

LINEAR REPRESENTATIONS OF FINITE GEOMETRIES AND ASSOCIATED LDPC CODES.

PETER SIN, JULIEN SORCI AND QING XIANG

ABSTRACT. The *linear representation* of a subset of a finite projective space is an incidence system of affine points and lines determined by the subset. In this paper we use character theory to show that the rank of the incidence matrix has a direct geometric interpretation in terms of certain hyperplanes. We consider the LDPC codes defined by taking the incidence matrix and its transpose as parity-check matrices, and in the former case prove a conjecture of Vandendriessche that the code is generated by words of minimum weight called plane words. In the latter case we compute the minimum weight in several cases and provide explicit constructions of minimum weight code-words.

1. INTRODUCTION

Codes with sparse parity-check matrix were first considered in the seminal PhD dissertation of Gallagher [7]. Such a code is called a low-density parity-check, or LDPC, code. While Gallagher's ideas were largely overlooked for many years, since the 1990s there has been a resurgence of interest in constructing LDPC codes due to their relatively fast decoding algorithms while still achieving high rates of transmission. One method of constructing these codes is by taking the incidence matrix of a finite geometry as parity-check matrix, which is the method that we shall consider here.

Let $q = p^e$ be a prime power and $n \geq 2$ be an integer. Let E be an $(n + 1)$ -dimensional vector space over $\mathbb{F}_q$ and V an n -dimensional subspace. Then $H := \text{PG}(V) \cong \text{PG}(n - 1, q)$ is a hyperplane of $\text{PG}(E) \cong \text{PG}(n, q)$. Let $P := \text{PG}(E) \setminus H \cong \text{AG}(n, q)$ be the complementary affine space, which we shall often view as a set of q^n affine points with H as the hyperplane at infinity. Additionally, if ℓ is an affine line of P with point at infinity u in H , we will at times refer to u as the direction of the line ℓ , and thus view the points of H as the directions of affine lines in P .

Fixing an arbitrary subset $\mathcal{K}$ of H , we consider the point-line incidence system whose point set is P and whose line set L is the set of affine lines of P whose direction is in $\mathcal{K}$. Each point of $\mathcal{K}$ is the direction of q^{n-1} parallel

The third author was partially supported by NSF grant DMS-1855723.

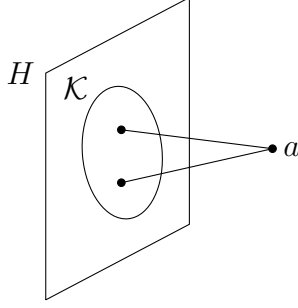


FIGURE 1. A point a and two lines in the geometry $T_{n-1}^*(\mathcal{K})$.

lines in L , so that $|L| = q^{n-1}|\mathcal{K}|$. A line in L will be viewed as a set of q points of P and we define a line to be incident to its points. This point-line incidence system is called the *linear representation* of the set $\mathcal{K}$ and is denoted by $T_{n-1}^*(\mathcal{K})$ (Figure 1). The idea for this definition can be traced back to certain constructions of generalized quadrangles by J. Tits, by R. W. Ahrens and G. Szekeres [1], and by M. Hall [8] in the 1960s. Later the definition was extended to arbitrary $\mathcal{K}$ and n in [5]. The automorphism group of a linear representation has been studied in [3] and [6].

We introduce homogeneous coordinates $X_0, X_1, \dots, X_n$ in $\text{PG}(n, q)$ and without loss of generality let H be the hyperplane $X_0 = 0$. We fix arbitrary orderings on P and L and let N denote the point-line incidence matrix. We have defined N as an integer matrix, but we can consider it as a matrix over any field. The following theorem gives a geometric interpretation of the rank of N .

Theorem 1.1. *Let F be a field in which $q \neq 0$. Then $\text{rank}_F N$ is equal to the number of functions in the dual space $V^* = \text{Hom}(V, \mathbb{F}_q)$ that take the value zero at some point of $\mathcal{K}$. In geometric terms, $\text{rank}_F N = 1 + (q - 1)h_{\mathcal{K}}$, where $h_{\mathcal{K}}$ is the number of hyperplanes in the projective space H that have nonempty intersection with $\mathcal{K}$.*

Remark 1.2. If $q = 0$ in F , then $\text{rank}_F N$ will be bounded above by $\text{rank}_{\mathbb{Q}} N$, which is given by the theorem. However, small examples show that inequality will be strict in general, and we have no conjecture yet for the exact value of $\text{rank}_F N$ in this case.

We define the F -codes $\mathcal{C}$ and $\mathcal{D}$ to be the codes with parity-check matrices N and N^T , respectively. Each row of N has weight $|\mathcal{K}|$ and each column has weight q so that $\mathcal{C}$ and $\mathcal{D}$ are LDPC codes of lengths $|\mathcal{K}|q^{n-1}$ and q^n , respectively, with dimensions given by Theorem 1.1. The code $\mathcal{C}$ has been studied in [16], [12] and [15]. In [17], Kou et al. considered the codes $\mathcal{C}$ and

$\mathcal{D}$ for $F = \mathbb{F}_2$ and when $\mathcal{K}$ is the set of all points of H , referring to $\mathcal{C}$ as the type-II geometry- $\mathbf{G}$ LDPC code and $\mathcal{D}$ as the type-I geometry- $\mathbf{G}$ LDPC code. In this case the geometry $T_{n-1}^*(\mathcal{K})$ is isomorphic to $\text{AG}(n, q)$, so the parity-check matrix for $\mathcal{C}$ is the full incidence matrix of $\text{AG}(n, q)$ and the parity-check matrix for $\mathcal{D}$ is the transpose. Kou et al. noted that $\mathcal{C}$ is one-step majority-logic decodable and can correct $\lfloor q/2 \rfloor$ errors with this decoding scheme, so that the minimum distance of $\mathcal{C}$ is at least $q + 1$. Similarly, they also noted that using the same decoding algorithm, one can correct $\lfloor |\mathcal{K}|/2 \rfloor$ errors in $\mathcal{D}$ so that $\mathcal{D}$ has minimum distance at least $|\mathcal{K}| + 1$.

