Beyond Worst-Case Reconstruction in Deterministic Compressed Sensing

Sina Jafarpour, *Member, IEEE,* Marco F. Duarte, *Member, IEEE,* and Robert Calderbank, *Fellow, IEEE*

***Abstract*—The role of random measurement in compressive sensing is analogous to the role of random codes in coding theory. In coding theory, decoders that can correct beyond the minimum distance of a code allow random codes to achieve the Shannon limit. In compressed sensing, the counterpart of minimum distance is the spark of the measurement matrix, i.e., the size of the smallest set of linearly dependent columns. This paper constructs a family of measurement matrices where the columns are formed by exponentiating codewords from a classical binary error-correcting code of block length M. The columns can be partitioned into mutually unbiased bases, and the spark of the corresponding measurement matrix is shown to be $O(\sqrt{M})$ by identifying a configuration of columns that plays a role similar to that of the Dirac comb in classical Fourier analysis. Further, an explicit basis for the null space of these measurement matrices is given in terms of indicator functions of binary self-dual codes. Reliable reconstruction of k-sparse inputs is shown for k of order $M/\log(M)$ which is best possible and far beyond the worst case lower bound provided by the spark.**

I. Introduction

The central goal of compressed sensing (CS) is to capture a signal using very few measurements. In most work to date, this broader objective is exemplified by the important special case in which the measurement data constitute a vector $f = \Phi\alpha + e$ where Φ is an $M \times N$ matrix called the sensing matrix, $\alpha \in \mathbb{R}^N$ is a k-sparse signal, and $e \in \mathbb{R}^M$ is the additive noise. There are two distinct CS frameworks with different objectives.

Worst-case CS [1], [2]: In the worst-case CS framework, the goal is to recover *every* k-sparse vector α from the corresponding measurement vector f. It is known that certain probabilistic processes generate sensing matrices that support worst-case CS [3]. However, the random sensing framework suffers from storage and computation limitations. As a result, there has been a significant amount of research on designing alternative deterministic matrices for worst-case CS framework over the last few years [4]. Most such constructions rely on the coherence between the columns of the matrix. When the coherence follows the Welch Bound $\mu = O\left(\frac{1}{\sqrt{M}}\right)$, the Gerschgorin Circle Theorem guarantees reconstruction of any k-sparse signal with $k = O(\sqrt{M})$.

Average-case CS [4]–[6]: In many practical applications, including wireless communications and radar, it is not necessary to reconstruct *every* sparse vector [7]. The goal of average-case CS is to recover *most* (in contrast to all) k-sparse vectors. Here, the sparse vector is often modeled to have a uniformly random support and random sign for the k non-zero entries. The average-case CS framework relies on the coherence and the spectral norm of the deterministic sensing matrix. The ideal case is when coherence follows the Welch bound [8], and different measurements are orthogonal. Then, as long as $k = O(M/\log N)$, with high probability a k-sparse vector has a unique sparse representation, and can be efficiently recovered from the compressive measurements [6].

In this paper, we construct an explicit basis for the null space of a large family of deterministic sensing matrices designed for the average-case CS framework (see [5] and the references therein) using the indicator vectors of binary self-dual codes. Characterizing the null space of these matrices makes it possible to investigate and analyze the geometric properties of these matrices more precisely.

More specifically, we introduce a family of deterministic sensing matrices called *the extended Delsarte-Goethals frames* (EDGFs) that hold the following three properties simultaneously: (i) as long as $k \leq c_1\sqrt{M}$ it is possible to recover every k-sparse vector α from the measurement vector $\Phi\alpha$; (ii) there exists a k-sparse vector α with $k = c_2\sqrt{M}$ such that no reconstruction algorithm can uniquely recover α from $\Phi\alpha$; and (iii) it is possible to recover most k-sparse vectors α from the measurement vector $\Phi\alpha$ as long as $k \leq c_3\frac{M}{\log N}$ (which can be much larger than $c_2\sqrt{M}$), where c_1, c_2, and c_3 are fixed constants. The EDGFs meet the coherence-based lower bound on worst-case reconstruction and the order-optimal upper bound on average-case reconstruction.

The rest of the paper is organized as follows. Section II reviews Delsarte-Goethals frames (DGFs). Section III explains the properties of the self-dual codes used in this paper. Sections IV and V characterize the null space of the DGFs and the EDGFs. The experiments are provided in Section VI. Section VII concludes the paper.

