[2] R. N. Daskalov and T. A. Gulliver, "New good quasi-cyclic ternary and quaternary linear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1647–1650, 1997.
[3] P. P. Greenough and R. Hill, "Optimal ternary quasi-cyclic codes," *Des., Codes Cryptogr.*, vol. 2, pp. 81–91, 1992.
[4] T. A. Gulliver and P. R. J. Östergård, "Improved bounds for ternary linear codes of dimension 7," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1377–1381, 1997.
[5] ——, "Improved bounds for ternary linear codes of dimension 8," *J. Heuristics*, to be published.
[6] T. A. Gulliver and V. K. Bhargava, "New good rate $(m-1)$/pm ternary and quaternary quasi-cyclic codes," *Des., Codes Cryptogr.*, vol. 7, pp. 223–233, 1996.
[7] R. Hill and P. P. Greenough, "Optimal quasi-twisted codes," in *Proc. Int. Workshop Algebraic and Combinatorial Coding Theory* (Voneshta Voda, Bulgaria, June 1992), pp. 92–97.
[8] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* New York: North-Holland, 1977.
[9] G. E. Séguin and G. Drolet, "The theory of 1-generator quasi-cyclic codes," Tech. Rep., Roy. Mil. College of Canada, Kingston, ON, Canada, 1991.

A Characterization of 1-Perfect Additive Codes

Joaquim Borges and Josep Rifà, *Member, IEEE*

***Abstract*—The characterization of perfect single error-correcting codes, or 1-perfect codes, has been an open question for a long time. Recently, Rifà has proved that a binary 1-perfect code can be viewed as a distance-compatible structure in F^n and a homomorphism θ: $F^n \to \Omega$, where Ω is a loop (a quasi-group with identity element). In this correspondence, we consider 1-perfect codes that are subgroups of F^n with a distance-compatible Abelian structure. We compute the set of admissible parameters and give a construction for each case. We also prove that two such codes are different if they have different parameters. The resulting codes are always systematic, and we prove their unicity. Therefore, we give a full characterization. Easy coding and decoding algorithms are also presented.**

***Index Terms*— Distance-compatible additive codes, perfect codes, translation-invariant propelinear codes.**

I. Introduction

Let F^n be the n-dimensional vector space over $\mathbb{Z}_2$. The *(Hamming) weight* $\mathrm{wt}\,(v)$ of a vector $v \in F^n$ is the number of nonzero coordinates of v. We define the *(Hamming) distance* between two vectors $v, u \in F^n$ as $d(v,u) = \mathrm{wt}\,(v+u)$. If $X \subset F^n$ and $v \in F^n$, we denote the *distance of* v *to* X as $d(v,X)$, that is,

$$d(v,X) = \min\{d(v,x) | x \in X\}.$$

We also define the sum $X+Y$ as the set of all vectors that can be expressed as the sum of a vector in X and a vector in Y. We write $X+x$ instead of $X+\{x\}$.

A (*binary*) *code* of length n is a subset of F^n. If this subset is a linear subspace, then the code is *linear*. If $\mathcal{C}$ is a code, then its elements are called *codewords*. We say that two codes, $\mathcal{C}_1$ and $\mathcal{C}_2$, are *isomorphic* if there exists a coordinate permutation π such that $\pi(\mathcal{C}_1) = \mathcal{C}_2$. Given two codes, $\mathcal{C}_1$ and $\mathcal{C}_2$, we say that $\mathcal{C}_1$ is a *translate* of $\mathcal{C}_2$ if there is a vector $v \in F^n$ such that $\mathcal{C}_1 = \mathcal{C}_2 + v$. Finally, we will say that two codes are *equivalent* if one of them is a translate of an isomorphic code to the other code. In this correspondence, we always consider codes containing the all-zero vector $\mathbf{0}$, unless stated otherwise.

The *code distance* in $\mathcal{C}$ is

$$d_{\mathcal{C}} = \min\{d(x,y) | x, y \in \mathcal{C}, x \neq y\}.$$

The *minimum weight* in $\mathcal{C}$ is

$$\mathrm{wt}_{\mathcal{C}} = \min\{\mathrm{wt}\,(x) | x \in \mathcal{C} \setminus \{\mathbf{0}\}\}.$$

A code $\mathcal{C}$ with $|C| = 2^m$ codewords is said to be *systematic* if there is a subset of m coordinate positions such that the restriction of $\mathcal{C}$ to these coordinates contains any vector of F^m.

A *perfect single error-correcting code* $\mathcal{C}$ *of length* $n \geq 3$ is a subset of F^n such that the code distance in $\mathcal{C}$ is 3, and $d(v,\mathcal{C}) \leq 1$, for every $v \in F^n$. Perfect single error-correcting codes exist exactly when $n = 2^t - 1$ for every positive integer $t > 1$, (see [10] or [16]). A perfect single error-correcting code is said to be a *1-perfect code*. For any $n = 2^t - 1 (t > 1)$, there exists exactly one 1-perfect linear code of length n, up to isomorphism, which is the well-known *Hamming code*.

In [10, p. 180], the problem of finding all 1-perfect codes is proposed. This question remains an open problem and appears to be quite challenging.

Let $e_0 = \mathbf{0}$ and e_i be the vector of weight 1 with the nonzero coordinate in the ith position, for $i = 1, \cdots, n$.

Let $*$ be a binary operation defined in F^n, such that $u * v \in F^n$ for $v, u \in F^n$. We consider the algebraic structure of F^n with such an operation.

The operation $*$: $F^n \times F^n \to F^n$ is said to be *distance-compatible* (see [12]) if for every $v \in F^n$ there exists a coordinate permutation π_v, such that

1) $v * e_i = v + e_{\pi_v(i)}$, for $i = 1, \cdots, n$;
2) $v * \mathbf{0} = \mathbf{0} * v = v$;
3) $v * e_i = w * e_i$ if and only if $w = v$.

If $(F^n, *)$ is a group and $*$ is distance-compatible, then we say that $(F^n, *)$ is a *distance-compatible group*.

We will say that $*$ is a *translation-invariant operation* if

$$d(x,y) = d(x * u, y * u) \qquad \forall x, y, u \in F^n$$

Definition 1: A *distance-compatible code* of length n is a subgroup of $(F^n, *)$, where $(F^n, *)$ is a distance-compatible group.

Definition 2: An *additive code* of length n is a subgroup of $(F^n, *)$, where $(F^n, *)$ is a distance-compatible Abelian group.

An additive code is defined in [3, p. 71] in terms of association schemes. The definition given in [3] requires that $*$ be a translation-invariant Abelian operation. In Section IV, we will prove that Definition 2 is equivalent to that of [3] in the case that the particular association scheme we consider is the Hamming scheme F^n. In the remainder of this correspondence, we will use Definition 2 for additive codes.

In [12], it is proved that any partition of F^n into 1-perfect codes can be viewed as a loop, that is, a quasi-group with identity

Manuscript received July 23, 1997; revised December 23, 1998. This work was supported in part by Spanish Grant TEL97-0663.

The authors are with the Department d'Informàtica, Universitat Autònoma de Barcelona, 08193-Bellaterra, Spain.

Communicated by T. Kløve, Associate Editor for Coding Theory.
Publisher Item Identifier S 0018-9448(99)04368-0.

element. Such a loop is a homomorphic image of $(F^n, *)$, where $*$ is a distance-compatible operation. Here, we will study the case where $(F^n, *)$ is a distance-compatible Abelian group. In particular, we focus our attention on 1-perfect additive codes and give a full characterization of such codes. These codes are isomorphic to subgroups of $\mathbb{Z}_2^{k_1} \oplus \mathbb{Z}_4^{k_2}$, where $k_1 + 2k_2 = n$. However, we study them as binary codes, but not as mixed group codes as in [7] and [9], where the Hamming distance is defined differently.

The correspondence is organized as follows. In Section II we give some basic definitions and properties about Steiner triple systems and propelinear codes. In Section III we study 1-perfect translation-invariant propelinear codes and prove they are additive codes using the characterization given in [14]. In Section IV we prove that any additive code (not necessarily 1-perfect code) is a translation-invariant propelinear code. In Section V we give the main results about existence, construction, unicity, and characterization of 1-perfect additive codes. Elementary coding and decoding algorithms are described in Section VI. Finally, we offer some conclusions in Section VII.

II. Preliminaries

A. Steiner Triple Systems

If S is a finite n-set with $n \geq 3$ and B is a collection of 3-subsets of S—called *blocks* or *triples*—such that every unordered pair of distinct elements of S is contained in exactly one triple, then B is called a *Steiner triple system* and denoted by $\mathrm{STS}(n)$. It is well known that an $\mathrm{STS}(n)$ exists if and only if $n \equiv 1$ or $3 \pmod 6$ (see [2] or [8]). If T is a k-subset of S, where $3 \leq k \leq n$, and D is a subset of B, such that D is a subcollection of triples of T and D is an $\mathrm{STS}(k)$, then D is called a *Steiner triple subsystem* of the system B. It is well known that, in this case, $2k + 1 \leq n$ or $k = n$ (see [4, p. 70], for instance).

Let $\mathcal{C}$ be a code in F^n containing the vector $\mathbf{0}$. It is easy to verify that if $\mathcal{C}$ is 1-perfect, then the minimum weight codewords (of weight 3) form an $\mathrm{STS}(n)$ if we identify the codewords with characteristic vectors of subsets of $\{1, \cdots, n\}$.

B. Propelinear Codes

The definitions and properties included in this subsection can be found in [13] and [14]. However, they have been included here in order to make the correspondence self-contained.

Let $\mathcal{S}_n$ denote the symmetric group of permutations of the set $\{1, 2, \cdots, n\}$. Let $\pi \in \mathcal{S}_n$. Then for any vector $v = (v_1, \cdots, v_n) \in F^n$, we write $\pi(v)$ to denote the vector $(v_{\pi^{-1}(1)}, \cdots, v_{\pi^{-1}(n)})$.

A code $\mathcal{C}$ of length n is said to be *propelinear* if for any codeword $x \in \mathcal{C}$ there exists $\pi_x \in \mathcal{S}_n$ satisfying the properties

1) $x + \pi_x(y) \in \mathcal{C}$ if and only if $y \in \mathcal{C}$.
2) $\pi_x \circ \pi_y = \pi_z \ \forall y \in \mathcal{C}$, where $z = x + \pi_x(y)$.