In [14], Tang et al. studied the codes $\mathcal{C}$ and $\mathcal{D}$ when $F = \mathbb{F}_2$ and for an arbitrary subset of points $\mathcal{K}$ in H , referring to $\mathcal{C}$ as the *Complementary Gallager Euclidean Geometry LDPC code* and $\mathcal{D}$ as the *Gallager Euclidean Geometry LDPC code*. For these more general codes, they also provided the lower bounds $q + 1$ and $|\mathcal{K}| + 1$ on the minimum distances of $\mathcal{C}$ and $\mathcal{D}$, respectively, by observing that any q columns of N or $|\mathcal{K}|$ columns of N^T are linearly independent. We note here that these lower bounds hold for any field F , and any prime power q by the following argument. A codeword of $\mathcal{D}$ is an element of F^P , whose support is a set of points in P . If x is a codeword of $\mathcal{D}$, and s is a point in the support of x , then each of the $|\mathcal{K}|$ lines meeting s must contain a further point of the support. These further points are distinct, hence the weight of x is at least $|\mathcal{K}| + 1$. Similarly, any codeword x in $\mathcal{C}$ has the property that if ℓ is a line in the support of x , then for each of the q points incident to ℓ there is a further line of the support meeting ℓ at that point, so that x must have weight at least $q + 1$.

The proof of Theorem 1.1 will be given in the next section, but we can sketch the main ideas now. We consider the actions of the additive group of the vector space V on the affine space P and on the set L of lines. These actions turn the spaces F^P and F^L of F -valued functions on P and L into FV -modules, where FV is the group algebra of V , and the incidence relation then defines an FV -module homomorphism $F^L \rightarrow F^P$, whose matrix is N . The regular action of V on P gives in addition an FV -isomorphism from F^P to the space F^V of F -valued functions on V . Thus, $\text{rank}_F N$ equals the dimension of the image in F^V of the composite homomorphism $F^L \rightarrow F^V$. The standard bases for F^L , F^P and F^V consist of the characteristic functions of elements. When F contains a primitive p -th root of unity, F^V has a second natural basis, namely the group $\hat{V} = \text{Hom}(V, F^\times)$ of F -characters of V . Using character theory, we can reduce the rank problem to one of counting certain characters. Finally, we make use of a natural bijection between $\hat{V}$ and the dual space $V^* = \text{Hom}_{\mathbb{F}_q}(V, \mathbb{F}_q)$ to arrive at Theorem 1.1.

In §3 we consider the F -codes $\mathcal{C}$ and $\mathcal{D}$ and prove a conjecture of P. Vanden-driessche [15] that the code $\mathcal{C}$ is generated by certain codewords called *plane*

words, which are known to be words of minimum weight. We will also define words in $\mathcal{D}$ analogous to plane words called *capacitor words* and show that they are codewords that span $\mathcal{D}$. In §4 we apply Theorem 1.1 in particular cases of n and $\mathcal{K}$. In the case where $\mathcal{K}$ is a rational normal curve minus its point at infinity the sets P and L form the bipartition of the vertex set of a bipartite graph called the Wenger graph. The Wenger graphs have been studied extensively; their automorphism groups have been found in [2] and their spectra determined in [4]. Our theorem shows that the rank of the adjacency matrix of a Wenger graph over any field F in which $q \neq 0$, is the same as the real rank, hence equal to the matrix size minus the multiplicity of zero as an eigenvalue of the adjacency matrix. The multiplicity of every eigenvalue has been computed in [4]. However we shall derive the rank independently, directly from Theorem 1.1. In the special case $n = 3$, we obtain a simple new proof for the dimensions of the LDPC codes called $\text{LU}(3, q)$, that were first computed in [13] for fields of characteristic 2, and for other fields with $q \neq 0$ in [16].

2. PROOF OF THEOREM 1.1

In proving Theorem 1.1 F may be replaced by any extension field F' , as $\text{rank}_{F'} N = \text{rank}_F N$, so we assume from now on that F is a field in which $q \neq 0$, that contains a primitive p -th root of unity ω . Let F^L be the space of F -valued functions on L and F^P the space of F -valued functions on P . These vector spaces are actually FV -permutation modules, as V permutes the sets L and P by an action that we now describe. A point of P has homogeneous coordinates $(1 : a_1 : \dots : a_n)$ with $a_i \in \mathbb{F}_q$ for all i , and together with $(0 : b_1 : \dots : b_n) \in \mathcal{K}$ determines an affine line consisting of the q points $(1 : a_1 + tb_1 : \dots : a_n + tb_n)$ for $t \in \mathbb{F}_q$. The additive group of V acts regularly on P via the action

$$(0, v_1, \dots, v_n) \cdot (1 : a_1 : \dots : a_n) = (1 : a_1 + v_1 : \dots : a_n + v_n)$$

From the description of lines in L above it follows that the action of the group V on P induces an incidence-preserving action on L , where the vector $(0, v_1, \dots, v_n) \in V$ moves the line $\{(1 : a_1 + tb_1 : \dots : a_n + tb_n) \mid t \in \mathbb{F}_q\}$ to the line $\{(1 : a_1 + v_1 + tb_1 : \dots : a_n + v_n + tb_n) \mid t \in \mathbb{F}_q\}$.

There are natural bases for the vector spaces F^L and F^P : For $\ell \in L$ we denote by $[\ell] \in F^L$ the characteristic function of the element ℓ , which takes value one at ℓ and zero elsewhere. These form a basis of F^L . Similarly, for $a \in P$, we let $\chi_a \in F^P$ denote the characteristic function of a , and the χ_a form a basis of F^P . More generally, for any subset A of P we let χ_A denote the characteristic function of A . We consider the incidence map $\eta : F^L \rightarrow F^P$, defined by $[\ell] \mapsto \chi_\ell = \sum_{a \in \ell} \chi_a$. With respect to the bases $\{[\ell]\}_{\ell \in L}$ of F^L and $\{\chi_a\}_{a \in P}$ of F^P , the matrix of η is N , considered as a matrix over F .

The set P is isomorphic, as a V -set, to the set V , with the left regular action, under the map taking each point $(1 : a_1 : a_2 : \cdots : a_n)$ to the vector $(0, a_1, a_2, \dots, a_n)$. Hence we have an induced FV -module isomorphism $\sigma : F^P \rightarrow F^V$, mapping the characteristic function of a point to the characteristic function of the corresponding vector. We shall study the composite map $\sigma\eta : F^L \rightarrow F^V$, since the dimension of its image is equal to $\text{rank}_F N$. The basis element $[\ell]$ for $\ell \in L$ is mapped to the sum of all the characteristic functions of vectors corresponding to points of ℓ .

