In particular, we prove the existence of nonbinary quantum codes that meet the Gilbert–Varshamov bound.

II. Basic Definitions

We start with the basic notions of classical and quantum coding theory. Denote by $\boldsymbol{F}_{p^m}$ the Galois field of p^m elements, where p is a prime number and m is an integer. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ denote the elements of a basis of $\boldsymbol{F}_{p^m}$ over $\boldsymbol{F}_p$. We fix a nonzero $\boldsymbol{F}_p$-linear functional $\mathrm{tr}\colon \boldsymbol{F}_{p^m} \to \boldsymbol{F}_p$ (called a *trace function*). Thus, tr satisfies

$$\mathrm{tr}(a+b) = \mathrm{tr}(a) + \mathrm{tr}(b) \tag{1}$$

$$\mathrm{tr}(\alpha a) = \alpha\, \mathrm{tr}(a) \tag{2}$$

for all $a, b \in \boldsymbol{F}_{p^m}$, $\alpha \in \boldsymbol{F}_p$. Note that for $x \in \boldsymbol{F}_{p^m}$, $\mathrm{tr}_x(a) = \mathrm{tr}(xa)$ defines another trace function, and that all such functions can be obtained this way. The standard trace function is the one defined by viewing $\boldsymbol{F}_{p^m}$ as an extension of $\boldsymbol{F}_p$ and letting $\mathrm{tr}(a) = \sum_{i=0}^{m-1} a^{p^i}$, [17, Ch. 2.3]. From the definition of tr if follows that

$$|\{a \in \boldsymbol{F}_{p^m}\colon \mathrm{tr}(a) = c\}| = p^{m-1}, \qquad \text{for any } c \in \boldsymbol{F}_p. \tag{3}$$

Let q be a prime power. The following equality holds for any $a, b \in \boldsymbol{F}_{q^m}$:

$$(a+b)^q = a^q + b^q. \tag{4}$$

Let t divide m.

Definition 1: A classical $\boldsymbol{F}_{p^t}$-linear code C over a field $\boldsymbol{F}_{p^m}$ of length n and size $(p^t)^k$ is a k-dimensional $\boldsymbol{F}_{p^t}$-linear subspace of the space $\boldsymbol{F}_{p^m}^n$.

In other words, for any $\boldsymbol{a}, \boldsymbol{b}$ from C and any $\alpha, \beta \in \boldsymbol{F}_{p^t}$, the vector $\alpha\boldsymbol{a} + \beta\boldsymbol{b}$ is also from C. (If C is p^m-linear we just call it linear.) We say that C is an $[n, k]_{p^t}$ code and denote by $d(C) = d$, $\delta(C) = \delta = \frac{d}{n}$, and $R_C = R = \frac{1}{n}\log_{p^m}|C|$ its minimum distance, relative minimum distance, and rate, respectively.

Let $*$ be an $\boldsymbol{F}_{p^t}$-bilinear form (an *inner product*). A code C is *self-orthogonal* for $*$ if for all vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ from C, the following property holds:

$$\boldsymbol{a} * \boldsymbol{b} = 0. \tag{5}$$

The code $C^\perp = \{\boldsymbol{v}\colon \boldsymbol{v} * \boldsymbol{a} = 0 \text{ for } \forall\, \boldsymbol{a} \in C\}$ is called the dual code of C with respect to (5).

The weight distribution of a linear code C of length n is defined as follows:

$$A_i(C) = |\{\boldsymbol{v} \in C\colon \mathrm{wt}(\boldsymbol{v}) = i\}|, \qquad i = 0, \ldots, n. \tag{6}$$

The polynomial

$$A_C(x, y) = \sum_{i=0}^{n} A_i(C) x^{n-i} y^i$$

where x and y are formal variables, is called the weight enumerator of C.

Remark: For an introduction to the theory of Galois fields and classical codes see, e.g., [17].

We also need the following notion from representation theory.

Let S be a finite group.

Definition 2: A homomorphism μ from S into $\boldsymbol{C}$ is called a *linear character* of S.

In particular, a linear character μ satisfies that

$$\mu(s_1)\mu(s_2) = \mu(s_1 s_2), \qquad s_1, s_2 \in S.$$

The number of linear characters of a finite Abelian group equals its order. Linear characters are pairwise orthogonal, i.e., if μ and γ are linear characters of S then

$$\sum_{s \in S} \mu(s)\overline{\gamma}(s) = \delta_{\mu,\gamma}|S|. \tag{7}$$

Since all characters of an Abelian group are linear, we omit the adjective "linear" for such groups.

Remark: For an introduction to representation theory see, e.g., [12], [22].

Definition 3: A q-ary quantum code Q of length n and size K is a K-dimensional subspace of a q^n-dimensional Hilbert space.

The rate of Q is given by $R_Q = \frac{1}{n}\log_q K$.

A q^n-dimensional Hilbert space is identified with the n-fold tensor product of q-dimensional Hilbert spaces. The q-dimensional spaces are thought of as the state spaces of q*-ary systems* in the same way as the values 0 and 1 can be thought of as the possible states of a bit in a bit string. We identify the state spaces with the q-dimensional complex linear space $\boldsymbol{C}_q$. An important characteristic of a quantum code is its *minimum distance*. If a code has minimum distance d then it can detect any $d-1$ and correct any $\lfloor \frac{d-1}{2} \rfloor$ errors. As a result, it is desirable to keep d as large as possible. A strict definition of the minimum distance is given in the next section after introducing error bases.

Remark: For introductions to the theory of quantum error correcting codes see, e.g., [15], [11], [16]. For a reader with a background in classical coding theory the papers [1], [2] have brief introductions to the field.

III. Error Bases

A general quantum error of a p^m-ary quantum system is a linear operator, say e, acting on the space $\boldsymbol{C}_{p^m}$. If $\boldsymbol{v}$ is a state (a unit vector in the space) of the system, then the effect of error e is to transform it to the state $e\boldsymbol{v}$. It is well known from the general theory of quantum codes that if a code can correct a given set $\mathcal{E}$ of error operators, then it can correct the linear span of $\mathcal{E}$. For this reason, it is enough to confine ourselves to errors that form a basis of the vector space of linear operators acting on $\boldsymbol{C}_{p^m}$. Let linear operators $e_1, e_2, \ldots, e_{p^{2m}}$ form such a basis.

Let $\boldsymbol{v}$ represent a state of n p^m-ary systems. A general error operator that can alter $\boldsymbol{v}$ is a linear operator acting on the n-fold tensor product of $\boldsymbol{C}_{p^m}$. Let us consider error operators of the form

$$E = \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n \tag{8}$$

where $\sigma_i \in \{e_1, e_2, \ldots, e_{p^{2m}}\}$. Since they form a basis we can again confine ourselves to them.