II. Background and Notation

A. Worst-case vs. Average-case Compressed Sensing

A vector is k-sparse if it has at most k non-zero entries. The support of a k-sparse vector indicates the positions of its non-zero entries. Let Φ be an $M \times N$ matrix. Then Φ is a tight-frame with redundancy ρ if $\Phi\Phi^\dagger = \rho I_{N\times N}$, where $\Phi^\dagger$ denotes the conjugate transpose of Φ. The spark of the measurement matrix Φ, denoted as spark(Φ), is the size of the smallest set

SJ is with the Multimedia Research Group, Yahoo! Research, email: *sina2jp@yahoo-inc.com*. MD is with the Department of Electrical and Computer Engineering, University of Massachusetts-Amherst, email: *md-uarte@ecs.umass.edu*. RC is with the Department of Computer Science, Duke University, email: *robert.calderbank@duke.edu*. The work is supported in part by ONR under grant N00014-08- 1-1110 and by AFOSR under grant FA 9550-09-1-0551.

of linearly dependent columns of Φ. Let φ_i denote the i^{th} column of Φ. The coherence between the columns of Φ is defined as the maximum inner product between two distinct columns of Φ:

$$\mu_\Phi \doteq \max_{i \neq j} \frac{|\varphi_i^\dagger \varphi_j|}{\|\varphi_i\|_2 \|\varphi_j\|_2}.$$

In this paper, we simplify the analysis by focusing on the noiseless CS problem, and note that it is straightforward to generalize the analysis to include noise. The following two theorems relate the maximum sparsity level k to the parameters of the sensing matrix Φ, so that it is possible to efficiently recover all (respectively most) k-sparse vectors α from $\Phi\alpha$.

Theorem 2.1 (Worst-case CS [9]): Let Φ be a an $M \times N$ sensing matrix with worst-case coherence μ_Φ. Then as long as $k = O(\mu_\Phi^{-1})$, it is possible to efficiently recover every k-sparse vector α from the measurement vector $\Phi\alpha$. In contrast, when $k \geq \frac{\mathrm{spark}(\Phi)}{2}$, there exist multiple k-sparse vectors that are mapped to the same measurement vector, rendering recovery impossible.

Theorem 2.2 (Average-case CS [6], [10]): Let Φ be a an $M \times N$ tight-frame with redundancy $\rho = N/M$ and with worst-case coherence μ_Φ. If $k = O\left(\min\left\{\frac{\mu_\Phi^{-2}}{\log N}, \frac{M}{\log N}\right\}\right)$, then with probability $1 - 1/N$, it is possible to efficiently recover a k-sparse vector α with uniformly random support and uniformly random sign from the measurement vector $\Phi\alpha$.

Throughout this paper, m denotes an integer and $M = 2^m$. Given a vector v with binary entries, we let $v(x)$ denote the entry of v indexed by x. The inner product of two binary vectors u, v is denoted by $u^\top v$.

B. Delsarte-Goethals Frames

The Delsarte-Goethals frames (DGFs) are a class of CS matrices that have been recently introduced by Calderbank, Howard, and Jafarpour [5]. Specifically, let m be an odd positive integer, and r to be an integer smaller than $\frac{m-1}{2}$. Next, let $DG(m,r)$ denote the Delsarte-Goethals set of binary symmetric matrices, as described in [11]. Then, given $M = 2^m$ and $N = M^{r+2}$, the $M \times N$ DGF Φ is constructed from $DG(m,r)$ in the following way. Index the rows of Φ by binary vectors $x \in \mathbb{F}_2^m$ and index the columns of Φ by pairs (P,b), where P ranges over all $2^{(r+1)m}$ binary symmetric matrices $DG(m,r)$ and b ranges over all members of $\mathbb{F}_2^m$. The entries of Φ are given by

$$\varphi_{(P,b)}(x) \doteq \frac{1}{\sqrt{M}} i^{x^\top P x + 2b^\top x},$$

where $i = \sqrt{-1}$. It is easy to see from this description that (i) DGFs are unions of orthonormal bases and (ii) each DGF can be represented as

$$\Phi = \frac{1}{\sqrt{M}}[\mathrm{D}_R\mathrm{H}, \cdots, \mathrm{D}_1\mathrm{H}, \mathrm{H}], \tag{1}$$

where $R = \frac{N-M}{M} = M^{r+1} - 1$, H is the Hadamard matrix, and each D_i is a diagonal matrix whose (x,x) entry is $i^{x^\top \mathrm{P}_i x}$ where P_i is a binary symmetric matrix from the Delsarte-Goethals set $DG(m,r)$. A DGF is a tight-frame with redundancy M^{r+1} and worst-case coherence $\mu_\Phi = \frac{2^r}{\sqrt{M}}$ [5]. As a result, it follows from Theorems 2.1 and 2.2 that as long as $k = O(\sqrt{M}/2^r)$, it is possible to recover every k-sparse vector α from $\Phi\alpha$ using the ℓ_1-minimization method.; moreover, even if $k = O\left(\frac{M}{2^{2r}\log N}\right)$ it is still possible to recover most k-sparse vectors α from $\Phi\alpha$ using the same techniques.