Next, we briefly present the character theory necessary for the proof of Theorem 1.1. For $\theta \in V^* = \text{Hom}_{\mathbb{F}_q}(V, \mathbb{F}_q)$, the dual vector space to V , we can define an F -character λ by $\lambda(v) = \omega^{\text{Tr}(\theta(v))}$, for all $v \in V$, where Tr is the trace map from $\mathbb{F}_q$ to $\mathbb{F}_p$. In this way we obtain a bijection from V^* to the group $\hat{V} := \text{Hom}(V, F^\times)$ of F -characters of V . We shall denote the character corresponding to the linear function θ by λ_θ and the linear function corresponding to a character λ by θ_λ . An element of V^* can be represented in the usual way using dual coordinates as $\theta = [c_1, c_2, \dots, c_n]$, where $\theta(0, v_1, \dots, v_n) = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$ for $(0, v_1, v_2, \dots, v_n) \in V$. Note that if W is an $\mathbb{F}_q$ -subspace of V , then for $\theta \in V^*$ we have

$$(1) \quad \text{Tr}(\theta(W)) = \begin{cases} \mathbb{F}_p, & \text{if } W \not\subseteq \ker(\theta), \\ \{0\}, & \text{if } W \subseteq \ker(\theta). \end{cases}$$

Lemma 2.1. *If W is an $\mathbb{F}_q$ -subspace of V and $\lambda \in \hat{V}$, then $\sum_{w \in W} \lambda(w) \neq 0$ if and only if $W \subseteq \ker(\theta_\lambda)$.*

Proof. Suppose $W \not\subseteq \ker(\theta_\lambda)$. Then by (1) the composite map $\text{Tr} \circ \theta_\lambda$ is a surjective group homomorphism from W to $\mathbb{F}_p$. Therefore, $\sum_{w \in W} \lambda(w)$ is a multiple of the complete sum of p -th roots of unity, hence equal to zero. Conversely, if $W \subseteq \ker(\theta_\lambda)$, then $\sum_{w \in W} \lambda(w) = |W| \neq 0$. $\square$

Let F^V be the space of F -valued functions on V , with V acting by the formula $(vg)(w) = g(w - v)$, for $v, w \in V$ and $g \in F^V$. The space F^V has two bases. The first is the set of characteristic functions δ_v of elements $v \in V$; the second is $\hat{V}$. By the orthogonality relations we have, for $v \in V$,

$$(2) \quad \delta_v = \frac{1}{|V|} \sum_{\lambda \in \hat{V}} \lambda(-v) \lambda$$

For $g \in F^V$, we write $g = \sum_{\lambda \in \hat{V}} a_\lambda \lambda$ as an F -linear combination of characters, where $a_\lambda = \frac{1}{|V|} \sum_{v \in V} g(v) \lambda(-v)$; if $a_\lambda = 0$ then we say that λ is a root of g ; otherwise λ is a nonroot of g . Define $\hat{V}_g := \{\lambda \in \hat{V} \mid a_\lambda \neq 0\}$, the set of nonroots of g .

The following is a special case of a well-known general principle.

Lemma 2.2. *For each $g \in F^V$, the set $\hat{V}_g$ is a basis for the FV -submodule of F^V generated by g .*

Proof. It is clear that the F -span of $\hat{V}_g$ contains g . Since the span is itself an FV -submodule, it contains the F -submodule FVg generated by g . Conversely, for each $\lambda \in \hat{V}$, FV contains the idempotent $e_\lambda = \frac{1}{|V|} \sum_{v \in V} \lambda(v)v$ and an easy computation using the orthogonality relations shows that e_λ acts as the identity on λ while annihilating all other characters. Thus $e_\lambda g = a_\lambda \lambda$, and so FVg contains $\hat{V}_g$. $\square$

Remark 2.3. In coding theory terms, for each $g \in F^V$, the F -submodule FVg is the ideal $F[V]$ generated by g . The lemma is saying that the dimension of FVg is the number of nonroots of g .

We are now ready to complete the proof of Theorem 1.1. If $\ell = \{(1 : a_1 + tb_1 : a_2 + tb_2 : \cdots : a_n + tb_n) \mid t \in \mathbb{F}_q\}$, then

$$(3) \quad \sigma\eta([\ell]) = \sum_{t \in \mathbb{F}_q} \delta_{(0, a_1 + tb_1, a_2 + tb_2, \dots, a_n + tb_n)}.$$

In particular, if $\ell \in L$ is a line through $(1 : 0 : 0 : \cdots : 0) \in P$ then the elements of V corresponding to points of ℓ form the one-dimensional subspace W_ℓ of V generated by $(0, b_1, b_2, \dots, b_n)$. Thus,

$$(4) \quad \begin{aligned} \sigma\eta([\ell]) &= \sum_{w \in W_\ell} \delta_w \\ &= \frac{1}{|V|} \sum_{w \in W_\ell} \sum_{\lambda \in \hat{V}} \lambda(-w) \lambda \\ &= \frac{1}{|V|} \sum_{\theta \in V^*} \left(\sum_{w \in W_\ell} \lambda_\theta(w) \right) \lambda_\theta, \end{aligned}$$

since $\hat{V}$ is parametrized by V^* and negation permutes W_ℓ . By Lemma 2.1, the coefficient of λ_θ is nonzero if and only if $W_\ell \subseteq \ker(\theta)$, or in other words,

$$\hat{V}_{\sigma\eta([\ell])} = \{\lambda_\theta \in \hat{V} \mid W_\ell \subseteq \ker(\theta)\}.$$

Let $L_0 \subset L$ denote the set of lines through $(1 : 0 : 0 : \cdots : 0) \in P$.