It is always possible to determine operators $e_1, e_2, \ldots, e_{p^{2m}}$ in such a way that one of them, say e_1, is the identity operator I_{p^m}. Define the *weight* of E in (8) as

$$\mathrm{wt}(E) = |\{\sigma_i \neq I_{p^m}\}|. \tag{9}$$

In the depolarizing channel model of errors [4], the operators $e_1, e_2, e_3, \ldots$ satisfy $\mathrm{Tr}(e_i^\dagger e_j) = p^m \delta_{i,j}$, where Tr is the trace of linear operators. When transmitting a qubit through a depolarizing channel, the probability that it is untouched (i.e., affected by the identity operator) is $1-\rho$ and the probability that it is affected by $e_i,\ i>1$, is $\rho/(p^{2m}-1)$. The standard assumption is that $1-\rho > \rho/(p^{2m}-1)$. Thus, the probability that an error of weight at least w occurs decreases exponentially with weight, a feature common to most realistic error models [16]. This explains why it is desirable to correct or detect all error operators up to some given weight.

Let P be the orthogonal projection operator onto Q. It can be shown that (see, e.g., [13]) an error operator E is *detectable* by Q iff

$$PEP = c_E P \tag{10}$$

where c_E is a constant depending on E.

Definition 4: The largest integer d such that every error of weight $d-1$ or less can be detected by a code is called its minimum distance.

We now define an explicit error basis for p^m-ary quantum codes. Let T and R be the linear operators acting on the space $\boldsymbol{C}_p$ that are defined by the matrices with entries

$$T_{i,j} = \delta_{i,\,j-1 \bmod p} \quad \text{and} \quad R_{i,j} = \xi^i \delta_{i,j}$$

where $\xi = e^{\iota 2\pi/p}$, $\iota = \sqrt{-1}$, and the indexes range from 0 to $p-1$ [23]. Let us define an inner product for operators as follows:

$$\langle A,\, B\rangle = \mathrm{Tr}(A^\dagger B). \tag{11}$$

The operators $T^i R^j$ generate a discrete Heisenberg group. It is readily seen that they form an orthogonal operator basis [23]. A proof is given for completeness and to establish some identities needed later.

Proposition 1: Operators $T^i R^j$ form an orthogonal basis under inner product (11).

Proof: It is easy to check that

$$TR = \xi RT$$

and, therefore,

$$T^i R^j = \xi^{ij} R^j T^i \tag{12}$$

$$(T^i R^j)\,(T^k R^l) = \xi^{il-jk} (T^k R^l)\,(T^i R^j) \tag{13}$$

$$(T^i R^j)\,(T^k R^l) = \xi^{-jk} T^{i+k} R^{j+l}. \tag{14}$$

The Hermitian transposes of T^i and R^i are obtained by raising to the power $p-1$

$$(T^i)^\dagger = (T^i)^{p-1} \qquad (R^i)^\dagger = (R^i)^{p-1} \tag{15}$$

and

$$T^p = R^p = I_p. \tag{16}$$

Now using the above expressions we obtain

$$\mathrm{Tr}\left((T^i R^j)^\dagger (T^k R^l)\right) = \mathrm{Tr}\left(\xi^{-(k-i)\,(l-j)} T^{(k-i)} R^{(l-j)}\right).$$

Noting that $\mathrm{Tr}(T^i R^j) = p\delta_{i,0}\delta_{j,0}$, we finish the proof. □

From (14) and (16) it follows that for $p > 2$

$$(T^i R^j)^p = \xi^{-ij(1+2+\cdots+(p-1))} = I_p. \tag{17}$$

Let $a,\, b \in \boldsymbol{F}_{p^m}$. Using a basis of $\boldsymbol{F}_{p^m}$ over $\boldsymbol{F}_p$, we can write uniquely

$$a = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_m\alpha_m$$

$$b = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m$$

with the a_i and b_i in $\boldsymbol{F}_p$. Define

$$T_a R_b = (T^{a_1} \otimes T^{a_2} \otimes \cdots \otimes T^{a_m})(R^{b_1} \otimes R^{b_2} \otimes \cdots \otimes R^{b_m}).$$

The multiplication rules given in the proof of Proposition 1 can be generalized. Define

$$\langle a,\, b\rangle = \sum_{i=1}^m a_i b_i \in \boldsymbol{F}_p. \tag{18}$$

From (14) and the identity $(A\otimes B)(C\otimes D) = AC \otimes BD$ it follows that

$$(T_a R_b)\,(T_c R_d) = \xi^{-\langle b,\, c\rangle} T_{a+c} R_{b+d}. \tag{19}$$

Equations (13) and (18) yield

$$(T_a R_b)\,(T_c R_d) = \xi^{\langle a,\, d\rangle - \langle b,\, c\rangle} (T_c R_d)\,(T_a R_b). \tag{20}$$

Now, using the same arguments as in Proposition 1, one can see that the operators $T_a R_b$ form an orthogonal basis of unitary operators acting on $\boldsymbol{C}_{p^m}$.

IV. Nonbinary Stabilizer Codes

Let $\boldsymbol{a}^\dagger = (a^{(1)},\, a^{(2)},\, \ldots,\, a^{(n)})$, $\boldsymbol{b}^\dagger = (b^{(1)},\, b^{(2)},\, \ldots,\, b^{(n)})$ be vectors from the space $\boldsymbol{F}_{p^m}^n$. (Throughout this section we use superscripts to label the systems.) As discussed in the previous section, it is enough to consider the error operators given by

$$E_{\boldsymbol{a},\boldsymbol{b}} = T_{a^{(1)}} R_{b^{(1)}} \otimes T_{a^{(2)}} R_{b^{(2)}} \otimes \cdots \otimes T_{a^{(n)}} R_{b^{(n)}}. \tag{21}$$

The set of operators

$$\mathcal{E} = \{\xi^i E_{\boldsymbol{a},\boldsymbol{b}} | 0 \le i \le p-1\}$$

form a group of order p^{2mn+1}. The center $\mathcal{Z}$ of $\mathcal{E}$ is generated by ξI and, therefore, has order p. For vectors $\boldsymbol{a},\, \boldsymbol{d} \in \boldsymbol{F}_{p^m}^n$ define an inner product by

$$\langle \boldsymbol{a},\, \boldsymbol{d}\rangle = \sum_{i=1}^n \left\langle a^{(i)},\, d^{(i)}\right\rangle \tag{22}$$

where $\langle a^{(i)},\, d^{(i)}\rangle$ is defined in (18). It follows from (20) that

$$E_{\boldsymbol{a},\boldsymbol{b}} E_{\boldsymbol{c},\boldsymbol{d}} = \xi^{\langle \boldsymbol{a},\boldsymbol{d}\rangle - \langle \boldsymbol{b},\boldsymbol{c}\rangle} E_{\boldsymbol{c},\boldsymbol{d}} E_{\boldsymbol{a},\boldsymbol{b}}. \tag{23}$$