For all $x \in \mathcal{C}$ and for all $y \in F^n$, denote by $*$ the binary operation such that $x * y = x + \pi_x(y)$. Then, $(\mathcal{C}, *)$ is a group, which is not Abelian in general. The vector $\mathbf{0}$ is always a codeword and $\pi_{\mathbf{0}}$ is the identity permutation. Hence, $\mathbf{0}$ is the identity element in $\mathcal{C}$ and $x^{-1} = \pi_x^{-1}(x)$, for all $x \in \mathcal{C}$ (see [13]). Note that $\Pi = \{\pi_x | x \in \mathcal{C}\}$ is a subgroup of $\mathcal{S}_n$ with the usual composition of permutations as the multiplication. By $(\mathcal{C}, \Pi)$ we shall mean the set of all pairs (x, π_x), where $x \in \mathcal{C}$.

Clearly, propelinear code class is more general than the linear code class.

Lemma 3: Let $(\mathcal{C}, *)$ be a propelinear code. Then

i) $d(u, v) = d(x * u, x * v) \ \forall x \in \mathcal{C} \ \forall u, v \in F^n$.

ii) $d_{\mathcal{C}} = \mathrm{wt}_{\mathcal{C}}$, that is, the code distance coincides with the minimum weight.

Proof:

i) We compute $d(x * u, x * v)$

$$\begin{aligned} d(x * u, x * v) &= \mathrm{wt}\,(x + \pi_x(u) + x + \pi_x(v)) \\ &= \mathrm{wt}\,(\pi_x(u + v)) = \mathrm{wt}\,(u + v) = d(u, v). \end{aligned}$$

ii) Let $x, y \in \mathcal{C}$ be two codewords such that $d(x, y) = d_{\mathcal{C}}$. By i), we have that

$$d(x, y) = d(x^{-1} * x, x^{-1} * y) = d(\mathbf{0}, x^{-1} * y) = \mathrm{wt}(x^{-1} * y).$$

Hence $d_{\mathcal{C}} \geq \mathrm{wt}_{\mathcal{C}}$ because $x^{-1} * y \in \mathcal{C}$.

Now, let us assume that $z \in \mathcal{C}$ has weight $\mathrm{wt}_{\mathcal{C}}$. Again by i), we can write

$$wt\,(z) = d(\mathbf{0}, z) = d(x, x * z)$$

for any codeword $x \in \mathcal{C}$. Since $x * z \in \mathcal{C}$, we conclude that $\mathrm{wt}_{\mathcal{C}} \geq d_{\mathcal{C}}$ and the result holds. □

Definition 4: A propelinear code $(\mathcal{C}, *)$ is said to be *translation-invariant* if

$$d(x, y) = d(x * u, y * u), \qquad \forall x, y \in \mathcal{C} \quad \forall u \in F^n.$$

Lemma 5: A propelinear code $(\mathcal{C}, *)$ is translation-invariant if and only if

$$\mathrm{wt}\,(v) = d(x, v * x), \qquad \forall x \in F^n \quad \forall v \in \mathcal{C}$$

Proof: If $\mathcal{C}$ is translation-invariant and $v \in \mathcal{C}$, then

$$\mathrm{wt}\,(v) = d(\mathbf{0}, v) = d(x, v * x), \qquad \forall x \in F^n$$

by the definition of translation-invariant code.

Conversely, by Lemma 3, given two codewords $u, v \in \mathcal{C}$, we know that

$$d(u, v) = d(\mathbf{0}, u^{-1} * v) = \mathrm{wt}\,(u^{-1} * v)$$

and, by hypothesis,

$$wt\,(u^{-1} * v) = d(x, u^{-1} * v * x), \qquad \forall x \in F^n$$

which is $d(u * x, v * x)$, again by Lemma 3. □

Example 6: A $\mathbb{Z}_4$-*linear* code of length n is an additive subgroup $\mathcal{C}$ of $\mathbb{Z}_4^n$. We will also call $\mathbb{Z}_4$-linear the binary code of length $2n$ obtained from $\mathcal{C}$ using the *Gray map*

$$\phi(0) = (0, 0) \quad \phi(1) = (0, 1) \quad \phi(2) = (1, 1) \quad \phi(3) = (1, 0).$$

For a complete discussion on $\mathbb{Z}_4$-linear codes, see [6]. Given a binary $\mathbb{Z}_4$-linear code $\mathcal{C}$, we assign a permutation π_x to each codeword $x = (x_1, \cdots, x_{2n})$, such that $\pi_x = \pi_{12} \circ \pi_{34} \circ \cdots \circ \pi_{2n-1\,2n}$, where

$$\pi_{ij} = \begin{cases} Id, & \text{if } (x_i, x_j) = (0, 0) \text{ or } (1, 1) \\ (i, j), & \text{otherwise.} \end{cases}$$

Then, given two codewords, $x = (x_1, \cdots, x_{2n})$ and $y = (y_1, \cdots, y_{2n})$, the operation

$$\begin{aligned} x * y = (&\phi(\phi^{-1}(x_1, x_2) + \phi^{-1}(y_1, y_2)), \cdots, \\ &\phi(\phi^{-1}(x_{2n-1}, x_{2n}) + \phi^{-1}(y_{2n-1}, y_{2n}))) \end{aligned}$$

(where the addition is modulo 4) is the same as $x + \pi_x(y)$. In fact, as can be seen in [14], $\Pi = \{\pi_x | x \in \mathcal{C}\}$ is a subgroup of $\mathcal{S}_n$ and $\mathcal{C}$ is a translation-invariant propelinear code which is Abelian.

Example 7: Let $a = (1010)$ and $b = (1001)$ (where customary commas have been deleted). Put $\pi_a = (1,2)(3,4)$ and $\pi_b = (1,3)(2,4)$. The propelinear code generated by a and b will be called the *quaternion propelinear code*; it contains eight elements

$$(\mathcal{C}, \Pi) = \{(\mathbf{0}, Id), (a, \pi_a), (a^2, Id), (a^3, \pi_a), (b, \pi_b), (a*b, \pi_{a*b}), (a^2*b, \pi_b), (a^3*b, \pi_{a*b})\}$$

This code is group-isomorphic to the *quaternion group* $\mathcal{Q}_8$ on eight elements. As can be seen in [14], $\mathcal{C}$ is a translation-invariant propelinear code which is not Abelian. Note that the eight codewords obtained are the six codewords of weight 2, the all-zero vector, and the all-one vector.

Lemma 8: Let $(\mathcal{C}, *)$ be the quaternion propelinear code as in Example 7. If $x, y \in \mathcal{C}$ are such that $\mathrm{wt}(x) = \mathrm{wt}(y) = 2$ and $x + y = (1111)$, then $x * y = (0000)$.

Proof: The pairs of codewords verifying the hypothesis are

$$(1010, (1,2)(3,4)) \quad \text{and} \quad (0101, (1,2)(3,4))$$
$$(1100, (1,4)(2,3)) \quad \text{and} \quad (0011, (1,4)(2,3))$$
$$(1001, (1,3)(2,4)) \quad \text{and} \quad (0110, (1,3)(2,4))$$

Since these pairs of codewords are inverses, the result holds. □

In [14], it is shown that any translation-invariant propelinear code of length n can be viewed as a group that is isomorphic to a subgroup of $\mathbb{Z}_2^{k_1} \oplus \mathbb{Z}_4^{k_2} \oplus \mathcal{Q}_8^{k_3}$; where $k_1 + 2k_2 + 4k_3 = n$. This means that we can partition the set of coordinate positions $\{1, \cdots, n\}$ into three subsets, X, Y, and Z, such that

i) $|X| = k_1, |Y| = 2k_2, |Z| = 4k_3$;
ii) $\mathcal{C}_X$, that is, the set of codewords truncated to the X-coordinates, is a linear code. $\mathcal{C}_Y$ is a $\mathbb{Z}_4$-linear code, and $\mathcal{C}_Z$ is a *quaternionic code*, that is, there are k_3 4-subsets of Z such that the code, restricted to anyone of these subsets, is the quaternion code.

We will say that these codes are of *type* (k_1, k_2, k_3). Since $\mathcal{Q}_8$ is not Abelian, we deduce that translation-invariant propelinear Abelian codes must be of type $(k_1, k_2, 0)$. Obviously, if a propelinear code is of type $(k_1, 0, 0)$, then it is linear; and if it is of type $(0, k_2, 0)$, then it is $\mathbb{Z}_4$-linear.

III. 1-Perfect Translation-Invariant Propelinear Codes

There exist 1-perfect translation-invariant propelinear (not linear) codes of length $n = 2^t - 1$, for all $t > 3$. In [14], a family of such codes are constructed. The codes presented in [14] are of type $((n-1)/2, (n+1)/4, 0)$.

For the case $t = 3$, we already know that the Hamming code of length 7 has the linear structure of type $(7, 0, 0)$. Below we show that it has also a structure of type $(3, 2, 0)$ and a structure of type $(3, 0, 1)$.

Lemma 9: The Hamming code of length 7 has propelinear structures of type $(3, 2, 0)$ and $(3, 0, 1)$.

Proof: Let $a = (1010100)$, $b = (1001010)$, and $c = (1111111)$, with associated permutations $\pi_a = (1,2)(3,4)$, $\pi_b = (1,2)(3,4)$, and $\pi_c = Id$ (identity permutation of length 7). The reader may verify that the code generated by $\langle a, b, c\rangle$ is a 1-perfect code and is of type $(3, 2, 0)$. On the other hand, if we substitute π_b with $(1,3)(2,4)$, then $\langle a, b, c\rangle$ is also a 1-perfect code but of type $(3, 0, 1)$. □

The main aim of this section is to study which are the admissible values of k_1, k_2, and k_3 for 1-perfect translation-invariant propelinear codes, in addition to those of [14].

Of course, k_1 cannot be 0 since n must be odd.

Lemma 10: Let $\mathcal{C}$ be a 1-perfect translation-invariant propelinear code of length $n \geq 7$ and type (k_1, k_2, k_3). Then, $k_3 \leq 1$.

Proof: Let X, Y, and Z be defined as above. Suppose that $k_3 > 1$. Let $Z_1, \cdots, Z_{k_3} \subset Z$ be the 4-sets of coordinate positions of the $\mathcal{Q}_8$ part. Consider a vector $x \in F^n$ such that $\mathrm{wt}(x) = 2$ and x has a one in Z_i and the other in Z_j $(i, j \in \{1, \cdots, k_3\}, i \neq j)$. Then, there is a codeword $z \in \mathcal{C}$ such that $d(z, x) = 1$. Clearly, z must have two ones in the same coordinate positions as x, and $\mathrm{wt}(z)$ must be 3. Thus either $\mathrm{wt}(z_{Z_i}) = 1$, or $\mathrm{wt}(z_{Z_j}) = 1$, or both. In any case, this is a contradiction because $\mathrm{wt}(a_{Z_k})$ must be even for any codeword $a \in \mathcal{C}$ and for any $k \in \{1, \cdots, k_3\}$. This is true because the code restricted to the Z_k-coordinates is the quaternion propelinear code as in Example 7, for any $k \in \{1, \cdots, k_3\}$. □

Proposition 11: With the above notation, if $k_3 = 1$, then $k_2 = 0$.