By Lemma 2.2 the set

$$(5) \quad \{\lambda_\theta \in \hat{V} \mid \exists \ell \in L_0, W_\ell \subseteq \ker(\theta)\}.$$

is a basis for the FV -submodule of F^V generated by the images of lines in L_0 . As every line of L is in the V -orbit of a line through $(1 : 0 : 0 : \cdots : 0) \in P$, this submodule is equal to $\text{Im}(\sigma\eta)$. For each $\ell \in L_0$, the subspace W_ℓ may be viewed as a point of H , and as such it is the point at infinity of the line ℓ . In

this way, L_0 corresponds bijectively with $\mathcal{K}$, so the basis (5) of $\text{Im}(\sigma\eta)$ is equal to

$$(6) \quad \{\lambda_\theta \in \hat{V} \mid \exists u \in \mathcal{K}, \theta(u) = 0\}.$$

Thus, the first statement of Theorem 1.1 is proved. For the last statement of Theorem 1.1 we observe that nonzero linear functions in V^* define hyperplanes of H and two such functions define the same hyperplane if and only if each is a nonzero scalar multiple of the other. Thus, if $h_{\mathcal{K}}$ denotes the number of hyperplanes of H that meet $\mathcal{K}$, there are $1 + (q - 1)h_{\mathcal{K}}$ linear functions in V^* that take the value 0 at some point of $\mathcal{K}$.

3. THE CODES ASSOCIATED TO $T_{n-1}^*(\mathcal{K})$

3.1. The Code $\mathcal{C}$. The code $\mathcal{C} \leq F^L$ is defined as the kernel of η . That is, $\mathcal{C}$ is the F -linear code with N as its parity check matrix, hence its dimension can be found immediately from Theorem 1.1. We shall next define a subcode $\mathcal{C}'$, following [15]. A *plane word* is defined as follows. Let u_1 and u_2 be two points of $\mathcal{K}$. There are $\frac{q^{n-1}-1}{q-1}$ planes in $\text{PG}(E)$ containing the line joining u_1 and u_2 , among which there are $\frac{q^{n-2}-1}{q-1}$ planes contained in H . We shall refer to these q^{n-2} planes not contained in H as *affine planes*. Let T be an affine plane of $\text{PG}(E)$ whose line at infinity in H is the line joining u_1 and u_2 . Then the plane word $w(u_1, u_2, T)$ is the sum of the characteristic functions of the q affine lines of T having u_1 as point at infinity minus the corresponding sum with respect to u_2 . A plane word clearly belongs to $\mathcal{C}$ and it has weight $2q$. We denote by $\mathcal{C}'$ the subcode spanned by the plane words for all possible choices of $u_1, u_2 \in \mathcal{K}$ and T .

Theorem 3.1. *Assume that $q \neq 0$ in F . Then $\mathcal{C}' = \mathcal{C}$.*

Proof. Since we will be considering geometries and codes associated with different subsets $\mathcal{K}$ of H , we will adopt notation such as $L_{\mathcal{K}}$ for the set of lines of the geometry $T_{n-1}^*(\mathcal{K})$, $\eta_{\mathcal{K}}$ for the incidence map, and $\mathcal{C}_{\mathcal{K}}, \mathcal{C}'_{\mathcal{K}}$ for the codes. We proceed by induction on $|\mathcal{K}|$, the case of $\mathcal{K} = \emptyset$ being trivial. Assume inductively that for some $\mathcal{K}$ we have $\mathcal{C}_{\mathcal{K}} = \mathcal{C}'_{\mathcal{K}}$ and let u_0 be a point of H outside $\mathcal{K}$, and $\mathcal{K}' = \mathcal{K} \cup \{u_0\}$. We shall consider the vector space $\overline{E} = E/u_0$ and its projective space $\text{PG}(\overline{E})$. We use the bar convention for images in $\text{PG}(\overline{E})$ of objects in $\text{PG}(E)$. Thus $\overline{H} = \text{PG}(\overline{V})$ is a hyperplane of $\text{PG}(\overline{E})$ and $\overline{P}$ is its affine complement. The image $\overline{\mathcal{K}}$ of $\mathcal{K}$ in $\overline{H}$ has one point for each line of H through u_0 that meets $\mathcal{K}$. Let $\overline{L}$ denote the set of affine lines in $\overline{P}$ with point at infinity in $\overline{\mathcal{K}}$. Under the projection from $\text{PG}(E)$ to $\text{PG}(\overline{E})$ the set $L_{\{u_0\}}$ maps bijectively to $\overline{P}$ and the set $\mathcal{T}$ of affine planes T with line at infinity passing

through u_0 and some point of $\mathcal{K}$ maps bijectively to $\bar{L}$. Naturally, these bijections preserve incidence. Thus, the incidence system $(L_{\{u_0\}}, \mathcal{T})$ is isomorphic to $T_{n-2}^*(\bar{\mathcal{K}})$. In the decomposition

$$(7) \quad F^{L_{\mathcal{K}'}} = F^{L_{\{u_0\}}} \oplus F^{L_{\mathcal{K}}}$$

let π be the projection onto the first summand $F^{L_{\{u_0\}}}$. Then the projection $\pi(w(u_0, u_1, T))$ of a plane word is just the sum of the characteristic functions of the q affine lines of T having u_0 as point at infinity. Then, under the isomorphism of $F^{L_{\{u_0\}}}$ with $F^{\bar{P}}$ induced by the above bijection, these q affine lines become the q points of the affine line $\bar{T}$, so the image in $F^{\bar{P}}$ of $\pi(w(u_0, u_1, T))$ is the characteristic function of $\bar{T}$. It follows that $\pi(\mathcal{C}'_{\mathcal{K}'})$ is isomorphic to the subspace of $F^{\bar{P}}$ spanned by the characteristic functions of lines in $\bar{L}$, so its dimension is equal to the rank of $\eta_{\bar{\mathcal{K}}} : F^{\bar{L}} \rightarrow F^{\bar{P}}$. Since $\pi(\mathcal{C}'_{\mathcal{K}}) = \{0\}$ we have shown that

$$(8) \quad \dim \mathcal{C}'_{\mathcal{K}'} - \dim \mathcal{C}'_{\mathcal{K}} \geq \text{rank}_F \eta_{\bar{\mathcal{K}}}.$$