From (19) we have

$$E_{\boldsymbol{a},\boldsymbol{b}} E_{\boldsymbol{c},\boldsymbol{d}} = \xi^{-\langle \boldsymbol{b},\boldsymbol{c}\rangle} E_{\boldsymbol{a}+\boldsymbol{c},\,\boldsymbol{b}+\boldsymbol{d}}. \tag{24}$$

From (21) and (17) it follows that for any $\boldsymbol{a}$ and $\boldsymbol{b}$ and $p > 2$

$$E_{\boldsymbol{a},\boldsymbol{b}}^p = I_{p^{mn}}. \tag{25}$$

Quantum *stabilizer codes* are defined as joint eigenspaces of the operators of a commutative subgroup S of $\mathcal{E}$. Without loss of generality, assume that $\mathcal{Z} \subseteq S$. If this is not the case, extend S by $\mathcal{Z}$. The order of S is a power of p, $|S| = p^{r+1}$. Let μ be a linear character that satisfies the constraint

$$\mu(\xi I) = \xi. \tag{26}$$

The number of characters satisfying (26) is p^r. To see this, recall that the characters of S with pointwise multiplication form a group, say $\hat{S}$, which is isomorphic to S. Let $\hat{\mathcal{Z}}$ be the group of linear characters of $\mathcal{Z}$. The map, say ψ, which restricts a character of S to a character of $\mathcal{Z}$ is a group homomorphism $\hat{S} \to \hat{\mathcal{Z}}$. The kernel of ψ consists of the characters ξ which satisfy $\mu(\xi I) = 1$. Such characters lift uniquely to characters of $\overline{S} = S/\mathcal{Z}$, so there are at most p^r many. This implies that the homomorphism ψ is onto, with kernel of size exactly p^r. The desired characters come from one of the cosets of the kernel.

Definition 5: The quantum stabilizer code $Q_{S,\mu}$ is the eigenspace of S associated with μ, i.e.,

$$Q_{S,\mu} = \{\boldsymbol{v} \in \boldsymbol{C}^n_{p^m}: E\boldsymbol{v} = \mu(E)\boldsymbol{v} \text{ for all } E \in S\}.$$

For any character γ of S, define

$$P_\gamma = \frac{1}{|S|} \sum_{E \in S} \overline{\gamma}(E)E.$$

Then P_γ is a projector, and

$$P_\gamma P'_\gamma = P_\gamma \delta_{\gamma,\gamma'} \tag{27}$$

(see, for example, [22, Ch. 2.6]).

Theorem 2: The operator P_μ is the projector onto $Q_{S,\mu}$, and the dimension of $Q_{S,\mu}$ is p^{mn-r}.

Proof: Let $E' \in S$ and γ a character of S. Then

$$\begin{aligned} E'P_\gamma &= \frac{1}{|S|} \sum_{E \in S} \overline{\gamma}(E)E'E \\ &= \frac{1}{|S|} \sum_{E \in S} \overline{\gamma}\left((E')^\dagger E\right)E \\ &= \gamma(E')P_\gamma \end{aligned} \tag{28}$$

where the last equality uses linearity of γ. From (28) it follows that $E'(P_\gamma \boldsymbol{v}) = \gamma(E')(P_\gamma \boldsymbol{v})$, hence, the range of P_μ is contained in $Q_{S,\mu}$. From the definitions of $Q_{S,\mu}$, P_γ, and the orthogonality of characters it follows that

$$P_\gamma \boldsymbol{v} = \boldsymbol{v}\delta_{\gamma,\mu}, \qquad \boldsymbol{v} \in Q_{S,\mu}. \tag{29}$$

Hence, P_γ is the orthogonal projection onto $Q_{S,\gamma}$ and $\dim(Q_{S,\mu}) = \mathrm{Tr}(P_\mu)$. Since for $E \in \mathcal{E} \setminus \mathcal{Z}$, $\mathrm{Tr}E = 0$, we have

$$\begin{aligned} \mathrm{Tr}P_\mu &= \frac{1}{|S|} \sum_{i=0}^{p-1} \overline{\mu}(\xi^i I)\mathrm{Tr}(\xi^i I) \\ &= \frac{1}{p^{r+1}} \sum_{i=0}^{p-1} p^{mn} \\ &= p^{mn-r} \end{aligned}$$

which implies the theorem. □

We say that $Q_{S,\mu}$ is an $[[n, n - r/m]]_{p^m}$ quantum stabilizer code.

V. Connection With Classical Codes

Let us establish a connection between quantum stabilizer and classical self-orthogonal codes. Note that since the error basis is obtained as a tensor product of p-ary error bases, stabilizer codes can be viewed as p-ary stabilizer codes. This situation is essentially the same as for classical linear codes over $\boldsymbol{F}_{p^m}$. However, since the goal is to protect against errors on p^m-ary systems, we wish to usefully relate p^m-ary stabilizer codes to classical codes over $\boldsymbol{F}_{p^m}$ and $\boldsymbol{F}_{p^{2m}}$.

First we show how to construct a classical code from a quantum code. Let φ be an automorphism of the vector space $\boldsymbol{F}^m_p$. For $\boldsymbol{a} = (a^{(1)}, a^{(2)}, \ldots, a^{(n)}) \in \boldsymbol{F}^n_{p^m}$ we define

$$\varphi(\boldsymbol{a}) = (\varphi(a^{(1)}), \varphi(a^{(2)}), \ldots, \varphi(a^{(n)})).$$

Clearly, the set $C = \{(\boldsymbol{a}, \varphi^{-1}\boldsymbol{b}) | E_{\boldsymbol{a},\boldsymbol{b}} \in S\}$ is an $\boldsymbol{F}_p$-linear code of length $2n$ and size p^r. (For each coset of $\mathcal{Z}$ in S, exactly one of its members contributes to C.) Moreover, since all operators from S commute and because of (23), the following property holds for any two vectors $(\boldsymbol{a}, \boldsymbol{b})$ and $(\boldsymbol{a}', \boldsymbol{b}')$ from C

$$\langle \boldsymbol{a}, \varphi(\boldsymbol{b}')\rangle - \langle \boldsymbol{a}', \varphi(\boldsymbol{b})\rangle = 0. \tag{30}$$

Thus, C is self-orthogonal with respect to the inner product defined by

$$(\boldsymbol{a}, \boldsymbol{b}) * (\boldsymbol{a}', \boldsymbol{b}') = \langle \boldsymbol{a}, \varphi(\boldsymbol{b}')\rangle - \langle \boldsymbol{a}', \varphi(\boldsymbol{b})\rangle.$$

Later, we will choose φ to relate the inner product to the structure of $\boldsymbol{F}_{p^m}$.

The minimum distance of a stabilizer code defined by S is related to the classical minimum weight of $C^\perp \setminus C$, where $C^\perp$ is the dual code of C with respect to (30). Define the weight of $\boldsymbol{v} = (\boldsymbol{a}, \boldsymbol{b}) \in \boldsymbol{F}^{2n}_{p^m}$ as

$$\mathrm{wt}(\boldsymbol{v}) = \left|\left\{i: a^{(i)} \neq 0 \quad \text{or } b^{(i)} \neq 0\right\}\right|.$$

Theorem 3: The minimum distance of a stabilizer code $Q_{S,\mu}$ equals $\min\{\mathrm{wt}(\boldsymbol{v}): \boldsymbol{v} \in C^\perp \setminus C\}$.