Proof: Otherwise, let $x \in F^n$ be a vector such that $\mathrm{wt}(x) = 2$, $\mathrm{wt}(x_Y) = 1$, and $\mathrm{wt}(x_Z) = 1$. Clearly, there must be a codeword a at distance 1 from x and $\mathrm{wt}(a_Y) = 1$, $\mathrm{wt}(a_Z) = 2$. Now, let $y \in F^n$ be another vector of weight 2 such that $y_Y = x_Y$ and $d(y_Z, a_Z) = 3$. The codeword $b \in \mathcal{C}$ such that $d(b, y) = 1$ will be such that $b_Y = a_Y$, $\mathrm{wt}(b_Z) = 2$, and $d(b_Z, a_Z) = 4$, that is, the supports of b_Z and a_Z do not intersect (otherwise, $d(a, b) < 3$).

Now, if we compute $c = a * b$, we will have $\mathrm{wt}(c_Y) = 2$ and $\mathrm{wt}(c_Z) = 0$ (by Lemma 8), thus c is a codeword of weight 2, a contradiction. □

Theorem 12: Let $\mathcal{C}$ be a 1-perfect propelinear code of type (k_1, k_2, k_3) of length n. Then, if $k_3 > 0$, $\mathcal{C}$ is the Hamming code of length 7 with a structure of type $(3, 0, 1)$.

Proof: Since there is no codeword $a \in \mathcal{C}$ such that $\mathrm{wt}(a_Z) = 1$, any vector of weight 2 with the two ones in the X-coordinates must be at distance one from a codeword of weight 3 with the three ones in the X-coordinates. That is, the codewords of weight 3, with the three ones in the X-coordinates form an STS$(n-4)$ which is a subsystem, thus $2(n-4)+1 \leq n$, that is, $n \leq 7$. Hence $n = 7$. We have seen in Lemma 9 that the Hamming code of length 7 has, effectively, a structure of type $(3, 0, 1)$. □

The Hamming code of length 3 has only two codewords: the all-zero vector and the all-one vector. Clearly, this code has a propelinear structure of type $(1, 1, 0)$. Recall that if a propelinear code of length n is of type (k_1, k_2, k_3), then $n = k_1 + 2k_2 + 4k_3$. Thus by Lemma 9 and Theorem 12, we have that any 1-perfect translation-invariant propelinear code of length $n \geq 3$ has a structure of type $(k, (n-k)/2, 0)$, i.e., it has an Abelian structure and is isomorphic to a subgroup of $\mathbb{Z}_2^k \oplus \mathbb{Z}_4^{(n-k)/2}$.

Let $\mathcal{C}$ be a 1-perfect propelinear code of length n and type $(k, (n-k)/2, 0)$, and let $X, Y \subset \{1, \cdots, n\}$ be the k coordinates of the $\mathbb{Z}_2$ part, and the $(n-k)$ coordinates of the $\mathbb{Z}_4$ part, respectively. Without loss of generality, we may assume that $X = \{1, \cdots, k\}$. Then, any vector $a \in F^n$ may be written as $a = (a_X | a_Y)$ (where "$|$" denotes concatenation). If $a_Y = (a_Y^{(1)}, \cdots, a_Y^{(n-k)})$, we also assume that the coordinates are "well" placed, in the sense that the corresponding vector in $\mathbb{Z}_4^{(n-k)/2}$ is

$$\phi^{-1}(a_Y) = (\phi^{-1}(a_Y^{(1)}, a_Y^{(2)}), \cdots, \phi^{-1}(a_Y^{(n-k-1)}, a_Y^{(n-k)}))$$

where ϕ is the Gray map defined in Example 6. Now we can define the operation $*$ for any pair of vectors $v, u \in F^n$, considering F^n as the group $\mathbb{Z}_2^k \oplus \mathbb{Z}_4^{(n-k)/2}$.

If $v = (v_1, \cdots, v_n)$ and $u = (u_1, \cdots, u_n)$, then

$$u * v = \Big(v_1 + u_1, \cdots, v_k + u_k, \phi(\phi^{-1}(v_{k+1}, v_{k+2}) + \phi^{-1}(u_{k+1}, u_{k+2})), \cdots, \phi(\phi^{-1}(v_{n-1}, v_n) + \phi^{-1}(u_{n-1}, u_n))\Big)$$

where the addition is modulo 2 for the first k coordinates, and modulo 4 for the remaining $n-k$. Again, ϕ is the Gray map as in Example 6.

Therefore, any 1-perfect translation-invariant propelinear code is a subgroup of F^n, where F^n is a distance-compatible Abelian group. Consequently, such a code must be an additive code (see Definition 2).

IV. 1-Perfect Additive Codes

A. Additive Structure of F^n

Now, let us assume that $(F^n, *)$ is a distance-compatible group.

Lemma 13: If $x \in F^n$ has weight $t > 0$, then there is a sequence of vectors of weight 1, $e_{i_1}, \cdots, e_{i_t}$, such that $x = e_{i_1} * e_{i_2} * \cdots * e_{i_t}$.

Proof: We have that $d(x, x * e_i) = 1$, for all $i = 1, \cdots, n$, and $x * e_i \neq x * e_j$, for all $i, j = 1, \cdots, n$ such that $i \neq j$. Thus the set $\{x * e_i\}_{i=1}^n$ is the set of all vectors at distance one apart from x.

We proceed by induction on t.

- If $t = 1$, the claim is trivial.
- If $t > 1$, then there is at least one vector of weight 1, say e_j, such that $\mathrm{wt}\,(x * e_j) = t - 1$.

Put $y = x * e_j$. Then, by hypothesis of induction, there are vectors $e_{i_1}, \cdots, e_{i_{t-1}}$ such that $y = e_{i_1} * \cdots * e_{i_{t-1}}$. Hence, $x = e_{i_1} * \cdots * e_{i_{t-1}} * e_{i_t}$, where $e_{i_t} = e_j^{-1}$. □

Note that the sequence $e_{i_1}, \cdots, e_{i_t}$ may be nonunique.

Proposition 14: Let $u, v \in F^n$. Then, $d(v, v * u) = \mathrm{wt}\,(u)$.

Proof: We proceed by induction on $t = \mathrm{wt}\,(u)$.

- If $t \leq 1$, the claim is trivially true.
- If $t > 1$, then by the argument of the proof of Lemma 13, there is a vector x, such that $\mathrm{wt}\,(x) = t - 1$ and $u = x * e_i$, for some $i \in \{1, \cdots, n\}$. Then, $v * u = v * x * e_i$. By hypothesis of induction, $d(v, v * x) = t - 1$. Therefore, $d(v, v * x * e_i) = t$ or $t - 2$. In the second case, however, there would be a sequence $e_{i_1}, \cdots, e_{i_{t-2}}$, such that $v * x * e_i = v * e_{i_1} * \cdots * e_{i_{t-2}}$, implying that $x * e_i = e_{i_1} * \cdots * e_{i_{t-2}}$, which is not possible because $\mathrm{wt}\,(x * e_i) = t$. □

Corollary 15: If $(F^n, *)$ is a distance-compatible group, then

i) $d(u, v) = d(x * u, x * v)$, for all $x, u, v \in F^n$.

ii) The operation $*$ is translation-invariant if and only if $\mathrm{wt}\,(v) = d(x, v * x)$, for all $x, v \in F^n$.

Proof:

i) Let $z = u^{-1} * v$. Then, $d(u, v) = d(u, u * z)$. By Proposition 14, we have that $d(u, u * z) = \mathrm{wt}\,(z) = \mathrm{wt}\,(u^{-1} * v)$. But

$$\mathrm{wt}\,(u^{-1} * v) = d(x * u, x * u * u^{-1} * v) = d(x * u, x * v)$$

again by Proposition 14.

ii) If $*$ is translation-invariant, then

$$\mathrm{wt}\,(v) = d(\mathbf{0}, v) = d(x, v * x), \qquad \forall v, x \in F^n.$$

Conversely, by i), we have that for all $x, y \in F^n$

$$d(x, y) = d(\mathbf{0}, x^{-1} * y) = \mathrm{wt}\,(x^{-1} * y).$$

Since we are assuming that $\mathrm{wt}\,(z) = d(v, z * v)$, for all $v, z \in F^n$, we have

$$\mathrm{wt}\,(x^{-1} * y) = d(v, x^{-1} * y * v) = d(x * v, y * v), \qquad \forall v \in F^n$$

Thus $d(x, y) = d(x * v, y * v)$, for all $x, y, v \in F^n$. □

Theorem 16: The distance-compatible group $(F^n, *)$ is Abelian if and only if the operation $*$ is translation-invariant.

Proof: Suppose that $(F^n, *)$ is Abelian, then for all $u, v \in F^n$, we have $d(v, v * u) = \mathrm{wt}\,(u)$, by Proposition 14. But $d(v, v * u) = d(v, u * v) = \mathrm{wt}\,(u)$. Then, by Corollary 15, $*$ is translation-invariant.

Conversely, if $*$ is translation-invariant, then

$$d(e_j, e_i * e_j) = d(e_i, e_i * e_j) = 1, \qquad \forall i, j \in \{0, 1, \cdots, n\}.$$

Thus $e_i * e_j = e_i + e_j$ or $\mathbf{0}$. In any case, $e_i * e_j = e_j * e_i$, for all $i, j \in \{0, \cdots, n\}$.

Now, let x and y be vectors of weight t and s, respectively. Then, by Lemma 13, we can write $x = e_{i_1} * \cdots * e_{i_t}$ and $y = e_{j_1} * \cdots * e_{j_s}$. Thus

$$\begin{aligned} x * y &= e_{i_1} * \cdots * e_{i_t} * e_{j_1} * \cdots * e_{j_s} \\ &= e_{j_1} * \cdots * e_{j_s} * e_{i_1} * \cdots * e_{i_t} = y * x \end{aligned}$$

And hence, $*$ is commutative. □

Proposition 17: If $(F^n, *)$ is a distance-compatible Abelian group, then any vector in F^n (different from $\mathbf{0}$) has order 2 or 4.