By our induction hypothesis and the fact that $\mathcal{C}'_{\mathcal{K}'} \subseteq \mathcal{C}_{\mathcal{K}'}$, we therefore obtain

$$(9) \quad \dim \mathcal{C}_{\mathcal{K}'} - \dim \mathcal{C}_{\mathcal{K}} \geq \text{rank}_F \eta_{\bar{\mathcal{K}}},$$

with equality if and only if $\mathcal{C}'_{\mathcal{K}'} = \mathcal{C}_{\mathcal{K}'}$. On the other hand by considering the linear maps $\eta_{\mathcal{K}}$ and $\eta_{\mathcal{K}'}$ we obtain

$$(10) \quad \dim \mathcal{C}_{\mathcal{K}'} - \dim \mathcal{C}_{\mathcal{K}} = q^{n-1} - (\text{rank}_F \eta_{\mathcal{K}'} - \text{rank}_F \eta_{\mathcal{K}}).$$

By Theorem 1.1,

$$(11) \quad \begin{aligned} \text{rank}_F \eta_{\mathcal{K}'} - \text{rank}_F \eta_{\mathcal{K}} &= \#\{\theta \in V^* \mid \theta(u_0) = 0 \text{ and } \theta(u) \neq 0 \text{ for all } u \in \mathcal{K}\} \\ &= \#\{\theta \in \bar{V}^* \mid \theta(u) \neq 0 \text{ for all } u \in \bar{\mathcal{K}}\} \\ &= q^{n-1} - \text{rank}_F \eta_{\bar{\mathcal{K}}}. \end{aligned}$$

Therefore, we have equality in (9) and the inductive step is established. $\square$

Remark 3.2. Theorem 3.1 was conjectured in [15], where the case $n = 3$ was proved. Now that Theorem 3.1 is established, the minimum weight of $\mathcal{C}$ (for arbitrary n and $\mathcal{K}$) is given by [15, Theorem 5.4]. As long as $\mathcal{C} \neq 0$ the plane words are words of minimum weight $2q$ (although not the only words of this weight in general). Thus, in all cases, $\mathcal{C}$ is generated by its words of minimum weight.

3.2. The Code $\mathcal{D}$. We begin our discussion of $\mathcal{D}$ by describing a generating set of codewords with a geometric flavor similar to that of the plane words. First we need some expressions regarding indicator functions. If U is an affine hyperplane of P , then we can view U as a coset of a subgroup U_0 of V , say

$U = U_0 + x$, where U_0 is the hyperplane parallel to U containing 0. We can express the indicator function of U_0 in terms of characters

$$(12) \quad \chi_{U_0} = \frac{1}{q} \sum_{\lambda \in \widehat{(V/U_0)}} \lambda$$

where $\widehat{(V/U_0)}$ denotes the characters of $\widehat{V}$ whose kernels contain U_0 . The indicator function of the coset $U = U_0 + x$ is simply the translate of χ_{U_0} by x , given by

$$(13) \quad \chi_U = \frac{1}{q} \sum_{\lambda \in \widehat{(V/U_0)}} \lambda(-x)\lambda.$$

Let U_1, U_2 be parallel affine hyperplanes of the affine space P whose common subspace at infinity is a hyperplane T of H that does not meet $\mathcal{K}$. We define a *capacitor word* of U_1, U_2 to be the indicator function of U_1 minus that of U_2 .

Theorem 3.3. *Let F be a field with $q \neq 0$. The capacitor words are codewords generating the F -code $\mathcal{D}_{\mathcal{K}}$.*

Proof. Let w be a capacitor word on the affine hyperplanes U_1 and U_2 . The condition on their common subspace at infinity implies that every line $\ell \in L$ meets the hyperplanes U_1 and U_2 each at exactly one point, so that w is indeed a codeword of $\mathcal{D}_{\mathcal{K}}$. To show that the capacitor words generate $\mathcal{D}_{\mathcal{K}}$, let W be the F -subspace generated by the capacitor words on all pairs of parallel affine hyperplanes with subspace at infinity not meeting $\mathcal{K}$. Then $W \subseteq \mathcal{D}_{\mathcal{K}}$. By Theorem 1.1 the dimension of $\mathcal{D}_{\mathcal{K}}$ is equal to the number of functions in V^* that do not contain any point of $\mathcal{K}$ in its kernel. Each such function $\theta \in V^*$ corresponds bijectively to a character $\lambda_{\theta} \in \widehat{V}$, as in §2. We shall show that each such character lies in W , which will imply that $\dim W \geq \dim \mathcal{D}_{\mathcal{K}}$, and hence that $W = \mathcal{D}_{\mathcal{K}}$. Thus, let $\theta \in V^*$ be given, such that no point of $\mathcal{K}$ lies in $\ker(\theta)$. Let $U_0 = \ker(\theta)$, an $\mathbb{F}_q$ -hyperplane of V , and let $\lambda_{\theta} \in \widehat{V/U_0}$ be the corresponding character. Our aim is to show that $\lambda_{\theta} \in W$. By Lemma 2.2, it suffices to show that λ_{θ} appears with nonzero coefficient in some capacitor word when the latter is expressed as a linear combination of characters. We can choose $x \in V$ such that $\lambda_{\theta}(x) \neq 1$. Then $x \notin U_0$, and from 12 and 13 we see that the indicator function of the capacitor word based on $U = U_0 + x$ and U_0 is equal to

$$(14) \quad \chi_U - \chi_{U_0} = \frac{1}{q} \sum_{\lambda \in \widehat{(V/U_0)}} (\lambda(-x) - 1)\lambda.$$

By choice of x , the coefficient of λ_{θ} in this expression is nonzero, and the proof is complete. $\square$

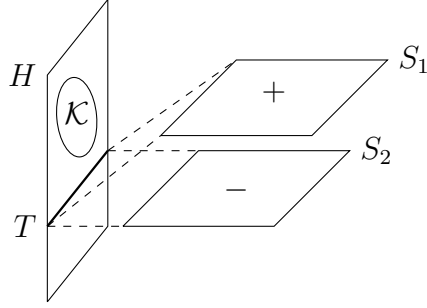


FIGURE 2. The hyperplanes giving a capacitor word.

Remark 3.4. The result in Theorem 3.3 is still valid if there are no capacitor words. In this case every hyperplane of H meets $\mathcal{K}$ so that N has full rank, and hence $\mathcal{D}_{\mathcal{K}} = \{0\}$.