Proof: Our proof generalizes the arguments of [6].

Denote by $S^\perp$ the group of operators in $\mathcal{E}$ that commute with all operators from S. Thus $S^\perp$ is given by

$$S^\perp = \{\xi^i E_{\boldsymbol{a},\boldsymbol{b}} | (\boldsymbol{a}, \varphi^{-1}\boldsymbol{b}) \in C^\perp\}.$$

The desired fact follows from the observation that $E' \in \mathcal{E}$ is detectable iff $E' \notin S^\perp \setminus S$. We consider three cases.

1) Let $E' \in S$. By (28)

$$P_\mu E' P_\mu = \mu(E')P_\mu$$

and hence E' is detectable.

2) Let $E' \notin S^\perp$. Let S_i, $0 \leq i < p$, be defined by $S_i = \{E \in S: E'E = \xi^i EE'\}$. Then from (23) it follows that $|S_i| = |S|/p$. Thus,

$$\begin{aligned} |S|P_\mu E' P_\mu &= \sum_{E \in S} \overline{\mu}(E)EE'P_\mu \\ &= E' \sum_{i=0}^{p-1} \sum_{E \in S_i} \xi^i \overline{\mu}(E)EP_\mu \\ &= E' \sum_{i=0}^{p-1} \sum_{E \in S_i} \xi^i P_\mu \\ &= E' \sum_{i=0}^{p-1} \xi^i P_\mu |S|/p \quad (31) \\ &= 0 \quad (32) \end{aligned}$$

where we used (28) in the third to last step. Again, E' is detectable.

3) Let $E' \in S^\perp \setminus S$. By taking T to be the commutative subgroup generated by S and E' and extending the character μ to T, a subcode Q' of Q is obtained corresponding to the extended character. The dimension of Q' is smaller by a factor of p, which implies that Q is not an eigenspace of E'. Since E' commutes with S, E' preserves Q. All of this implies that $P_\mu E' P_\mu$ is not proportional to P_μ. □

The inner product defined in (30) depends on the automorphism φ. The choice of φ is primarily one of convenience. We now standardize this choice to simplify the construction of large minimum-distance codes. With respect to our distinguished basis of $\boldsymbol{F}_{p^m}$, φ is given by an $m \times m$ matrix M over $\boldsymbol{F}_p$. Choose M by defining

$$M_{i,j} = \mathrm{tr}(\alpha_i \alpha_j).$$

With $a^T = (a_1, a_2, \ldots, a_m)$, $b^T = (b_1, b_2, \ldots, b_m) \in \boldsymbol{F}_{p^m}$, we compute

$$\begin{aligned} a^T M b &= \sum_{i=1}^m \sum_{j=1}^m a_i b_j \mathrm{tr}(\alpha_i \alpha_j) \\ &= \sum_{i=1}^m \sum_{j=1}^m \mathrm{tr}(a_i b_j \alpha_i \alpha_j) \\ &= \mathrm{tr}\left(\left(\sum_{i=1}^m a_i \alpha_i\right)\left(\sum_{i=1}^m b_i \alpha_i\right)\right) \\ &= \mathrm{tr}(ab) \end{aligned}$$

where the product in the trace is multiplication in $\boldsymbol{F}_{p^m}$. For vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\boldsymbol{F}_{p^m}^n$, let $\langle \boldsymbol{a}, \boldsymbol{b} \rangle_* = \sum_i a^{(i)} b^{(i)}$. With this choice of φ, C is, therefore, self-orthogonal with respect to the inner product defined by

$$(\boldsymbol{a}, \boldsymbol{b}) * (\boldsymbol{a}', \boldsymbol{b}') = \mathrm{tr}\left(\langle \boldsymbol{a}, \boldsymbol{b}' \rangle_* - \langle \boldsymbol{a}', \boldsymbol{b} \rangle_*\right). \tag{33}$$

We can now construct a quantum stabilizer code of size p^{mn-r} from a classical self-orthogonal code C, $|C| = p^r$. Let vectors $\boldsymbol{v}_i = (\boldsymbol{a}_i, \boldsymbol{b}_i)$, $0 \le i \le r-1$, form a basis of C over $\boldsymbol{F}_p$. Then the p^r operators $E_{\boldsymbol{a}_i, \varphi \boldsymbol{b}_i}$ together with $\xi I_{p^{mn}}$ generate a group of commuting operators of order p^{r+1}, which defines $[[n, n - r/m]]_{p^m}$ stabilizer codes with minimum distance $d = \min\{\mathrm{wt}(\boldsymbol{v}) \colon \boldsymbol{v} \in C^\perp \setminus C\}$.

In [5], a number of families of good classical codes that are self-orthogonal with respect to the inner product

$$(\boldsymbol{a}, \boldsymbol{b}) *_1 (\boldsymbol{a}', \boldsymbol{b}') = \langle \boldsymbol{a}, \boldsymbol{b}' \rangle_* - \langle \boldsymbol{a}', \boldsymbol{b} \rangle_* \tag{34}$$

were constructed. Since a code that is self-orthogonal with respect to (34) is also self-orthogonal with respect to (33), our results establish a previously missing connection between the classical codes defined in [5] and quantum codes. Thus, we already have many good nonbinary stabilizer codes. It is shown in [5] that for any $q = p^m$ we can obtain quantum stabilizer codes with parameters

$[[q^r, q^r - (r+2), 3]]_q$

$[[q^2+1, q^2-3, 3]]_q$ ($r \ge 2$ is any integer)

$[[(q^{r+2}-1)/(q^2-1),$
$(q^{r+2}-1)/(q^2-1) - (r+2), 3]]_q$ ($r \ge 2$ is any even integer)

$[[q^3(q^{r-1}-1)/(q^2-1),$
$q^3(q^{r-1}-1)/(q^2-1) - (r+2), 3]]_q$ ($r \ge 1$ is any odd integer)

and others.

Note that if C is $\boldsymbol{F}_{p^m}$-linear and if it is self-orthogonal with respect to (33) then it is automatically self-orthogonal with respect to (34). Indeed,, if $(\boldsymbol{a}, \boldsymbol{b}) \in C$ then $(\alpha \boldsymbol{a}, \alpha \boldsymbol{b}) \in C$, $\alpha \in \boldsymbol{F}_{p^m}$, and from (1), (2), and (33), it follows

$$\begin{aligned} (\alpha \boldsymbol{a}, \alpha \boldsymbol{b}) * (\boldsymbol{a}', \boldsymbol{b}') &= \mathrm{tr}(\alpha(\langle \boldsymbol{a}, \boldsymbol{b}' \rangle_* - \langle \boldsymbol{a}', \boldsymbol{b} \rangle_*)) \\ &= \mathrm{tr}(\alpha (\boldsymbol{a}, \boldsymbol{b}) *_1 (\boldsymbol{a}', \boldsymbol{b}')) \\ &= 0, \qquad \text{for all } \alpha \in \boldsymbol{F}_{p^m} \end{aligned}$$

and $\mathrm{tr}(\alpha x) = 0$ for all $\alpha \in \boldsymbol{F}_{p^m}$ implies $x = 0$. Since the above implication does not hold for general $\boldsymbol{F}_p$-linear codes, one expects to find better codes self-orthogonal with respect to (33) in this class.