Proof: Since $d(e_i, e_i * e_i) = 1$, for all $i = 1, \cdots, n$, we have that $e_i * e_i = \mathbf{0}$ or $e_i + e_j$, for some $j \in \{1, \cdots, n\}$, $j \neq i$. In the first case, we would have that e_i has order 2. For the second case, suppose that $e_j * e_i \neq \mathbf{0}$, then $e_j * e_i = e_j + e_i = e_i * e_i$, but this would imply that $i = j$. Thus e_i and e_j are inverses and $e_i * e_i = e_i + e_j = e_j * e_j$; hence $e_i^4 = \mathbf{0}$. Now, for any vector $x \in F^n$, we know that x can be expressed as $x = e_{i_1} * \cdots * e_{i_t}$, where $t = \mathrm{wt}\,(x)$. Therefore, $x^4 = \mathbf{0}$. On the other hand, there is no order 3 element, because if $x^3 = \mathbf{0}$, since $x^4 = \mathbf{0}$, the only possibility would be $x = \mathbf{0}$. □

Therefore, if $(F^n, *)$ is a distance-compatible group, we have proved that the following statements are equivalent.

- The group $(F^n, *)$ is Abelian.
- The operation $*$ is translation-invariant.
- F^n is isomorphic to $\mathbb{Z}_2^{k_1} \oplus \mathbb{Z}_4^{k_2}$, for some k_1, k_2, such that $k_1 + 2k_2 = n$.

We can conclude that if $(F^n, *)$ is a distance-compatible Abelian group, then any subgroup of $(F^n, *)$, that is, any additive code can be viewed (under group isomorphism) as a translation-invariant propelinear code of type $(k_1, k_2, 0)$. From now on, we will not distinguish between 1-perfect additive codes and 1-perfect translation-invariant propelinear codes.

B. The Quotient Group Ω

A *1-perfect partition* of F^n is a partition of F^n into 1-perfect codes. If $n = 2^t - 1 (t \geq 2)$, then a 1-perfect code of length n has 2^{n-t} codewords, and therefore, a 1-perfect partition will have $2^t = n + 1$ classes. Given a 1-perfect code $\mathcal{C}$, we can always define a 1-perfect partition as $\{\mathcal{C} + e_i\}_{i=0}^n$. This is a "natural" partition if $\mathcal{C}$ is a linear code because this partition is, in fact, the quotient group $F^n/\mathcal{C}$. However, it is possible to define other 1-perfect partitions, if we consider other translation-invariant operations defined in F^n.

Lemma 18: Let $\mathcal{C} \subset F^n$ be a 1-perfect code, where $(F^n, *)$ is a distance-compatible Abelian group, then

$$\Omega = \{\mathcal{C} = C_0, C_1, \cdots, C_n\} \quad \text{where}$$
$$C_i = \mathcal{C} * e_i = \{x * e_i\}_{x \in \mathcal{C}}, \qquad \forall i = 0, \cdots, n$$

is a 1-perfect partition.

Proof: Given two different codewords, $x, y \in \mathcal{C}$, it is clear that $x * e_i \neq y * e_i$, for all $i = 0, \cdots, n$. Hence, $|C_i| = |\mathcal{C}|$. Since $*$ is translation-invariant (see Theorem 16), $d(x * e_i, y * e_i) = d(x, y)$. Thus the code distance in C_i is 3 as in $\mathcal{C}$. Consequently, C_i is a 1-perfect code, for all $i = 0, \cdots, n$.

We could then define an operation $\cdot : \Omega \times \Omega \to \Omega$ such that

$$C_i \cdot C_j = C_k \Leftrightarrow e_i * e_j \in C_k, \qquad \forall i, j = 0, \cdots, n.$$

Note that if $e_i * e_j$ and $e_i * e_k$ are in the same class, then $j = k$; otherwise, $d(e_i * e_j, e_i * e_k) = 2$. Thus if $C_i \cdot C_j = C_i \cdot C_k$, then $C_j = C_k$. Also, if $C_i \cdot C_j = C_k \cdot C_j$, then $C_i = C_k$. Therefore, $(\Omega, \cdot)$ is at least a quasi-group.

Proposition 19: Let $(F^n, *)$ be a distance-compatible Abelian group, let $\mathcal{C} \subset F^n$ be a 1-perfect code, and let $C_i = \mathcal{C} * e_i$, for all $i = 0, \cdots, n$. The following statements will then be equivalent.

i) $\mathcal{C}$ is additive.
ii) $\forall v \in C_i \quad \forall x \in \mathcal{C}$, $x * v \in C_i$ $(i \in \{0, \cdots, n\})$.
iii) $\forall v \in C_i \quad \forall u \in C_j$, $v * u \in C_k$, where C_k is the class containing $e_i * e_j$.

Proof:

i) $\Rightarrow$ ii): For all $v \in C_i$, there is a codeword $y \in \mathcal{C}$, such that $y * e_i = v$. Then, $x * v = x * y * e_i$ and, since $x * y \in \mathcal{C}$, we have that $x * v \in C_i$.

ii) $\Rightarrow$ iii): If $v \in C_i$ and $u \in C_j$, then there are codewords $x, y \in \mathcal{C}$ such that $v = x * e_i$ and $u = y * e_j$. Thus $v * u = x * y * e_i * e_j$. By substituting $i = 0$ in ii), we obtain that $x * y \in \mathcal{C}$. Since $e_i * e_j \in C_k$, we have that $(x * y) * (e_i * e_j) \in C_k$ again by ii). Consequently, $v * u \in C_k$.

iii) $\Rightarrow$ i): Trivial, substituting $i = j = 0$. □

Corollary 20: Let $\mathcal{C}$ be a 1-perfect additive code and let Ω be the 1-perfect partition defined as in Lemma 18. Then, $(\Omega, \cdot)$ is also an Abelian group, where $\mathcal{C}$ is the zero element, and the nonzero elements have order 2 or 4.

Proof: In fact, Ω is the quotient group $F^n/\mathcal{C}$. The order of its elements is 2 or 4, because the order of the elements of F^n is 2 or 4 (see Proposition 17). □

Therefore, there are natural numbers α and β, not both zero, such that $\Omega \cong \mathbb{Z}_2^\alpha \oplus \mathbb{Z}_4^\beta$.

We can also state the following equivalence.

Theorem 21: Let $(F^n, *)$ be a distance-compatible Abelian group. Let $\mathcal{C} \subset F^n$ be a 1-perfect code and let Ω be the 1-perfect partition defined as in Lemma 18. The following statements are then equivalent.

i) $\mathcal{C}$ is an additive code.
ii) The map θ: $F^n \to \Omega$ such that for all $v \in F^n$, $\theta(v) = C_i$ if and only if $v \in C_i$ is a group homomorphism of F^n onto Ω.

Proof: Suppose that $\mathcal{C}$ is additive and let $u \in C_i$ and $v \in C_j$. By Proposition 19, $u * v \in C_i \cdot C_j$, hence $\theta(u * v) = C_i \cdot C_j$, which is $\theta(u) \cdot \theta(v)$. It is also true that $\theta(\mathbf{0}) = \mathcal{C}$. Therefore, θ is a group homomorphism of F^n onto Ω.

Conversely, if θ is a homomorphism of F^n onto Ω, then let $x, y \in \mathcal{C}$. Since $\mathcal{C} \cdot \mathcal{C} = \mathcal{C}$, we have $\theta(x) \cdot \theta(y) = \mathcal{C}$. Thus $\theta(x * y) = \mathcal{C}$ which implies $x * y \in \mathcal{C}$. Alternatively, and easier, $\mathcal{C}$ must be a subgroup of F^n, since $\mathcal{C} = \mathrm{Ker}\,\theta$, where θ is the natural projection of F^n onto the quotient group $F^n/\mathcal{C}$. □

V. The Full Characterization

A. Allowable Parameters

From now on, let $\mathcal{C}$ be a 1-perfect additive code or, equivalently, a 1-perfect translation-invariant propelinear code, of length n and type $(k, (n-k)/2, 0)$. Assume, as before, that $X = \{1, \cdots, k\}$ and $Y = \{k+1, \cdots, n\}$ are the k coordinates of the $\mathbb{Z}_2$ part, and the $(n-k)$ coordinates of the $\mathbb{Z}_4$ part, respectively. Without loss of generality, we also assume that the Y-coordinates are "well" placed as at the end of Section III.

We will search for admissible values of k.

Lemma 22: If $a \in \mathcal{C}$ is a codeword, then $\mathrm{wt}\,(a_Y) > 1$ or a_Y is the all-zero vector.

Proof: The claim is trivial, otherwise $a * a$ should be a codeword of weight 2. □

Proposition 23: Let $\mathcal{C}$ be a 1-perfect additive code of type $(k, (n-k/2), 0)$. If $n > 3$, then $k \geq 3$.

Proof: Of course, k must be odd (otherwise n would be even). Suppose that $k = 1$. If $n > 3$, then $n \geq 7$ (recall that $n = 2^t - 1$ for some nonnegative integer t). Let X and Y be as defined above. Let $z = (\mathbf{0}|z_Y)$ be a codeword of weight three, then $z * z$ will have weight six. Without loss of generality, take $z * z = (0\ 11\ 11\ 11 \cdots 00)$. Consider the vectors $a = (0\ 11\ 00 \cdots 00)$ and $b = (0\ 00\ 11\ 00 \cdots\ 00)$. These two vectors have weight two, hence they are not codewords and they must be at distance one from two codewords of weight three, respectively. These two codewords will be $x = (1\ 11\ 00 \cdots 00)$ and $y = (1\ 00\ 11\ 00 \cdots 00)$, otherwise $\mathrm{wt}\,(x * x) = \mathrm{wt}\,(y * y) = 2$. This, however, yields a contradiction, since $d(z * z, x * y) = 2$. □

Theorem 24: Let $(\mathcal{C}, *)$ be a 1-perfect additive code of type $(k, (n-k)/2, 0)$, where $n = 2^t - 1$ and $t \geq 3$. There will then be a natural number r, such that $2 \leq r \leq t \leq 2r$ and

i) $k = 2^r - 1$, i.e., $\mathcal{C}$ is of type $(2^r - 1, 2^{t-1} - 2^{r-1}, 0)$;
ii) $\Omega \cong \mathbb{Z}_2^{2r-t} \oplus \mathbb{Z}_4^{t-r}$, where Ω is the quotient group $F^n/\mathcal{C}$, defined in Section IV-B.

Proof: Let $\mathcal{C}' = \{a \in \mathcal{C} | \mathrm{wt}\,(a_Y) = 0\}$ and let

$$S = \{a \in \mathcal{C} | \mathrm{wt}\,(a_X) = 3, \mathrm{wt}\,(a_Y) = 0\}.$$

Any vector $b = (b_X|b_Y)$ with $\mathrm{wt}\,(b_X) = 2$ and $\mathrm{wt}\,(b_Y) = 0$ must be at distance 1 apart from a codeword $z \in \mathcal{C}$ of weight 3 and $\mathrm{wt}\,(z_X) \geq 2$. Since $\mathrm{wt}\,(z_Y) = 1$ is not possible by Lemma 22, we conclude that $\mathrm{wt}\,(z_X) = 3$. Thus S restricted to the X-coordinates is an STS$\,(k)$, which is a subsystem of an STS$\,(n)$. Hence, $n = k$ or $n \geq 2k + 1$. Moreover, the subcode $\mathcal{C}'$ is linear and S is the set of minimum-weight codewords of $\mathcal{C}'$. Since S is an STS, it follows that $\mathcal{C}'$ (restricted to the X coordinates) is a Hamming code of length k (see [1], for example). Thus $k = 2^r - 1$ for some $r > 1$ and $(n-k)/2 = 2^{t-1} - 2^{r-1}$, which is i).