Similarly, we can define a codeword whose weight depends on the dimension of the subspace spanned by the points of $\mathcal{K}$. Suppose that the points of $\mathcal{K}$ span a d -dimensional subspace of H denoted by X , and let Y be an affine $(d+1)$ -dimensional space whose intersection at infinity is X , and T a hyperplane of X not meeting $\mathcal{K}$. Then let S_1 and S_2 be hyperplanes of Y whose intersections at infinity are both T . Define a d -capacitor word as the sum of the characteristic functions of points of S_1 minus the corresponding sum of points of S_2 . It is easy to see that this is indeed a codeword of $\mathcal{D}$ of weight $2q^d$: given any line ℓ of L , either ℓ has no points contained in Y , or ℓ is totally contained in Y . In the latter case, ℓ must meet each of the hyperplanes S_1 and S_2 in exactly one point, and hence we have a sum of zero when summing over the points incident to ℓ .

Now we turn to the question of the minimum weight of $\mathcal{D}$. As mentioned in the Introduction, we have a general lower bound of $|\mathcal{K}| + 1$ for the minimum weight, but this bound may be improved for certain fields and subsets $\mathcal{K}$. We record in the next lemma a few facts about the code $\mathcal{D}_{\mathcal{K}}$ that will be useful.

Lemma 3.5. *Let F be any field which may or may not have $q = 0$.*

- (1) *If $\mathcal{K}' \subseteq \mathcal{K}$ then $\mathcal{D}_{\mathcal{K}} \subseteq \mathcal{D}_{\mathcal{K}'}$, and hence $d(\mathcal{D}_{\mathcal{K}}) \geq d(\mathcal{D}_{\mathcal{K}'})$.*
- (2) *If w is a codeword of $\mathcal{D}_{\mathcal{K}}$, then any line ℓ in $L_{\mathcal{K}}$ is either skew to the support of w , or contains at least two points of the support of w .*
- (3) *$d(\mathcal{D}_{\mathcal{K}}) \geq |\mathcal{K}| + 1$.*
- (4) *If w is a codeword of $\mathcal{D}_{\mathcal{K}}$ for $\mathcal{K}$ properly contained in a line, then either $wt(w) \geq 2|\mathcal{K}| + 2$ or the support of w is contained in an affine plane whose line at infinity contains $\mathcal{K}$.*

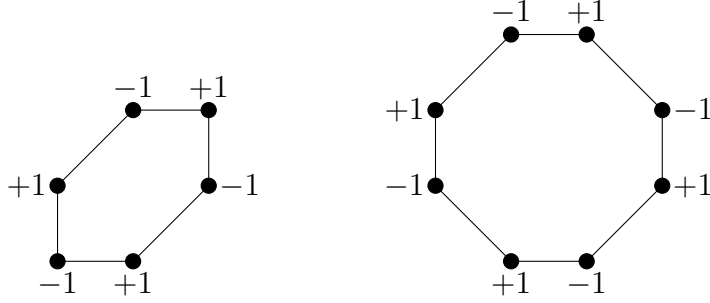


FIGURE 3. The codewords w (left) and w' (right) of weights 6 and 8, respectively, given in Example 3.9.

Proof. For (1), if a word satisfies the parity-check equations of $N_{\mathcal{K}}^T$ then it also satisfies the equations of $N_{\mathcal{K}'}^T$ for any $\mathcal{K}' \subseteq \mathcal{K}$.

For (2), if a line $\ell \in L_{\mathcal{K}}$ meets exactly one point of the support of w , then w does not satisfy the parity-check equation associated to ℓ and hence cannot be a codeword.

Part (3) was proven in the Introduction, so it only remains to prove part (4). Suppose there are two affine planes π_1, π_2 with line at infinity containing $\mathcal{K}$ and both containing a point in the support of w . Using the same argument as the proof of part (3), each plane π_1, π_2 must contain at least $|\mathcal{K}| + 1$ distinct points so that w has weight at least $2|\mathcal{K}| + 2$. $\square$

Remark 3.6. Part (4) of Lemma 3.5 implies that when $\mathcal{K}$ is properly contained in a line, either $d(\mathcal{D}_{\mathcal{K}}) \geq 2|\mathcal{K}| + 2$ or $\mathcal{D}_{\mathcal{K}}$ has the same minimum distance as the code given by the parity-check matrix $N_{\mathcal{K}}^T$ restricted to the lines of $T_1^*(\mathcal{K})$.

Theorem 3.7. *Let F be a totally ordered field. Then the minimum distance of the F -code $\mathcal{D}_{\mathcal{K}}$ is at least $2|\mathcal{K}|$. In particular, when $F = \mathbb{Q}$ or $\mathbb{R}$ then $d(\mathcal{D}_{\mathcal{K}}) \geq 2|\mathcal{K}|$.*

Proof. Let w be a codeword in $\mathcal{D}_{\mathcal{K}}$, and let s be a point in the support of w , where without loss of generality we assume that $w(s) > 0$. There are $|\mathcal{K}|$ lines containing s , each of which contains a point t in the support of w with $w(t) < 0$. Each such point t is also met by $|\mathcal{K}|$ lines, each of which contains a point u in the support of w with $w(u) > 0$. Therefore the sets

$$\{x \in P : w(x) > 0\}$$

$$\{x \in P : w(x) < 0\}$$

both have cardinality at least $|\mathcal{K}|$. $\square$

TABLE 1

i	a_i	b_i
1	(0,0)	(0,0)
2	(0,1)	(0,1)
3	(1,2)	(1,2)
4	(2,2)	(2,2)
5	(2,1)	(3,1)
6	(1,0)	(3,0)
7	—	(2,-1)
8	—	(1,-1)

Example 3.8. Let ℓ be a line of $\text{PG}(n, q)$ contained in H , and let v be a point on ℓ . Let $\mathcal{K}$ be a subset of $\ell \setminus \{v\}$. If ℓ_1 and ℓ_2 are any two affine lines with point at infinity v , and both lying in an affine plane with line at infinity ℓ , then the word $\chi_{\ell_1} - \chi_{\ell_2}$ is a codeword of $\mathcal{D}_{\mathcal{K}}$ of weight $2q$.