To explicitly construct a family of good quantum codes we give a specific form of the relationship which takes us from classical to quantum codes. Let q be a prime power. Let $\gamma_0 \in \boldsymbol{F}_q$ and choose $\gamma \in \boldsymbol{F}_{q^2} \setminus \boldsymbol{F}_q$ so that $\gamma^q = -\gamma + \gamma_0$. Such a γ exists (except when $\gamma_0 = 0$ and q is even) since $\gamma \to \gamma^q + \gamma$ is the standard $\boldsymbol{F}_q$-linear trace function of $\boldsymbol{F}_{q^2}$ onto $\boldsymbol{F}_q$ and because of (3). Then $\{1, \gamma\}$ is an $\boldsymbol{F}_q$-linear basis of $\boldsymbol{F}_{q^2}$. Let D be a linear code of length n over $\boldsymbol{F}_{q^2}$ and $D^\perp$ be its dual code with respect to the inner product

$$\boldsymbol{v}\boldsymbol{w} = \sum_{i=1}^n v_i w_i^q; \qquad \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{F}_{q^2}^n. \tag{35}$$

Note that since w_i^q is the conjugate of w_i (as defined for quadratic extension fields [17, Ch. 4], (35) is the generalization of the inner product used in [7]. Also, if we take $\gamma_0 = 0$ than the conjugate takes the conventional form: if $\alpha = a_0 + a_1\gamma$, $a_0, a_1 \in \boldsymbol{F}_q$ then $\alpha^q = a_0 - a_1\gamma$.

Let C and $C^\perp$ be codes of length $2n$ over $\boldsymbol{F}_q$ obtained by expanding each symbol of D and $D^\perp$ (respectively) in the basis $\{1, \gamma\}$.

Theorem 4: If $D \subseteq D^\perp$ then $C \subseteq C^\perp$ and $C^\perp$ is the dual code of C with respect to inner product (33).

Proof: The first statement is obvious. Let us prove the second statement.

Let $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in D$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_n) \in D^\perp$. Let $v_i = v_i^{(1)} + \gamma v_i^{(2)}$ and $w_i = w_i^{(1)} + \gamma w_i^{(2)}$. Then

$$\begin{aligned} \boldsymbol{v}\boldsymbol{w} &= \sum_{i=1}^n v_i w_i^q \\ &= \sum_{i=1}^n \left(v_i^{(1)} + \gamma v_i^{(2)}\right)\left(w_i^{(1)} + (\gamma_0 - \gamma) w_i^{(2)}\right) \\ &= \sum_{i=1}^n v_i^{(1)} w_i^{(1)} + \gamma\left(v_i^{(2)} w_i^{(1)} - v_i^{(1)} w_i^{(2)}\right) \\ &\quad + \gamma_0 v_i^{(1)} w_i^{(2)} + \gamma(\gamma_0 - \gamma) v_i^{(2)} w_i^{(2)}. \end{aligned} \tag{36}$$

Since this expression is 0 and $\gamma(\gamma_0 - \gamma) = \gamma\gamma^q \in \boldsymbol{F}_q$, it follows that

$$\sum_{i=1}^n v_i^{(2)} w_i^{(1)} - v_i^{(1)} w_i^{(2)} = 0$$

and the assertion follows. □

From this theorem, Corollary 1 follows.

Corollary 1: Let D be an $[n, (n-k)/2]_{q^2}$ self-orthogonal code over $\boldsymbol{F}_{q^2}$ and let $d = \min\{\mathrm{wt}(\boldsymbol{v}) \colon \boldsymbol{v} \in D^\perp \setminus D\}$. Then there exists an $[[n, k]]_q$ quantum stabilizer code with minimum distance d.

VI. Good Nonbinary Codes Exist

Any quantum code Q has two weight enumerators with coefficients given by $B_i(Q)$ and $B_i^\perp(Q)$, respectively. If P is the orthogonal projection onto Q and K is the code's size then

$$B_i(Q) = \frac{1}{K^2} \sum_{\mathrm{wt}(E)=i} \mathrm{Tr}(EP)^2$$

and

$$B_i^\perp(Q) = \frac{1}{K} \sum_{\mathrm{wt}(E)=i} \mathrm{Tr}(EPEP).$$

If Q is a q-ary quantum code associated with linear codes D and $D^\perp$ then $B_i(Q) = A_i(D)$ and $B_i^\perp(Q) = A_i(D^\perp)$ where $A_i(D)$ and $A_i(D^\perp)$ are defined in (6) [21].

Weight distributions of classical and quantum codes are probably their most important characteristics. In particular, they allow one to estimate the probability of undetected error [2]. In the classical case, the weight distribution also allows one to estimate the probability of decoding error and we believe that this is the case for quantum codes as well.

Classical coding theory asserts that there exist r-ary linear codes of rate R such that

$$\frac{1}{n}\log_r A_i \le H_r(i) + R - 1 + o(1)$$

where

$$H_r(x) = -x\log_r(x) - (1-x)\log_r(1-x) + x\log_r(r-1).$$

Such a weight distribution is called binomial. This result in particular allows one to obtain the lower bound on the reliability function [9, Ch. 5] and to derive the Gilbert–Varshamov bound for minimum-distance d codes. The Gilbert–Varshamov bound says that there exist r-ary codes, with relative minimum distance $\delta = \frac{d}{n}$ and rate R such that

$$R \ge 1 - H_r(\delta). \tag{37}$$

In [2], the existence of binary quantum codes that have binomial weight distribution and, as a consequence, meet the quantum version of the Gilbert–Varshamov bound was proven. Here we extend this result for p-ary quantum codes. We will prove the existence of p-ary quantum codes Q such that

$$B_i^{\perp} \le H_p\left(\frac{i}{n}\right) + \frac{R_Q}{2} - \frac{1}{2}. \tag{38}$$

Hence, we obtain the quantum Gilbert–Varshamov bound

$$R_Q \ge 1 - 2H_{p^2}(\delta). \tag{39}$$

Let Q be an $[[n, k_Q]]_q$ quantum stabilizer code associated with linear codes D and $D^{\perp}$ over $\boldsymbol{F}_{q^2}$. Recall that $R_Q = \frac{k_Q}{n}$ is the rate of Q and $R_{D^{\perp}}$ is the rate of $D^{\perp}$. It is easy to see that

$$R_Q = 2R_{D^{\perp}} - 1. \tag{40}$$

Let $\mathcal{T}$ be the set of $[n, k]_{p^2}$ self-orthogonal codes and $\mathcal{T}^{\perp}$ be the set of corresponding dual codes.