Since $\Omega \cong \mathbb{Z}_2^\alpha \oplus \mathbb{Z}_4^\beta$ (see Corollary 20), we have that $|\Omega| = 2^\alpha 4^\beta = 2^{\alpha+2\beta}$. On the other hand, the quotient group Ω is a 1-perfect partition (see Lemma 18), and therefore the number of elements of Ω is $n + 1 = 2^t$. Thus $2^{\alpha+2\beta} = 2^t$, and therefore, $\alpha + 2\beta = t$.

Assuming, as always, that the first k coordinates are those of the $\mathbb{Z}_2$ part, we have that $e_i * e_i = \mathbf{0}$, if and only if $i \leq k$. So, $C_i \cdot C_i = \mathcal{C}$ if and only if $i \leq k$. Also, $\mathcal{C} \cdot \mathcal{C} = \mathcal{C}$. Thus the number of order 2 elements in Ω is $k + 1$. The number of order 2 elements in $\mathbb{Z}_2^\alpha \oplus \mathbb{Z}_4^\beta$ is $2^\alpha 2^\beta$. Hence, we have that $2^{\alpha+\beta} = k + 1 = 2^r$.

Hence $\alpha+\beta=r$ and $\alpha+2\beta=t$, so we obtain that $\alpha=2r-t$ and $\beta=t-r$, which is ii).

The inequalities $2\le r\le t\le 2r$ come from the fact that $k\ge 3$, $\beta\ge 0$, and $\alpha\ge 0$. □

In view of the last theorem, we have that all 1-perfect additive codes of length 15 must be of type $(15,0,0),(7,4,0)$ or $(3,6,0)$. Following the classification of the 80 nonisomorphic STS(15) given in [17], the Hamming code of length 15 has the STS(15) number 1. The example given in [14], of type $(7,4,0)$, has the STS(15) number 2. In Section VI-C, we will give an example of type $(3,6,0)$, whose STS(15) is the number 7.

Therefore, we have the following admissible parameters for 1-perfect additive codes:

t	r	Type: $(2^r-1,2^{t-1}-2^{r-1},0)$
2	2	$(1,1,0),(3,0,0)$
3	2,3	$(3,2,0),(7,0,0)$
4	2,3,4	$(3,6,0),(7,4,0),(15,0,0)$
5	3,4,5	$(7,12,0),(15,8,0),(31,0,0)$
6	3,4,5,6	$(7,28,0),(15,24,0),(31,16,0),(63,0,0)$
...	...	...

Note that two such codes of the same length could have the same set of codewords, up to coordinate permutation, but that they would have different algebraic structures. This is true for $n=3$ and for $n=7$; however, we will see that there are no more cases.

Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be 1-perfect additive codes of length n and of types $(k,(n-k)/2,0)$ and $(\ell,(n-\ell)/2,0)$, respectively. Suppose that there is a coordinate permutation $\sigma\in\mathcal{S}_n$ such that $\mathcal{C}_1=\sigma(\mathcal{C}_2)=\mathcal{C}$. Then, let $*$ be the operation such that $(\mathcal{C},*)$ is isomorphic to a subgroup of $\mathbb{Z}_2^k\oplus\mathbb{Z}_4^{(n-k)/2}$ and let $\perp$ be the operation such that $(\mathcal{C},\perp)$ is isomorphic to a subgroup of $\mathbb{Z}_2^\ell\oplus\mathbb{Z}_4^{(n-\ell)/2}$.

Lemma 25: With the above assumptions and notation, if $\ell<k$ and $n>7$, then $(\mathcal{C},\perp)$ has more than four coordinates in the $\mathbb{Z}_4$ part and at least four of such coordinates would be in the $\mathbb{Z}_2$ part of $(\mathcal{C},*)$.

Proof: We have that $n>7$, $\ell\ge 3$ and $n\ge 2\ell+1$, hence $n-\ell>4$. On the other hand, since $k+1$ and $\ell+1$ are powers of 2, and $\ell\ge 3$, we obtain that $k-\ell\ge 4$. □

Theorem 26: Let $\mathcal{C}$ be a 1-perfect additive code of types $(k,(n-k)/2,0)$ and $(\ell,(n-\ell)/2,0)$. Then, it is not possible that $\ell<k$ and $n>7$.

Proof: Assume that $\ell>k$ and $n>7$. Then, let us apply the last lemma. Let $A\subset\{1,\cdots,n\}$ be the set of coordinates in the $\mathbb{Z}_2$ part of $(\mathcal{C},*)$ and in the $\mathbb{Z}_4$ part of $(\mathcal{C},\perp)$.

Let us consider the case that there is no codeword of weight 3 with the three ones in the A-coordinates. Without loss of generality, we assume that the first four coordinates are in A. We also assume that the fifth and sixth coordinates are in the $\mathbb{Z}_4$ part of $(C,\perp)$ and $e_5\perp e_6=0$. Suppose that $e_1\perp e_3\ne 0$. Let $a=(10\,10\,00\,0\cdots0)$ and $x\in\mathcal{C}$ be the codeword such that $d(x,a)=1$. Then $x=(10\,10\,00\,0\cdots1\cdots0)$. Note that the third coordinate in x is in the $\mathbb{Z}_2$ part of $(\mathcal{C},*)$; otherwise, $x*x$ would be of weight 2. Thus this third coordinate is also in the $\mathbb{Z}_2$ part of $(\mathcal{C},\perp)$; otherwise, x would have the three ones in the A-coordinates. Now, let $b=(10\,00\,10\,0\cdots0)$ and let $y\in\mathcal{C}$ be the codeword such that $d(y,b)=1$. We examine the possibilities for y.

- If y has a one at the second, third, or fourth coordinate, then the fifth coordinate must be in the $\mathbb{Z}_2$ part of $(\mathcal{C},*)$; otherwise, $y*y$ would be of weight 2. But then, y would have the three ones in the A-coordinates.
- If $y=(10\,00\,11\,0\cdots0)$, then $\mathrm{wt}\,(y\perp y)=2$, a contradiction.
- Hence, the only possibility is $y=(10\;00\;10\;0\cdots1\cdots0)$, where the third coordinate is not in the same position as the third coordinate of x (otherwise, $d(x,y)=2$). Call j the coordinate position of the third coordinate of y. Now, we have the following possibilities:
  - If j is in the $\mathbb{Z}_2$ part of $(\mathcal{C},\perp)$, then it is easy to check that $d(x*y,x\perp y)=2$, a contradiction.
  - If j is in the $\mathbb{Z}_4$ part of $(\mathcal{C},\perp)$, then it cannot be the inverse coordinate of the first one (otherwise, $\mathrm{wt}\,(y\perp y)=2$). If j is neither the inverse coordinate of the third one, then $d(x*y,x\perp y)=2$ again.
  - Finally, if j is the inverse coordinate of the third one, then consider the vector $b'=(10\;00\;01\;0\cdots0)$. Now, the codeword y' such that $d(b',y')=1$ will be $y'=(10\,00\,01\,0\cdots1\cdots0)$, where the third one is neither the inverse coordinate of the third one, nor the inverse of the first one. Now, however, we have a contradiction because $d(x*y,x\perp y)=2$.

Now, consider the case that there is some codeword of weight 3 with the three ones in the A-coordinates. Let x be one such codeword. Without loss of generality, assume that $x=(10\;10\;10\;0\cdots0)$, where the first six coordinates are "well" placed in $(\mathcal{C},\perp)$. Note that the second, fourth, and sixth coordinates could be in A or not. Let $a=(10\;01\;00\;0\cdots0)$. The codeword $y\in\mathcal{C}$ such that $d(y,a)=1$ cannot have a one in the second coordinate and neither in the third one, otherwise, $\mathrm{wt}\,(y\perp y)=2$. If $y=(10\;01\;10\;0\cdots0)$, then $d(x,y)=2$; and if $y=(10\;01\;01\;0\cdots0)$, then $\mathrm{wt}\,(x\perp y)=2$. Therefore, if the third one of y is at the ith coordinate position, we have that $i>6$. If i is in the $\mathbb{Z}_2$ part of $(\mathcal{C},\perp)$, then $\mathrm{wt}\,(x\perp x\perp y\perp y)=2$. We conclude that i is in the $\mathbb{Z}_4$ part of $(\mathcal{C},\perp)$. Without loss of generality, we can assume that $y=(10\;01\;00\;10\;0\cdots0)$ and the eighth coordinate is the inverse of the seventh one in $(\mathcal{C},\perp)$. Now, we have

$$u=x\perp y=(11\;00\;10\;10\;0\cdots0)$$
$$v=u*x=(01\;10\;00\;10\;0\cdots0)$$
$$w=v\perp y=(00\;00\;00\;11\;0\cdots0)$$

Thus w is a codeword of weight 2, a contradiction. □

Do all these codes exist? Are they unique? In the next subsection we will give affirmative answers to both questions. Thus we will be able to conclude that the number of 1-perfect additive codes of length $n=2^t-1$ is $\lfloor(t+2)/2\rfloor$, for all $t>3$, and 1 for $t=2$ and $t=3$.

B. Existence, Construction, and Unicity

Let r and t be natural numbers such that $2\le r\le t\le 2r$. Consider F^n with the additive propelinear structure such that it is isomorphic to the group $\mathbb{Z}_2^{2^r-1}\oplus\mathbb{Z}_4^{2^{t-1}-2^{r-1}}$ and the coordinates are "well" placed as in Section III.

Recall (see Lemma 13) that any vector $x\in F^n$, of weight $\mathrm{wt}\,(x)=t$, can be expressed as

$$x=e_{i_1}*\cdots*e_{i_t}=\lambda_1e_1*\cdots*\lambda_ne_n$$

where λe_i stands for $e_i*\cdots*e_i$, λ times. In this representation of x, we have that $\lambda_i\in\{0,1\}$, for all $i=1,\cdots,2^r-1$, and $\lambda_i\in\{0,1,2\}$, for all $i=2^r,\cdots,n$. The representation can be nonunique; for example, the vector $e_{n-1}+e_n$ could be expressed as $e_{n-1}*e_{n-1}$ or e_n*e_n (if $r<t$). We call G the group $\mathbb{Z}_2^{2r-t}\oplus\mathbb{Z}_4^{t-r}$.