Example 3.9. When $\mathcal{K}$ is a particular set of 3 or 4 points contained in a line we can give explicit codewords of weight 6 and 8, respectively. As proved in Lemma 3.5 these words are contained in an affine plane. In these examples, we assume that the characteristic of $\mathbb{F}_q$ is large enough for the given points to be distinct. Let

$$\begin{aligned}\mathcal{K} &= \{\langle(1, 0)\rangle, \langle(0, 1)\rangle, \langle(1, 1)\rangle\}, \\ \mathcal{K}' &= \{\langle(1, 0)\rangle, \langle(0, 1)\rangle, \langle(1, -1)\rangle, \langle(1, 1)\rangle\},\end{aligned}$$

Then set $w = \sum_{i=1}^6 (-1)^i \delta_{a_i}$, $w' = \sum_{i=1}^8 (-1)^i \delta_{b_i}$, where the points a_i, b_i are given in Table 1. It is an easy exercise to check that $w \in \mathcal{D}_{\mathcal{K}}$ and $w' \in \mathcal{D}_{\mathcal{K}'}$, and that $\text{wt}(w) = 6$, $\text{wt}(w') = 8$. It would be interesting if there were a general construction for similar codewords of $\mathcal{D}_{\mathcal{K}}$ of weight $2|\mathcal{K}|$ for any set of points $\mathcal{K}$ contained in a line.

The next examples shows that the minimum distance of $\mathcal{D}_{\mathcal{K}}$ can be less than $2|\mathcal{K}|$ when $q = 0$ in F .

Example 3.10. Let ℓ be a line of $\text{PG}(n, q)$ contained in H and $\mathcal{K} = \ell$. Choose a point $v \in \mathcal{K}$. If ℓ_1 and ℓ_2 are any two affine lines with v as point at infinity, and both lying in a plane with ℓ as line at infinity then, in the special case that $q = 0$, the word $\chi_{\ell_1} - \chi_{\ell_2}$ is a codeword of $\mathcal{D}_{\mathcal{K}}$ of weight $2q = 2|\mathcal{K}| - 2$.

Example 3.11. In $\text{PG}(n, 2)$, let $\mathcal{K} = H$ and let w be the sum of the characteristic functions of the points of $P = \text{PG}(n, 2) \setminus H$. That is, the support of w is the set of points of the affine space $\text{AG}(n, 2)$. Every affine line has two points, so when $F = \mathbb{F}_2$ then w is a codeword in $\mathcal{D}_{\mathcal{K}}$ of weight $|\mathcal{K}| + 1$, meeting the bound of [17]. Of course, in this case $\mathcal{D}_{\mathcal{K}} = \{0, w\}$ is not a very interesting code.

4. APPLICATIONS

In this section we retain the general hypotheses of Theorems 1.1 and 3.1 and consider the implications of these theorems in some special cases.

Corollary 4.1. *Assume $q \neq 0$ in F . If $\mathcal{K}$ contains a line of H then N has full rank q^n .*

Proof. This is immediate since every hyperplane of H must meet $\mathcal{K}$. $\square$

Remark 4.2. In the special case $\mathcal{K} = \text{PG}(n-1, q)$ of this corollary, $T_{n-1}^*(\mathcal{K})$ is the 2-design of points and lines in affine space.

Remark 4.3. A *blocking set* in $\text{PG}(2, q)$ is any set of points that meets every line. (Sets containing a line are considered to be trivial examples of blocking sets.) Clearly, by Theorem 1.1, the set $\mathcal{K}$ is a blocking set if and only if N has full rank q^n . Baer subplanes are nontrivial examples.

4.1. Wenger graphs. Let $\mathcal{K} = \{(0 : 1 : u : u^2 : \dots : u^{n-1}) \mid u \in \mathbb{F}_q\}$. Then the bipartite graph having P and L as the bipartition of its vertex set, with adjacency defined by point-line incidence, is called the Wenger graph $W_{n-1}(q)$. Thus the matrix

$$A = \begin{pmatrix} 0 & N \\ N^T & 0 \end{pmatrix}$$

is an adjacency matrix of $W_{n-1}(q)$. This graph has many alternative descriptions. (See [4].)

From Theorem 1.1 we can derive the following formula for $\text{rank}_F N$.

Corollary 4.4. *Let F be a field in which $q \neq 0$. Then $\text{rank}_F N$ is equal to the number of polynomials in $\mathbb{F}_q[X]$ of degree $\leq n-1$ having a root in $\mathbb{F}_q$.*

Proof. As $H = \text{PG}(V)$, a point of $\mathcal{K}$ is a one-dimensional subspace of V of the form $\langle(0, 1, u, u^2, \dots, u^{n-1})\rangle$. A hyperplane of H is defined by a nonzero linear function on V , which we can write in dual coordinates as $\theta = [c_1 : c_2 : \dots : c_n]$. The point lies on the hyperplane if and only if $0 = \theta((0, 1, u, u^2, \dots, u^{n-1})) = c_1 + c_2 u + \dots + c_n u^{n-1}$. Thus, the hyperplane meets $\mathcal{K}$ if and only if the polynomial $c_1 + c_2 X + \dots + c_n X^{n-1}$ has a root in $\mathbb{F}_q$. Since a nonzero scalar multiple of a linear function defines the same hyperplane and a nonzero scalar multiple of a polynomial has the same roots, the corollary now follows from Theorem 1.1. $\square$

Remark 4.5. It is an easy exercise to count polynomials of degree at most $n-1$ having no root in $\mathbb{F}_q$. The number of them is

$$(15) \quad (q-1) \sum_{d=0}^{n-1} \sum_{k=0}^d (-1)^k \binom{q}{k} q^{d-k}.$$

(cf. [4, Lemma 2.2].) It then follows that

$$\text{rank}_F N = q^n - (q-1) \sum_{d=0}^{n-1} \sum_{k=0}^d (-1)^k \binom{q}{k} q^{d-k}$$

Remark 4.6. As pointed out in [2], the incidence system $T_{n-1}^*(\mathcal{K})$ is dual, in the sense of interchanging the roles of points and lines, to the system (actually several isomorphic systems) described in [4]. Of course, dual systems give rise to isomorphic bipartite graphs, so [2] and [4] are studying the same bipartite graphs.