III. Binary Self-Dual Codes

We start by defining a binary self-dual code and explaining some of its properties.

Definition 3.1 (Binary Self-Dual Code): A binary code $\mathcal{C}$ is self-dual if

$$C^\perp \doteq \left\{u : u^\top w = 0 \,\forall\, w \in \mathcal{C}\right\} = \mathcal{C}, \tag{2}$$

Let $\mathcal{C}$ be a self-dual code of length m, and let b be a binary m-tuple vector in the finite field $\mathbb{F}_2^m$. Throughout the paper, by $b \oplus \mathcal{C}$ we mean $\{b \oplus c : c \in \mathcal{C}\}$.

Definition 3.2: Let $\mathcal{C}$ be a binary self-dual code of length m. The binary vector v of length $M = 2^m$ with entries $v(x) = 1$ if $x \in \mathcal{C}$ and $v(x) = 0$ if $x \notin \mathcal{C}$ is called the indicator of $\mathcal{C}$.

A direct calculation, captured in the following lemma, shows that the indicator of a self-dual code can be viewed as the binary counterpart of the Dirac comb in Fourier analysis.

Lemma 3.1: Let $\mathcal{C}$ be a binary self-dual linear code of length m, and let $v \in \{0,1\}^M$ be the indicator of $\mathcal{C}$. Let H be the $M \times M$ Hadamard matrix. Then $\mathrm{H}v = |\mathcal{C}|v$.

Next, we use the vector v to construct a sparse vector in the null space of the matrix $\Phi_0 = \left[\mathrm{I}, \frac{1}{\sqrt{M}}\mathrm{H}\right]$.

Theorem 3.1: Let $\Phi_0 = \left[\mathrm{I}, \frac{1}{\sqrt{M}}\mathrm{H}\right]$ be an $M \times 2M$ matrix generated from concatenating the identity matrix and the normalized Hadamard matrix. Let $\mathcal{C}$ be a binary linear self-dual code with indicator v. Define $v_2 \doteq [-v, v]^\top$. Then

I. Φ_0 is a tight frame with redundancy 2.
II. v_2 is a $2\sqrt{M}$-sparse vector in the null space of Φ_0. Therefore $\mathrm{spark}(\Phi_0) \leq 2\sqrt{M}$.

Proof: We prove each part separately:

I. Φ_0 is a union of two orthonormal bases, therefore $\Phi\Phi^\dagger = 2\mathrm{I}_{M\times M}$.
II. Every self-dual code of length m has dimension $\frac{m}{2}$, and hence exactly $\sqrt{M}$ different codewords. Therefore, v is $\sqrt{M}$-sparse, and by construction, v_2 has exactly $2\sqrt{M}$ non-zero entries. Moreover, it follows from Lemma 3.1 that $\mathrm{H}v = \sqrt{M}\,\mathrm{I}v$ or equivalently $\left[\mathrm{I}, \frac{1}{\sqrt{M}}\mathrm{H}\right][-v, v]^\top = 0$.

■

Corollary 3.1: Let D be an $M \times M$ diagonal matrix, and let $\Phi_1 = \frac{1}{\sqrt{M}}[\mathrm{DH}, \mathrm{H}]$ be an $M \times 2M$ matrix generated from concatenating the modulated Hadamard (DH) matrix and the Hadamard (H) matrix. Define $v_\mathcal{N} \doteq \left[-\frac{1}{\sqrt{M}}\mathrm{H}\left(\mathrm{D}^{-1}v\right); v\right]$. Then $v_\mathcal{N}$ is in the null space of Φ_1.

Proof: Theorem 3.1 guarantees that $\frac{1}{\sqrt{M}}\mathrm{H}\,v - \mathrm{I}\,v = 0$. On the other hand, since the Hadamard matrix has orthogonal binary-valued rows, we have $H^{-1} = \frac{1}{M}H$, and therefore

$$\mathrm{H}\,v - \mathrm{DH}\left(\frac{1}{\sqrt{M}}\mathrm{H}D^{-1}v\right) = \mathrm{H}\,v - \sqrt{M}\mathrm{DH}\,(\mathrm{DH})^{-1}\,v = 0.$$

■

Finally, we consider the sparsity degree of $v_{\mathcal{N}}$. Since v is $\sqrt{M}$-sparse, only $\sqrt{M}$ of the second M entries of $v_{\mathcal{N}}$ are non-zero. Now we analyze the first M entries of $v_{\mathcal{N}}$. Let Δ denote the diagonal of D^{-1}, let $\xi \doteq \mathrm{D}^{-1}v$, and denote $\eta \doteq \mathrm{H}\xi$. Here we focus on a special but important case where there exists an $m \times m$ binary symmetric matrix P such that $\Delta(x) = i^{-x^\top \mathrm{P}x}$ for every $x \in \mathbb{F}_2^m$. This is the case for a large class of CS matrices, including the DGFs.