Since the number of order 2 nonzero elements of G is the same number of vectors e_i of order 2 in F^n, we can always define a map

$\vartheta\colon F^n \to G$, verifying:

$$\vartheta(e_i) \neq \vartheta(e_j) \qquad (i \neq j) \text{ and } \vartheta(\mathbf{0}) = \mathbf{0} \tag{1}$$

$$\vartheta(e_i) = -\vartheta(e_j) \Leftrightarrow e_i * e_j = \mathbf{0} \tag{2}$$

$$\vartheta(x) = \sum_{j=1}^{t} \vartheta(e_{i_j}), \qquad \text{for all } x = e_{i_1} * \cdots * e_{i_t} \text{ in } F^n. \tag{3}$$

We remark that (3) is equivalent to

$$\vartheta(x) = \sum_{i=1}^{n} \lambda_i \vartheta(e_i), \qquad \text{for all } x = \lambda_1 e_1 * \cdots * \lambda_n e_n \text{ in } F^n.$$

Note that ϑ is well-defined, that is, the image $\vartheta(x)$ does not depend on the selected representation of $x \in F^n$. The representation of the $2^r - 1$ coordinates of the $\mathbb{Z}_2$ part is unique

$$\lambda_1 e_1 * \cdots * \lambda_{2^r-1} e_{2^r-1} = \lambda_1 e_1 + \cdots + \lambda_{2^r-1} e_{2^r-1}$$

The only problem could be in the $\mathbb{Z}_4$ part. We take, for instance, the last two coordinates:

$$\begin{aligned} 00 &= 0e_{n-1} * 0e_n = 1e_{n-1} * 1e_n = 2e_{n-1} * 2e_n \\ 01 &= 0e_{n-1} * 1e_n = 1e_{n-1} * 2e_n \\ 10 &= 1e_{n-1} * 0e_n = 2e_{n-1} * 1e_n \\ 11 &= 2e_{n-1} * 0e_n = 0e_{n-1} * 2e_n. \end{aligned}$$

Since $e_{n-1} * e_n = \mathbf{0}$, we have that $\vartheta(e_{n-1}) = -\vartheta(e_n)$. For the first case, we have that $\vartheta(e_{n-1}) + \vartheta(e_n) = 2\vartheta(e_{n-1}) + 2\vartheta(e_n) = \mathbf{0}$. In the second case, we have $\vartheta(e_n) = \vartheta(e_{n-1}) + 2\vartheta(e_n)$. Similarly, for the third case. Finally, $2\vartheta(e_{n-1}) = -2\vartheta(e_n) = 2\vartheta(e_n)$. The last equality is due to the fact that all the elements of G have order 2 or 4.

Lemma 27: Let G be the group $\mathbb{Z}_2^{2r-t} \oplus \mathbb{Z}_4^{t-r}$. A map $\vartheta\colon F^n \to G$ verifying (1)–(3) is a homomorphism.

Proof: Let $x, y \in F^n$; we choose representations for both elements

$$x = \lambda_1 e_1 * \cdots * \lambda_n e_n \quad \text{and} \quad y = \mu_1 e_1 * \cdots * \mu_n e_n$$

Then

$$\begin{aligned} \vartheta(x * y) &= \vartheta((\lambda_1 + \mu_1)e_1 * \cdots * (\lambda_n + \mu_n)e_n) \\ &= \sum_{i=1}^{n} (\lambda_i + \mu_i)\vartheta(e_i) \\ &= \sum_{i=1}^{n} \lambda_i \vartheta(e_i) + \sum_{i=1}^{n} \mu_i \vartheta(e_i) = \vartheta(x) + \vartheta(y). \end{aligned}$$ □

The next theorem gives a method to find any 1-perfect additive code with admissible parameters. The proof can be found in [12] with a slighly different notation. We also include the proof here to make the correspondence self-contained.

Theorem 28: Let G and $\vartheta\colon F^n \to G$ be as in Lemma 27. Then $\mathcal{C} = \operatorname{Ker}\vartheta$ is a 1-perfect additive code of type $(2^r - 1, 2^{t-1} - 2^{r-1}, 0)$.

Proof: $\operatorname{Ker}\vartheta$ is a subgroup of F^n, hence $\mathcal{C}$ is an additive code of type

$$(2^r - 1, 2^{t-1} - 2^{r-1}, 0).$$

The number of codewords will be

$$|\operatorname{Ker}\vartheta| = \frac{|F^n|}{|G|} = \frac{2^n}{2^t} = 2^{n-t}$$

which is the number of codewords of any 1-perfect code of length $n = 2^t - 1$. In order to see that $\mathcal{C}$ is 1-perfect it suffices to see that the code distance in $\mathcal{C}$ is 3. By Lemma 3, it will be enough to check that the minimum weight is 3.

By the definition of ϑ, we have that $\vartheta(e_i) \neq \mathbf{0}$, $\forall i = 1, \cdots, n$. Thus there is no codeword of weight 1. Now, suppose there is a codeword x of weight 2. Then, $x = e_i * e_j$ for some $i, j = 1, \cdots, n$. We have that $\vartheta(x) = \mathbf{0} = \vartheta(e_i) + \vartheta(e_j)$. Thus $\vartheta(e_i) = -\vartheta(e_j)$, which implies $e_i * e_j = \mathbf{0}$, getting a contradiction since $e_i * e_j = x$. □

If we consider the 1-perfect partition

$$\Omega = F^n/\mathcal{C} = \{\mathcal{C} * e_i\}_{i=0}^{n}$$

with the operation $\cdot$, such that $C_i \cdot C_j = C_k$ if and only if $e_i * e_j \in C_k$, then $(\Omega, \cdot) \cong (G, +)$ and $C_i = \vartheta^{-1}(e_i)$, for all $i = 0, \cdots, n$. Thus ϑ is like the natural projection.

Now, we will see that the 1-perfect additive code constructed is unique, up to isomorphism, with the given parameters.

Corollary 29: For all r and t, such that $2 \leq r \leq t \leq 2r$, there is exactly one 1-perfect additive code of type $(2^r - 1, 2^{t-1} - 2^{r-1}, 0)$, up to isomorphism.

Proof: If $\mathcal{C}$ is a 1-perfect additive code of type $(2^r - 1, 2^{t-1} - 2^{r-1}, 0)$, we always can define $\vartheta\colon F^n \to G$ such that $\mathcal{C} = \operatorname{Ker}\vartheta$. Now, suppose that ϑ' is another homomorphism of F^n onto G, and let $\mathcal{C}' = \operatorname{Ker}\vartheta'$ be the associated code. We want to see that $\mathcal{C}$ and $\mathcal{C}'$ are isomorphic.

Let $X = \{1, \cdots, 2^r - 1\}$ and $Y = \{2^r, \cdots, n\}$. Suppose that $\mathcal{C}$ has its coordinates "well" placed, in the sense that the X-coordinates are those of the $\mathbb{Z}_2$ part, the Y-coordinates are those of the $\mathbb{Z}_4$ part, and $e_i * e_{i+1} = \mathbf{0}$ for all even i in Y. Clearly, there exists a permutation $\sigma \in \mathcal{S}_n$ such that $\vartheta'(e_i) = \vartheta(e_{\sigma(i)})$, for all $i = 1, \cdots, n$. Note that $\sigma(X) = X$ and $\sigma(Y) = Y$. Without loss of generality, we may also suppose that $\mathcal{C}'$ has its coordinates "well" placed as well. Let x be any vector of F^n, consider a representation $x = \lambda_1 e_1 * \cdots * \lambda_n e_n$. Then, we have

$$\begin{aligned} \vartheta'(x) &= \sum_{i=1}^{n} \lambda_i \vartheta'(e_i) = \sum_{i=1}^{n} \lambda_i \vartheta(e_{\sigma(i)}) \\ &= \sum_{i=1}^{n} \lambda_{\sigma^{-1}(i)} \vartheta(e_i) = \vartheta(\sigma^{-1}(x)) \end{aligned}$$

Therefore, $\vartheta'(x) = \mathbf{0}$ if and only if $\vartheta(\sigma^{-1}(x)) = \mathbf{0}$, hence $\mathcal{C} = \sigma(\mathcal{C}')$. □

A 1-perfect code $\mathcal{C}$ in F^n is of *full rank* if $\langle \mathcal{C} \rangle = F^n$, where $\langle \cdot \rangle$ denotes the linear span. We will see that any 1-perfect additive code is of type $P_2(n)$ in the sense of [5], or it is of full rank. A code $\mathcal{C}$ of length $n + 1 = 2^t$ is *extended 1-perfect* if it is obtained from a 1-perfect code of length n by adding an even or odd parity coordinate. As in [5], let E^n denote the set of all even-weight vectors of F^n. Let $C_0^0, \cdots, C_n^0$ and $C_0^1, \cdots, C_n^1$ be two partitions of E^{n+1} and $F^{n+1} \setminus E^{n+1}$, respectively, into extended 1-perfect codes. Let π be a permutation on the set $\{0, 1, \cdots, n\}$. Then, a construction due to Phelps [11] and Solov'eva [15] states that the code

$$\mathcal{C} = \{(c_0|c_1) | c_0 \in C_i^0, c_1 \in C_j^1, \pi(i) = j\}$$

is an extended 1-perfect code. Puncturing any coordinate of $\mathcal{C}$ gives a 1-perfect code of length $2n + 1$. The set of 1-perfect codes of length $2n + 1$ that can be obtained in this way is denoted by $P_2(2n + 1)$. We denote by $\mathcal{C}^\perp$ the subspace of F^n consisting of those vectors that are orthogonal to all codewords of $\mathcal{C}$.

Assume that $\mathcal{C}$ is not of full rank, then let $w \in \mathcal{C}^\perp$ of weight $\nu + 1 = (n+1)/2$ (all the nonzero vectors in $\mathcal{C}^\perp$ have this weight, see [5]), suppose, without loss of generality, that the nonzero coordinates of w are the first $\nu + 1$. Then, for all $u \in E^{\nu+1}$, define

$$T(u) = \{v \in F^\nu | (u|v) \in \mathcal{C}\}.$$

The sets $\{T(u)\}_{u\in F^\nu}$ are *noninterlaced* if $T(u)\cap T(u')=\emptyset$ or $T(u)=T(u')$, for all $u,u'\in E^{\nu+1}$.

In [5], it is proved that a 1-perfect code of length n is of type $P_2(n)$ if and only if the sets $\{T(u)\}_{u\in F^\nu}$ are noninterlaced for some $w\in\mathcal{C}^\perp$.