In what follows we will need the following lemmas.

Lemma 5: Let $D^{\perp}$ be an $[n, k]_{p^2}$ code such that $D \subseteq D^{\perp}$. Then the number of self-orthogonal vectors in $D^{\perp}$ equals

$$\frac{1}{p}\left((p^2)^k + (p-1)(-p)^n\right).$$

See the proof of the Lemma in the Appendix.

We will say that a vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_n) \in \boldsymbol{F}_{p^2}^n$ is self-orthogonal if $\boldsymbol{aa} = \sum_{i=1}^n a_i a_i^p = 0$.

Lemma 6: Let D be a self-orthogonal code and $D^{\perp}$ be its dual. Let $\boldsymbol{a} \in D^{\perp}$. Then, if $\boldsymbol{a}$ is self-orthogonal (non self-orthogonal) then $(\boldsymbol{a}+\boldsymbol{c})$, $\boldsymbol{c} \in D$, is also self-orthogonal (non self-orthogonal).

Proof:

$$\begin{aligned}(\boldsymbol{a}+\boldsymbol{c})(\boldsymbol{a}+\boldsymbol{c}) &= \sum_{i=1}^n (a_i+c_i)(a_i+c_i)^p \\ &= \sum_{i=1}^n a_i a_i^p + a_i c_i^p + c_i a_i^p + c_i c_i^p \\ &= \sum_{i=1}^n a_i a_i^p. \end{aligned}$$

□

Lemma 7: The number of self-orthogonal codes that contain a self-orthogonal vector does not depend on the vector.

Proof: Straightforward generalization of the proof of Lemma 4 from [3]. □

Lemma 8: The number of self-orthogonal vectors $\boldsymbol{a} \in \boldsymbol{F}_{p^2}^n$ of weight t equals

$$(p+1)\binom{n}{t}\frac{(p-1)^t + (-1)^t(p-1)}{p}.$$

Proof: Let us consider the equation

$$b_1 + b_2 + \cdots + b_t = 0,\ b_i \in \boldsymbol{F}_{p^2} \setminus 0.$$

We will prove by induction that the number of solutions, say R_t, of the equation is

$$R_t = \frac{(p-1)^t + (-1)^t(p-1)}{p}. \tag{41}$$

Indeed, when $t = 1$ or 2 it is easy to check that it is the case. Now noting that R_t equals the number of sets $(b_1, b_2, \ldots, b_t)$, $b_i \in \boldsymbol{F}_{p^2} \setminus 0$ that do not satisfy the equation, we obtain $R_t = (p-1)^{t-1} - R_{t-1}$, and computing R_t for even and odd t we obtain (41). Taking into account that for a given $c \in \boldsymbol{F}_p$ there are exactly $p+1$ elements $a \in \boldsymbol{F}_{p^2}$ such that $a^{p+1} = c$, we finish the proof. □

Let $A_D(x, y) = \sum_{i=0}^n A_i(D)x^{n-i}y^i$ be the weight enumerator of a code D. The MacWilliams identities [17, Ch. 5.6, Theorem 13] state that

$$A_{D^{\perp}}(x, y) = \frac{1}{|D|}A_D(x + (p^2-1), x - y). \tag{42}$$

Let $N = |\mathcal{T}|$. Let us denote by $\tilde{A}(x, y)$ the average weight enumerator of a code from $\mathcal{T}$

$$\tilde{A}(x, y) = \frac{1}{N}\sum_{D\in\mathcal{T}}\sum_{i=0}^n A_i(D)x^{n-i}y^i = \sum_{i=0}^n \tilde{A}_i x^{n-i}y^i.$$

Let, similarly, $\tilde{A}^{\perp}(x, y)$ be the average weight enumerator of the family $\mathcal{T}^{\perp}$.

Theorem 9:

$$\begin{aligned}\tilde{A}(x, y) &= (1-\alpha)x^n + \frac{\alpha}{p}\left((x+(p^2-1)y)^n + (p-1)(x-(p+1)y)^n\right) \\ \tilde{A}^{\perp}(x, y) &= \frac{1}{p^{2k}}\Big[(1-\alpha)(x+(p^2-1)y)^n \\ &\quad + \frac{\alpha}{p}\left(p^{2n}x^n + (p-1)p^n((p+1)y-x)^n\right)\Big]\end{aligned}$$

where

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

Proof: Let L be the number of self-orthogonal codes in which a self-orthogonal vector $\boldsymbol{a}$ is contained. Taking into account the number of self-orthogonal vectors from Lemma 8 and the fact that by Lemma 7 L is a constant, we obtain

$$\begin{aligned}\tilde{A}(x, y) &= \frac{1}{N}\sum_{D\in\mathcal{T}}\sum_{i=0}^n A_i(D)x^{n-i}y^i \\ &= \frac{1}{N}\sum_{i=0}^n x^{n-i}y^i \sum_{\boldsymbol{a}:\,\boldsymbol{aa}=0}\ \sum_{\substack{D\in\mathcal{T}\\ \boldsymbol{a}\in D}} 1\, x^n \\ &\quad + \frac{L}{N}\sum_{i \text{ is even}} x^{n-i}y^i(p+1)^i\binom{n}{i}\frac{(p-1)^i + (p-1)}{p} \\ &\quad + \frac{L}{N}\sum_{i \text{ is odd}} x^{n-i}y^i(p+1)^i\binom{n}{i}\frac{(p-1)^i - (p-1)}{p}.\end{aligned}$$

After simple computations we obtain

$$\tilde{A}(x, y) = \left(1 - \frac{L}{N}\right) x^n + \frac{1}{p}\frac{L}{N}\left((x+(p^2-1)y)^n + (p-1)(x-(p+1)y)^n\right).$$

Now, using the boundary condition $\tilde{A}(1, 1) = p^{2k}$, we compute $\alpha = \frac{L}{N}$. With the help of MacWilliams identities (42) we compute $\tilde{A}^{\perp}(x, y)$ and finish the proof. □

Using the Chebyshev inequality, we obtain that in $\mathcal{T}^{\perp}$ there exist a code $D^{\perp}$ such that $A_i(D)^{\perp} \le n^2 \tilde{A}_i^{\perp}$. From this, after simple computations, it follows that there exists a code $D^{\perp}$ such that

$$\frac{1}{n}\log_{p^2}\tilde{A}_i(D^{\perp}) \le H_{p^2}\left(\frac{i}{n}\right) + R_{D^{\perp}} - 1 + o(n).$$

From this and (40) the bound (38) follows. The exponent $\frac{1}{n}\log_{p^2}\tilde{A}_i^{\perp}$ becomes negative when $i \le H_{p^2}^{-1}(1 - R_{D^{\perp}}) = \delta_{GV}(R_{D^{\perp}})$. Since $D^{\perp} \in \mathcal{T}^{\perp}$ is a linear code, all the coefficients of its weight distribution should be integers and hence $A_i(D^{\perp}) = 0, i \le \delta_{GV}(R_{D^{\perp}})$. Thus, we have the following corollary.