Remark 4.7. In [4], a proposed open problem was to determine the parameters of the linear codes whose Tanner graphs are the Wenger graphs. These correspond to the code we have called $\mathcal{C}$ and its dual. The minimum weight of $\mathcal{C}$ is $2q$, by Theorem 3.1. Our corollary shows that (15) gives the dimension of such a code over a field where $q \neq 0$, as the code is defined as the nullspace of N (or N^T). As mentioned in the Introduction, the dimension could also be deduced by combining Theorem 1.1 with the known multiplicity from [4] of the eigenvalue zero. This is because Theorem 1.1 shows that the rank is the same for all fields where $q \neq 0$ and is in particular equal to the rank in characteristic zero, which, for a symmetric real matrix, is the matrix size minus the algebraic multiplicity of zero. The problem of computing $\text{rank}_F N$ has also been considered in unpublished work of M. Tait and C. Timmons, where the formula for $\text{rank}_F N$ in the case $n = 4$ was correctly conjectured.

Remark 4.8. If $n = 3$, the Wenger graphs coincide with the graphs in [10] and the codes having these as their Tanner graphs are denoted $\text{LU}(3, q)$ and $\text{LU}(3, q)^D$. They were studied in [9] where a conjecture for the dimension of the binary code was made. The conjecture was proved in [13] and the result generalized to other fields in [16], where also the connection to $T_2^*(\mathcal{K})$ was made.

4.2. Hyperovals. Assume $n = 3$, and let q be a power of 2 and $\mathcal{K}$ a hyperoval in $H \cong \text{PG}(2, q)$. Then it is well known that $T_2^*(\mathcal{K})$ is a generalized quadrangle of order $(q-1, q+1)$ ([11, 3.1.3]). By definition of a hyperoval, each line of H meets $\mathcal{K}$ in 0 or 2 points, and so there are $\binom{q+2}{2}$ lines that meet $\mathcal{K}$. Therefore,

$$\text{rank}_F N = 1 + (q-1) \binom{q+2}{2}$$

for any field F that is not of characteristic 2. This rank formula does not depend on the choice of hyperoval $\mathcal{K}$. Note that if we now modify $\mathcal{K}$ by removing any point, then the rank will be unchanged, since certain secant lines through the point just become tangent lines, and the same lines meet $\mathcal{K}$. These ranks were previously computed in [15] by a different method.

Finally, we could drop our assumption of coprime characteristics and consider the $\mathbb{F}_2$ -rank of N for hyperovals. This seems to be an interesting and difficult problem with possible applications in coding theory. Some computer results are tabulated in [15], and the $\mathbb{F}_2$ -ranks will depend on the choice of hyperoval $\mathcal{K}$.

ACKNOWLEDGEMENTS

The authors would like to thank the referees for their helpful suggestions for clarifying several points in the paper.

REFERENCES

- [1] R.W. Ahrens and G. Szekeres, *On a combinatorial generalization of 27 lines associated with a cubic surface*, J. Austral. Math. Soc. **10** (1969), 485–492.
- [2] P. Cara, S. Rottey, G. Van de Voorde, *A construction for infinite families of semisymmetric graphs revealing their full automorphism group*, J. Alg. Combinatorics **39** (2014), 967–988.
- [3] P. Cara, S. Rottey, G. Van de Voorde, *The isomorphism problem for linear representations and their graphs*, Advances in Geometry, **14** (2014), 353–367.
- [4] S. Cioaba, F. Lazebnik, W. Li, *On the spectrum of Wenger graphs*, J. Combinatorial Theory B, **107** (2014), 132–139.
- [5] F. De Clerck, *Een kombinatorische studie van de eindige partiele meetkunden*, PhD Thesis. Ghent University, 1978.
- [6] S. De Winter, S. Rottey, G. Van de Voorde, *linear representations of subgeometries*. Designs, Codes and Cryptography **77** (2015), 203–215.
- [7] R. G. Gallager, *Low-density parity-check codes*, IRE Trans. Inform. Theory, vol. IT-18, pp.21–28, Jan. 1962.
- [8] M. Hall, Jr. *Affine generalized quadrilaterals*, Studies in Pure Mathematics (ed. L. Mirsky), Academic Press, London (1971), 113–116.
- [9] J. L. Kim, U. Peled, I. Perepelitsa, V. Pless and S. Friedland, *Explicit construction of families of LDPC codes with no 4-cycles*, IEEE Trans. Inform. Theory, **50** (2004), 2378–2388.
- [10] F. Lazebnik and V. Ustimenko, *Explicit construction of graphs with arbitrary large girth and of large size*, Discrete Appl. Math. **60** (1997), 275–284.
- [11] S. E. Payne, J. A. Thas, *Finite Generalized Quadrangles*, (1985), Pitman, New York.
- [12] V. Pepe, L. Storme, G. Van de Voorde, *Small weight codewords in the LDPC codes arising from linear representations of geometries*, J. Combin. Des. **17** (1) (2009), 1–24.
- [13] P. Sin, Q. Xiang, *On the dimension of certain LDPC codes based on q -regular bipartite graphs*, IEEE Trans. Inform. Theory **52** no. 8, (2006), 3735–3737.
- [14] H. Tang, J. Xu, Y. Kou, S. Lin, K. Abdel-Ghaffar, *On Algebraic Construction of Gallager and Circulant Low-Density Parity-Check Codes*, IEEE Trans. Inform. Theory **50** no. 6, (2004), 1269–1279.
- [15] P. Vandendriessche, *LDPC codes associated with linear representations of geometries*, Advances in Mathematics of Communications **4** (3) (2010), 405–417.
- [16] P. Vandendriessche, *Some low-density parity-check codes derived from finite geometries*, Designs, Codes and Cryptography **54** (3) (2010), 287–297.

- [17] Y. Kou, S. Lin, and M. P.C. Fossorier, *Low-Density Parity-Check Codes Based on Finite Geometries: A Rediscovery and New Results*, IEEE Trans. Inform. Theory **47** no. 7, (2001), 2711-2736

PETER SIN, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, P. O. BOX 118105, GAINESVILLE FL 32611, USA

Email address: sin@ufl.edu

JULIEN SORCI, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, P. O. BOX 118105, GAINESVILLE FL 32611, USA

Email address: jsorci@ufl.edu

QING XIANG, DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF DELAWARE, NEWARK, DE 19716, USA

Email address: qxiang@udel.edu