Theorem 30: Any 1-perfect additive code $\mathcal{C}$ of length n is of type $P_2(n)$ or it is of full rank.

Proof: Assume that $\mathcal{C}$ is not of full rank. It suffices to see that the sets $\{T(u)\}_{u\in F^\nu}$ are noninterlaced. Let ϑ be the homomorphism of F^n onto G such that $\mathcal{C}=\operatorname{Ker}\vartheta$. We have

$$\begin{aligned}T(u)&=\{v\in F^\nu|(u|v)\in\mathcal{C}\}\\&=\{v\in F^\nu|\vartheta(\mathbf{0}|v)=-\vartheta(u|\mathbf{0})\}.\end{aligned}$$

If $T(u)\cap T(u')\neq\emptyset$, then there is a vector $x\in F^\nu$ such that $(u|x)\in\mathcal{C}$ and $(u'|x)\in\mathcal{C}$, but then $\vartheta(\mathbf{0}|x)=-\vartheta(u|\mathbf{0})$ and $\vartheta(\mathbf{0}|x)=-\vartheta(u'|\mathbf{0})$, hence $\vartheta(u|\mathbf{0})=\vartheta(u'|\mathbf{0})$ that implies $T(u)=T(u')$. □

VI. Coding and Decoding

Designing a coding–decoding scheme is very simple using a 1-perfect additive code. As we are going to see, this comes from the fact that all these codes are systematic and we can easily compute a syndrome for any received vector.

Theorem 31: Any 1-perfect additive code is systematic.

Proof: Let $\mathcal{C}$ be a 1-perfect additive code of type $(2^r-1,2^{t-1}-2^{r-1},0)$. As before, let G be the group $\mathbb{Z}_2^{2r-t}\oplus\mathbb{Z}_4^{t-r}$, and let ϑ be a homomorphism of F^n onto G such that $\mathcal{C}=\operatorname{Ker}\vartheta$. Now, consider the set of elements $B\subset G$, $B=\{z_1,\cdots,z_r\}$, where

$$\begin{aligned}z_1&=(1,0,0,\cdots,0|0,\cdots,0)\\z_2&=(0,1,0,\cdots,0|0,\cdots,0)\\&\cdots\\z_{2r-t}&=(0,0,\cdots,0,1|0,\cdots,0)\\z_{2r-t+1}&=(0,\cdots,0|1,0,0,\cdots,0)\\z_{2r-t+2}&=(0,\cdots,0|0,1,0,\cdots,0)\\&\cdots\\z_r&=(0,\cdots,0|0,0,\cdots,0,1).\end{aligned}$$

The symbol "|" separates the binary and quaternary parts. It is easy to check that any element $u=(u_1,\cdots,u_r)\in G$ can be expressed as $u=\Sigma_{i=1}^r\mu_i z_i$, for some $\mu_1,\cdots,\mu_{2r-t}\in\{0,1\}$ and $\mu_{2r-t+1},\cdots,\mu_r\in\{0,1,2,3\}$. This representation is unique. In fact, we have that $\mu_i=u_i$, for all $i=1,\cdots,r$.

We reorder the coordinate positions of the vectors in F^n such that

$$\vartheta(e_{n-t+i})=z_i,\qquad\forall i=1,\cdots,2r-t$$

and

$$\begin{aligned}\vartheta(e_{n+2r-2t+1})&=3z_{2r-t+1}\\\vartheta(e_{n+2r-2t+2})&=z_{2r-t+1}\\\vartheta(e_{n+2r-2t+3})&=3z_{2r-t+2}\\\vartheta(e_{n+2r-2t+4})&=z_{2r-t+2}\\&\cdots\\\vartheta(e_{n-1})&=3z_r\\\vartheta(e_n)&=z_r.\end{aligned}$$

Now, let $x\in F^{n-t}$. We must prove that there exists $y\in F^t$ such that $a=(x|y)$ is a codeword.

Let $\mu_1,\cdots,\mu_r\in\{0,1,2,3\}$ be such that

$$\sum_{i=1}^r\mu_i z_i=-\vartheta(x|\mathbf{0}).$$

Define

$$\begin{aligned}(\mathbf{0}|y)=\mu_1e_{n-t+1}*\cdots*\mu_{2r-t}e_{n+2r-2t}*\mu_{2r-t+1}e_{n+2r-2t+2}\\*\mu_{2r-t+2}e_{n+2r-2t+4}*\cdots*\mu_re_n\end{aligned}$$

where $y\in F^t$. Then

$$\begin{aligned}\vartheta(a)=\vartheta(x|y)&=\vartheta(x|\mathbf{0})+\vartheta(\mathbf{0}|y)\\&=\vartheta(x|\mathbf{0})+\sum_{i=1}^{2r-t}\mu_i\vartheta(e_{n-t+i})\\&\quad+\sum_{i=1}^{t-r}\mu_{2r-t+i}\vartheta(e_{n+2r-2t+2i})\\&=\vartheta(x|\mathbf{0})+\sum_{i=1}^r\mu_iz_i=\vartheta(x|\mathbf{0})-\vartheta(x|\mathbf{0})=\mathbf{0}\end{aligned}$$

hence a is a codeword. □

We can use ϑ as a syndrome map, since for all $v\in F^n$

$$\vartheta(v)=\vartheta(e_i)\Leftrightarrow v\in C_i\Leftrightarrow\exists x\in\mathcal{C}\text{ such that }x*e_i=v$$

and $\vartheta(v)=\mathbf{0}$ if and only if $v\in\mathcal{C}$. Then, we only need an array with the images $\vartheta(e_i)$ for all $i=1,\cdots,n$; in order to perform the decoding process.

Suppose that we want to transmit with a coding–decoding scheme, using a 1-perfect additive code of type $(2^r-1,2^{t-1}-2^{r-1},0)$, where $n=2^t-1$. Note that the homomorphism ϑ and the code $\mathcal{C}$ are uniquely defined by the images $\vartheta(e_1),\cdots,\vartheta(e_n)$. Hence, we will mantain a static array with the elements $\vartheta(e_1),\cdots,\vartheta(e_n)$ of $\mathbb{Z}_2^{2r-t}\oplus\mathbb{Z}_4^{t-r}$, instead of generator and parity-check matrices as in a linear code. As we will see, this array (with n positions, each position being an r-vector) is the only static data we need for coding and decoding.

We will use the following arrangement of the elements of G:

1) The first elements will be all order 2 elements not specified at 3.
2) All order 4 elements not specified at 4. These elements will be arranged such that the first and the second are inverses, the third and the fourth, and so on.
3) The $2r-t$ elements with a "1" in the $\mathbb{Z}_2$ part and zeros elsewhere. For example, we choose the order

$$\begin{aligned}z_1&=(1,0,0,\cdots,0|0,\cdots,0)\\z_2&=(0,1,0,\cdots,0|0,\cdots,0)\\&\cdots\\z_{2r-t}&=(0,0,\cdots,0,1|0,\cdots,0).\end{aligned}$$

4) The elements

$$\begin{aligned}&(0,\cdots,0|3,0,0,\cdots,0)\\z_{2r-t+1}&=(0,\cdots,0|1,0,0,\cdots,0)\\&(0,\cdots,0|0,3,0,\cdots,0)\\z_{2r-t+2}&=(0,\cdots,0|0,1,0,\cdots,0)\\&\cdots\\&(0,\cdots,0|0,0,\cdots,0,3)\\z_r&=(0,\cdots,0|0,0,\cdots,0,1).\end{aligned}$$

Now, assume that this order corresponds to the images $\vartheta(e_1), \cdots, \vartheta(e_n)$. Thus the first $n-t$ coordinates will be the systematic ones. The first $\alpha = 2^r - 2r + t - 1$ are coordinates in the $\mathbb{Z}_2$ part, whereas the following $n - t - \alpha = 2^t + 2t - 2^r + 2r$ are in the $\mathbb{Z}_4$ part. For the last nonsystematic t coordinates, we have that those of the $\mathbb{Z}_2$ part are the first $2r - t$.

Note that given an element $x \in G$, there exist $\mu_1, \cdots, \mu_r$, such that $x = \Sigma_{i=1}^r \mu_i z_i$, where $\mu_i \in \{0, 1, 2, 3\}$, for all $i = 1, \cdots, r$. As in the proof of Theorem 31, the representation is unique.

Given a vector $a = (a_1, \cdots, a_{n-t})$, we consider the vector $(a|\mathbf{0}) \in F^n$ and a representation:

$$(a|\mathbf{0}) = \beta_1 e_1 * \cdots * \beta_n e_n.$$

A. Coding Algorithm

Step 0: Let $a = (a_1, \cdots, a_{n-t})$ be the binary information vector to encode.

Step 1: Compute $\vartheta(a|\mathbf{0})$

$$\vartheta(a|\mathbf{0}) = \sum_{i=1}^{n} \beta_i \vartheta(e_i).$$

Step 2: Compute $b = -\vartheta(a|\mathbf{0})$. Note that this can be done simply by replacing 1's with 3's (and *vice versa*) in the quaternary part.

Step 3: Find $\mu_1, \cdots, \mu_r$ such that $b = \Sigma_{i=1}^r \mu_i z_i$ (see proof of Theorem 31), and compute

$$\begin{aligned} c = {} & \mu_1 e_{n-t+1} * \cdots * \mu_{2r-t} e_{n+2r-2t} \\ & * \mu_{2r-t+1} e_{n+2r-2t+2} \\ & * \mu_{2r-t+2} e_{n+2r-2t+4} * \cdots * \mu_r e_n \end{aligned}$$

Since $\vartheta(a|\mathbf{0}) + \vartheta(\mathbf{0}|c) = \mathbf{0}$, $x = (a|c)$ is the codeword encoding a.

Now, suppose we transmit the codeword x. Since we are using a single error-correcting code, we want to find a decoding process that obtains x if the received vector y has less than two errors. That is, the received vector is $y = x * e_i$, for some $i \in \{0, 1, \cdots, n\}$.

B. Decoding Algorithm

Step 0: Suppose we have received the vector $y \in F^n$.

Step 1: Compute a representation for y, i.e., find $\lambda_1, \cdots, \lambda_n$, such that

$$y = \lambda_1 e_1 * \cdots * \lambda_n e_n.$$

Now, compute $\vartheta(y) = \Sigma_{i=1}^n \lambda_i \vartheta(e_i)$.

Step 2: If $\vartheta(y) = \mathbf{0}$, then y is a codeword. In this case, put $x = y$ and skip the following step.