Corollary 2: There exists $[[n, k]]_{p^2}$ quantum stabilizer codes that meet Gilbert–Varshamov bound 39.

Remark: The proof given here can be generalized to the case of the q-ary Gilber-Varshamov for arbitrary prime powers q.

APPENDIX

Proof (Lemma 5): Let $\mu \in \boldsymbol{F}_{p^2}$, and $\xi = e^{2\pi i/p}$. Then $\chi(\mu) = \xi^{\mathrm{tr}(\mu)}$ is an additive character of $\boldsymbol{F}_{p^2}$.

Let α be a primitive element of $\boldsymbol{F}_{p^2}$. We define $\alpha^{\infty} = 0$. Let us denote by $\boldsymbol{t_u} = \boldsymbol{t} = (t_{\infty}, t_0, \ldots, t_{p^2-2})$ the composition of a vector $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \boldsymbol{F}_{p^2}^n$, that is, t_i is the number of components u_j equal to α^i. We denote by $A_{\boldsymbol{t}}$ and $A_{\boldsymbol{t}}^{\perp}$ the number of codewords of D and $D^{\perp}$, respectively, with the composition $\boldsymbol{t}$. From the standard arguments of the proof of MacWilliams identities for complete weight enumerators [17, Ch. 5.6, Theorem 10] and from the definition of duality (35) it follows that

$$\sum_{\boldsymbol{t}} A_{\boldsymbol{t}}^{\perp} z_{\infty}^{t_\infty} z_0^{t_0} \cdots z_{p^2-2}^{t_{p^2-2}} = \frac{1}{|D|}\sum_{\boldsymbol{t}} A_{\boldsymbol{t}} \prod_{s=\infty}^{p^2-2}\left[\sum_{i=\infty}^{p^2-2}\chi(\alpha^{sp}\alpha^i)z_i\right]^{t_s}. \quad (43)$$

We can rewrite the right-hand side of (43) as follows:

$$\frac{1}{|D|}\sum_{\boldsymbol{t}} A_{\boldsymbol{t}}\left[\sum_{i=\infty}^{p^2-2} z_i\right]^{t_\infty} \prod_{s=0}^{p^2-2}\left[z_\infty + \sum_{j=0}^{p-2}\sum_{i=0}^{p}\chi\left(\alpha^{sp}\alpha^{j+i(p-1)}\right)z_{j+i(p-1)}\right]^{t_s}.$$

Let us put $z_{j+i(p-1)} = y_j$, $0 \le i \le p$; $0 \le j \le p-2$, $y_\infty = z_\infty$, $r_j = \sum_{i=0}^{p} t_{j+i(p-1)}$, and $r_\infty = t_\infty$. We will say that $\boldsymbol{r} = (r_\infty, r_0, \ldots, r_{p-2})$ is the reduced composition of a vector. Then we have

$$\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}^{\perp} y_\infty^{r_\infty} y_0^{r_0}\cdots y_{p-2}^{r_{p-2}} = \frac{1}{|D|}\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}\left[y_\infty + \sum_{j=0}^{p-2} y_j(p+1)\right]^{r_\infty} \prod_{s=0}^{p-2}\left[y_\infty + \sum_{j=0}^{p-2} y_j \sum_{i=0}^{p}\chi\left(\alpha^s\alpha^{j+i(p-1)}\right)\right]^{r_s}. \quad (44)$$

Now, for $l = \infty, 0, \ldots, p-2$, let us put $y_j = \xi^{\alpha^{(j-l)(p+1)}}$. Then, for the left-hand side of (44) we have

$$\sum_{l=\infty}^{p-2}\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}^{\perp} y^{\infty} y^0 \cdots y^{r_{p-2}} = \sum_{\boldsymbol{r}} A_{\boldsymbol{r}}^{\perp}\left(1 + \sum_{l=0}^{p-2}\xi^{f(l)}\right) \quad (45)$$

where

$$f(l) = \alpha^{-l(p+1)}\sum_{j=0}^{p-2} r_j\alpha^{j(p+1)}.$$

If $\boldsymbol{r}$ is the reduced composition of a self-orthogonal vector then

$$\sum_{j=0}^{p-2} r_j\,\alpha^{j(p+1)} = 0. \quad (46)$$

From this it follows that (45) equals $p\sum A_{\boldsymbol{r}}^{\perp}$, where the sum runs over reduced compositions of all self-orthogonal vectors from $D^{\perp}$.

Now let us estimate the sum

$$\sum_{l=\infty}^{p-2} \text{right-hand side of (44).} \quad (47)$$

The first term in the sum, i.e., when $l = \infty$, equals

$$\frac{1}{|D|}\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}(p^2)^{r_\infty}0^{r_0}\cdots 0^{r_{p-2}} = \frac{1}{|D|}(p^2)^n. \quad (48)$$

For any $l \ge 0$, we have

$$y_\infty + \sum_{j=0}^{p-2} y_j(p+1) = 1 + (p+1)\sum_{j=0}^{p-2}\xi^{\alpha^{(j-l)(p+1)}} = 1 + (p+1)\sum_{j=1}^{p-1}\xi^j = -p. \quad (49)$$

Further, for any $l \ge 0$ and any $s \ge 0$ we have

$$y_\infty + \sum_{j=0}^{p-2} y_j\sum_{i=0}^{p}\chi\left(\alpha^{s+j+i(p-1)}\right) = 1 + \sum_{j=0}^{p-2}\sum_{i=0}^{p}\xi^{\mathrm{tr}\left(\alpha^{s+j+i(p-1)}\right)+\alpha^{(j-l)(p+1)}} = 1 + \sum_{j=0}^{p-2}\sum_{i=0}^{p}\xi^{\mathrm{tr}\left(\alpha^{s+j+i(p-1)}+\frac{p+1}{2}\alpha^{(j-l)(p+1)}\right)}.$$

Making the change of variables $t = s + j$ and noting that

$$\alpha^{i(p-1)(p+1)} = (\alpha^i)^{(p^2-1)} = 1$$

we obtain

$$1 + \sum_{t=0}^{p-2}\sum_{i=0}^{p}\xi^{\mathrm{tr}\left(\alpha^{t+i(p-1)}+\frac{p+1}{2}\frac{\alpha^{(t+i(p-1))(p+1)}}{\alpha^{(l+s)(p+1)}}\right)} = \sum_{\mu\in\boldsymbol{F}_{p^2}}\xi^{\mathrm{tr}\left(\mu+\frac{p+1}{2}\frac{1}{\alpha^{(l+s)(p+1)}}\mu^{p+1}\right)}.$$