Step 3: If $\vartheta(y) \neq \mathbf{0}$, then search in the array for the value $\vartheta(y)$. Let j be the array position where $\vartheta(y)$ is. Then, $\vartheta(y) = \vartheta(e_j)$. Compute $x = y * e_j^{-1}$ (if e_j is in the $\mathbb{Z}_2$ part, then $e_j^{-1} = e_j$; otherwise, $e_j^{-1} = e_{j+1}$ or $e_j^{-1} = e_{j-1}$). Since $\vartheta(x) = \vartheta(v) - \vartheta(e_j) = \mathbf{0}$, x is a codeword and it is at distance one apart from y.

Step 4: Retrieve the information vector a simply by taking the first $n - t$ coordinates of x.

C. Example

Suppose we want to use a 1-perfect additive code of type $(3, 6, 0)$. That is, $r = 2$, $t = 4$, and $n = 15$. Then, we arrange the elements of $G = \mathbb{Z}_4^2$

$$(0,2), (2,0), (2,2)$$

$$(3,3), (1,1), (3,1), (1,3), (3,2), (1,2), (2,3), (2,1)$$

$$(3,0), z_1 = (1,0), (0,3), z_2 = (0,1).$$

Note that, in this case, all the nonsystematic coordinates will be in the $\mathbb{Z}_4$ part.

Let $a = (1,0,1,1,0,0,1,1,0,1,1)$ be the information vector to encode.

1) We choose a representation for $(a|\mathbf{0})$

$$(a|\mathbf{0}) = e_1 * e_3 * e_4 * e_7 * e_8 * e_{10} * e_{10}$$

Then,

$$\begin{aligned} \vartheta(a|\mathbf{0}) &= \vartheta(e_1) + \vartheta(e_3) + \vartheta(e_4) + \vartheta(e_7) \\ &\quad + \vartheta(e_8) + 2\vartheta(e_{10}) \\ &= (0,2) + (2,2) + (3,3) + (1,3) \\ &\quad + (3,2) + (2,3) + (2,3) = (1,2). \end{aligned}$$

2) $b = -\vartheta(a|\mathbf{0}) = (3,2)$.

3) $b = (3,2) = 3z_1 + 2z_2$, then $c = 3e_{13} * 2e_{15} = (1,0,1,1)$. The codeword for a is

$$x = (a|c) = (1,0,1,1,0,0,1,1,0,1,1,1,0,1,1).$$

Now, suppose we receive the vector

$$y = (1,0,1,1,0,0,1,1,0,1,0,1,0,1,1).$$

1) A representation for y may be

$$y = e_1 * e_3 * e_4 * e_7 * e_8 * e_{10} * e_{12} * 2e_{14}.$$

Thus

$$\begin{aligned} \vartheta(y) &= \vartheta(e_1) + \vartheta(e_3) + \vartheta(e_4) + \vartheta(e_7) \\ &\quad + \vartheta(e_8) + \vartheta(e_{10}) + \vartheta(e_{12}) + 2\vartheta(e_{14}) \\ &= (0,2) + (2,2) + (3,3) + (1,3) \\ &\quad + (3,2) + (2,3) + (3,0) + (0,3) \\ &\quad + (0,3) = (2,1). \end{aligned}$$

2) The value $\vartheta(y) = (2,1)$ is at the 11th position in the array, hence $\vartheta(y) = \vartheta(e_{11})$. Thus the codeword at distance one apart from y is

$$\begin{aligned} x = y * e_{11}^{-1} = \\ y * e_{10} = (1,0,1,1,0,0,1,1,0,1,1,1,0,1,1). \end{aligned}$$

3) Finally, the information vector is retrieved by taking the first 11 coordinates $a = (1,0,1,1,0,0,1,1,0,1,1)$.

VII. Conclusions

The characterization of binary 1-perfect codes and 1-perfect partitions is a difficult and fascinating subject in coding theory. Using the results of [12] as a starting point, we have studied the class of 1-perfect additive codes, that is, 1-perfect codes that are subgroups of F^n with a distance-compatible Abelian group structure.

We have shown that any additive code can be viewed as a translation-invariant propelinear code. These codes have been extensively analyzed in [14]. For 1-perfect additive codes we have given a full characterization. We have seen that a 1-perfect additive code exists exactly when it is of type $(2^r - 1, 2^{t-1} - 2^{r-1}, 0)$ for any r and t such that $2 \leq r \leq t \leq 2r$. Moreover, we have given a construction for any r and t and have proved that the constructed code is unique, up to isomorphism. We have also shown that if two 1-perfect additive codes have length greater than 7, then they are isomorphic if and only if they have the same parameters r and t. On the other hand, we have proved that each of these codes is of type $P_2(n)$ (see [5]) or of full rank.

In addition, we have seen that these 1-perfect codes are systematic and that easy algorithms can be used for a coding–decoding scheme.

Future research should include the non-Abelian case. The study is also interesting if F^n is a quasi-group instead of a group. One of these cases was studied in [12].

ACKNOWLEDGMENT

The authors wish to thank the referees for their comments, which have proved to be helpful in enabling us to greatly improve the presentation of this correspondence. We would also like to thank Prof. I. J. Dejter for several corrections and remarks he made with regards to an earlier version.

REFERENCES

[1] E. F. Assmus and H. F. Mattson, "Coding and combinatorics," *SIAM Rev.*, vol. 16, pp. 349-388, 1974.
[2] T. Beth, D. Junickel, and H. Lenz, *Design Theory*. Manheim, Germany: Wiessenschaftverlag, 1985; Cambridge, U.K.: Cambridge Univ. Press, 1986.
[3] A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance Regular Graphs*. Berlin, Germany: Springer-Verlag, 1989.
[4] C. J. Colbourn and J. H. Dinitz, *The CRC Handbook of Combinatorial Designs*. Boca Raton, FL: CRC, 1996.
[5] T. Etzion and A. Vardy, "Perfect binary codes: Constructions, properties, and enumeration," *IEEE Trans. Inform. Theory*, vol. 40, pp. 754–763, 1994.
[6] A. R. Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The Z_4-linearity of kerdock, preparata, goethals and related codes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 301–319, 1994.
[7] O. Heden, "A new construction of group and nongroup perfect codes," *Inform. Contr.*, vol. 34, pp. 314–323, 1977.
[8] D. R. Hughes and F. C. Piper, *Design Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[9] B. Lindström, "Group partitions and mixed perfect codes," *Canad. Math. Bull.*, vol. 18, pp. 57–60, 1975.
[10] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.
[11] K. T. Phelps, "A combinatorial construction of perfect codes," *SIAM J. Alg. Discr. Meth.*, vol. 5, pp. 224–228, 1984.
[12] J. Rifà, "Well-ordered Steiner triple systems and 1-perfect partitions of the n-cube," *SIAM J. Discrete Mathematics*, 1998, to be published.
[13] J. Rifà, J. M. Basart, and L. Huguet, "On completely regular propelinear codes," in *Proc. 6th Int. Conf., AAECC-6, Lecture Notes in Computer Science*, vol. 375. Berlin, Germany: Springer-Verlag, 1989, pp. 341–355.
[14] J. Rifà and J. Pujol, "Translation invariant propelinear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 590–598, 1997.
[15] F. I. Solov'eva, "On binary nongroup codes," *Methodi Diskr. Analiza*, vol. 37, pp. 65–76, 1981, in Russian.
[16] A. Tietäväinen, "On the nonexistence of perfect codes over finite fields," *SIAM J. Appl. Math.*, vol. 24, pp. 88–96, 1973.
[17] H. S. White, F. N. Cole, and L. D. Cummings, "Complete classification of the triad systems of fifteen elements," *Mem. Math. Acad. Sci.*, USA, no. 14, 2nd memoir, pp. 1–89, 1919.

An Encoder to Match Reed–Solomon Codes Over GF(q) to a Subalphabet of GF(q)

Claude Le Dantec, *Member, IEEE*, and Philippe Piret, *Member, IEEE*

***Abstract*— One describes procedures to generate the codewords of a Reed–Solomon code over GF(2^m) having all their symbols in a GF(2)-subspace of GF(2^m). Some of the described encoders are systematic binary encoders and some are only partly systematic.**

***Index Terms*— Partly nonsystematic encoders, Reed–Solomon codes, subalphabet subcodes, systematic encoders.**

I. INTRODUCTION

As a motivation, we consider the case where a Reed–Solomon (RS) decoder over the alphabet GF(2^8) is available. Since GF(2), GF(2^2), and GF(2^4) are subfields of GF(2^8), Bose–Chaudhury–Hocquenghem (BCH) codes of length ≤ 255 of appropriate dimension over those smaller fields can be decoded by this decoder. However, if the number of symbols in the alphabet is another power of 2, like $2^3, 2^5, 2^6$, or 2^7, the situation is different because the corresponding exponents ($3, 5, 6.$ or 7) are not divisors of 8. Thus linear codes over GF(2^3), GF(2^5), GF(2^6), or GF(2^7) are never subfield subcodes of RS codes over GF(2^8) and they cannot be decoded by means of a decoder working in GF(2^8).

Consider in particular the case where the natural alphabet of the problem contains $2^6 = 64$ symbols. Since GF(2^6) is not matched to a decoder working over GF(2^8), a possible solution is to use a Reed–Solomon code over GF(2^6) of length (if extended) at most equal to 65 (or 66 in some specific cases). However, this may be too small for several applications.

Another solution is to use sufficiently long BCH codes over GF(2^6). This requires a decoder working in the field GF(2^{12}) (or larger) which is not a classical technology. Moreover, the necessary redundancy is essentially twice as large as the one needed with maximum distance separable (or MDS) [1] codes.

In the sequel we discuss another type of solution to this problem. Given an additive subalphabet S of GF(2^m) containing 2^t symbols $(t < m)$, we investigate several encoding procedures to generate the codewords of an MDS code over GF(2^m) for which all symbols are in S.

Such codes were first considered in [2]. In this reference, the subalphabet S of GF(2^m) was chosen to be its GF(2)-subspace generated by $\{1, \omega, \cdots, \omega^{t-1}\}$ where ω is a primitive element of GF(2^m). The main result of [2] was a formula giving the GF(2) dimension of the codes having this structure.

In [3], a heuristic approach was described to encode information in such a way that the produced codewords are over a linear subspace of GF(2^m), but no claim was made of optimality.

More recently [4]–[7], the subalphabet subcodes of Reed–Solomon codes (SSRS codes) were thoroughly revisited, the dimension formula was derived along a different approach, the so-called "exceptional"

Manuscript received April 20, 1998; revised December 10, 1998. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Cambridge, MA, August 16–21, 1998.

The authors are with Information and Modulation Department, Canon Research Centre France, F-35517 Cesson-Sévigné, France.

Communicated by I. F. Blake, Asociate Editor for Coding Theory.
Publisher Item Identifier S 0018-9448(99)04374-6.