From [8, Theorem 3] it follows that the last sum equals

$$(-p)\cdot\xi^{(p-1)\alpha^{(l+s)(p+1)}}. \quad (50)$$

Substituting (49), (48), and (50) into (47), and using the fact that D is self-orthogonal, we obtain

$$\sum_{l=\infty}^{p-2} \text{right-hand side of (44)}$$
$$= \frac{1}{|D|}\left[(p^2)^n + \sum_{l=0}^{p-2}\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}(-p)^{\sum_{s=\infty}^{p-2} r_s}\xi^{h(l)}\right]$$
$$\left(\text{where } h(l) = (p-1)\alpha^{l(p+1)}\sum_{s=0}^{p-2} r_s\alpha^{s(p+1)}\right)$$
$$= \frac{1}{|D|}\left[(p^2)^n + \sum_{l=0}^{p-2}\sum_{\boldsymbol{r}} A_{\boldsymbol{r}}(-p)^n\right]$$
$$= \frac{1}{|D|}(p^2)^n + (p-1)(-p)^n$$

and the assertion of the lemma follows. □

REFERENCES

[1] A. Ashikhmin and S. Litsyn, "Upper bounds of the size of quantum codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1205–1215, May 1999.

[2] A. Ashikhmin, A. Barg, E. Knill, and S. Litsyn, "Quantum error detection I: Statement of the problem," *IEEE Trans. Inform. Theory*, pp. 778–788, May 2000.

[3] ——, "Quantum error detection II: Lower and upper bounds," *IEEE Trans. Inform. Theory*, pp. 789–800, May 2000.

[4] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed state entanglement and quantum error-correcting codes," *Phys. Rev. A*, vol. 54, p. 3824, 1996.

[5] J. Bierbrauer and Y. Edel. (1998) Quantum twisted codes. Preprint. [Online]. Available: http://www.math.mtu.edu/~jbierbra/

[6] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction and orthogonal geometry," *Phys. Rev. Lett*, vol. 78, pp. 405–409, 1997.

[7] ——, "Quantum errors correction via codes over GF (4)," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1369–1387, July 1998.

[8] L. Carlitz, "Evaluation of some exponential sums over a finite field," *Math. Nachrichtentech.*, vol. 96, pp. 13–20, 1980.

[9] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[10] D. Gottesman, "A class of quantum error-correcting codes saturating the quantum Hamming bound," *Phys. Rev. A*, vol. 54, pp. 1862–1868, 1996.

[11] ——, "Stabilizer codes and quantum error correction," Ph.D. dissertation, Calif. Inst, Technol., Pasadena, CA, 1997.

[12] G. James and M. Liebeck, *Representation and Characters of Groups*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[13] E. Knill, "Non-binary unitary error bases and quantum codes," LANL Preprint, quant-ph/9608048, 1996.

[14] ——, "Group representations, error bases and quantum codes," LANL Preprint, quant-ph/9608049, 1996.

[15] E. Knill and R. Laflamme, "A theory of quantum error correcting codes," *Phys. Rev. A*, vol. 55, pp. 900–911, 1997.

[16] E. Knill, R. Laflamme, and L. Viola, "Theory of quantum error correction for general noise," *Phys. Rev. Lett.*, vol. 84, pp. 2525–2528, 2000.

[17] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[18] E. Rains, "Nonbinary quantum codes," LANL e-print, quant-ph/9703048.

[19] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," in *Proc. 35th Annu. Symp. Foundations of Computer Science*, S. Goldwasser, Ed. Los Alamitos, CA: IEEE Comput. Soc. Press, 1994, p. 124.

[20] ——, "Scheme for reducing decoherence in quantum memory," *Phys. Rev. A*, vol. 52, p. 2493, 1995.

[21] 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.

[22] J. P. Serre, *Linear Representation of Finite Groups*. Berlin, Germany: Springer-Verlag, 1977.

[23] J. Schwinger, "Unitary operator bases," *Proc. Nat. Acad. Sci.*, vol. 46, pp. 570–579, 1960.

[24] A. M. Steane, "Simple quantum error correcting codes," *Phys. Rev. Lett.*, vol. 77, pp. 793–797, 1996.

[25] ——, "Multiple particle interference and quantum error correction," *Proc. Roy. Soc. London A*, vol. 452, pp. 2551–2577, 1996.

The Worst Additive Noise Under a Covariance Constraint

Suhas N. Diggavi, *Member, IEEE*, and Thomas M. Cover, *Fellow, IEEE*

***Abstract*—The maximum entropy noise under a lag p autocorrelation constraint is known by Burg's theorem to be the pth order Gauss–Markov process satisfying these constraints. The question is, what is the worst additive noise for a communication channel given these constraints? Is it the maximum entropy noise?**

The problem becomes one of extremizing the mutual information over all noise processes with covariances satisfying the correlation constraints $R_0, \ldots, R_p$. For high signal powers, the worst additive noise is Gauss–Markov of order p as expected. But for low powers, the worst additive noise is Gaussian with a covariance matrix in a convex set which depends on the signal power.

***Index Terms*—Burg's theorem, mutual information game, worst additive noise.**

I. INTRODUCTION

This correspondence treats a simple problem. What is the noisiest noise under certain constraints? There are two possible contexts in which we might ask this question. One is, what is the noisiest random process satisfying, for example, a lag covariance constraint, $\mathbb{E}[Z_i Z_{i+k}] = R_k,\ k = 0, \ldots p$. Thus, we ask for the maximum entropy rate for such a process. It is well known from Burg's work [1], [2] that the maximum-entropy noise process under p lag constraints is the pth-order Gauss–Markov process satisfying these constraints, i.e., it is the process that has minimal dependency on the past given the covariance constraints.

Another context in which we might ask this question is for an additive noise channel $Y = X + Z$, where the noise Z has covariance constraints $R_0, \ldots, R_p$ and the signal X has a power constraint P. What is the worst possible additive noise subject to these constraints? We expect the answer to be the maximum-entropy noise, as in the first problem. Indeed, we find this is the case, but only when the signal power is high enough to fill the spectrum of the maximum-entropy noise (yielding a white noise sum).

Consider the channel

$$Y_k = X_k + Z_k \tag{1}$$

Manuscript received September 9, 1999; revised September 25, 2000. This work was supported in part by the National Science Foundation under Grant NSF CCR-9973134 and by ARMY (MURI) DAAD19-99-1-0252. The material in this correspondence was presented in part at the International Symposium on Information Theory (ISIT), Ulm, Germany, June 1997.

S. N. Diggavi is with AT&T Shannon Laboratories, Florham Park, New Jersey, NJ 07932 USA (e-mail: suhas@research.att.com).

T. M. Cover is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA (e-mail: cover@isl.stanford.edu).

Communicated by S. Shamai, Associate Editor for Shannon Theory.
Publisher Item Identifier S 0018-9448(01)08962-3.

0018–9448/01$10.00 © 2001 IEEE