For every binary m-tuple x we have

$$\xi(x) = \begin{cases} \Delta(x) & \text{if } x \in \mathcal{C} \\ 0 & \text{otherwise} \end{cases},$$

and $\eta(x) =$

$$\sum_{y\in\mathbb{F}_2^m}(-1)^{x^\top y}\xi(y) = \sum_{y\in\mathcal{C}}(-1)^{x^\top y}\Delta(y) = \sum_{y\in\mathcal{C}}(-1)^{x^\top y}i^{-y^\top \mathrm{P}y}.$$

Note that the calculation $y^\top \mathrm{P}y + 2x^\top y$ is now over Z_4 and not Z_2. Let $\overline{\eta(x)}$ denote the complex conjugate of $\eta(x)$. Observe that $\eta(x)$ is zero if and only if $\overline{\eta(x)}$ is zero. Therefore, it is sufficient to analyze $\overline{\eta(x)}$.

Theorem 3.2: Let P be an $m \times m$ binary symmetric matrix, and let E be the null space of P. Define

$$\overline{\eta(x)} \doteq \sum_{y\in\mathcal{C}} i^{y^\top \mathrm{P}y + 2x^\top y}, \text{ where } x \in \mathbb{F}_2^m.$$

Let d_{P} denote the diagonal of P and assume that $d_{\mathrm{P}} \in \mathcal{C}$. This assumption is easily satisfied if we only consider zero-diagonal matrices in the construction of the DGFs. Define

$$\mathcal{C}_0 \doteq \{z \in \mathcal{C} : \mathrm{P}z \in \mathcal{C}\}. \tag{3}$$

Then $\overline{\eta(x)} \neq 0$ for at most 2^t values of x, where $t = m - \dim(\mathcal{C}_0)$.

Proof: We have

$$\begin{aligned}\overline{\eta(x)}^2 &= \sum_{y,y'\in\mathcal{C}} i^{y^\top \mathrm{P}y + y'^\top \mathrm{P}y' + 2x^\top(y+y')} \\ &= \sum_{y,y'\in\mathcal{C}} i^{(y+y')^\top \mathrm{P}(y+y') + 2y^\top \mathrm{P}y' + 2x^\top(y+y')}.\end{aligned}$$

Changing variables to $z = y + y'$ and y gives

$$\begin{aligned}\overline{\eta(x)}^2 &= \sum_{z\in\mathcal{C}} i^{z^\top \mathrm{P}z + 2x^\top z}\sum_{y\in\mathcal{C}}(-1)^{(z+y)^\top \mathrm{P}y} \\ &= \sum_{z\in\mathcal{C}} i^{z^\top \mathrm{P}z + 2x^\top z}\sum_{y\in\mathcal{C}}(-1)^{(d_{\mathrm{P}}+\mathrm{P}z)^\top y}.\end{aligned} \tag{4}$$

The inner sum is zero if $(d_{\mathrm{P}} + \mathrm{P}z) \notin \mathcal{C}$. Otherwise the inner sum is $|\mathcal{C}|$, and we have

$$\overline{\eta(x)}^2 = |\mathcal{C}| \sum_{z\in\mathcal{C}_0} i^{z^\top \mathrm{P}z + 2x^\top z}.$$

Now observe that for any $z_1, z_2 \in \mathcal{C}_0$,

$$\begin{aligned}&(z_1+z_2)^\top \mathrm{P}(z_1+z_2) + 2x^\top(z_1+z_2) \\ &= z_1^\top \mathrm{P}z_1^\top + 2x^\top z_1 + z_2^\top \mathrm{P}z_2 + 2x^\top z_2 + 2z_1^\top \mathrm{P}z_2 \\ &= z_1^\top \mathrm{P}z_1^\top + 2x^\top z_1 + z_2^\top \mathrm{P}z_2 + 2x^\top z_2 \,(\mathrm{mod}\,4),\end{aligned} \tag{5}$$

where the second equality follows from the fact that both $\mathrm{P}z_1$ and z_2 are codewords of $\mathcal{C}$. If this linear map is the zero map on x, then $|\overline{\eta(x)}|^2 = |\mathcal{C}||\mathcal{C}_0|$, and otherwise $\overline{\eta(x)} = 0$. As a result, $\eta(x)$ vanishes for all but 2^t values of x, where $t = m - \dim(\mathcal{C}_0)$. ■

IV. The Null Space of Delsarte-Goethals Frames

In this section, we construct a basis for the null space of matrices of the form of Equation (1). To do this, we first provide an orthogonal basis for the null space of the matrix $\Phi_0 \doteq \left[I, \frac{1}{\sqrt{M}}\mathrm{H},\right]$.

Let a be a binary m-dimensional vector. The time-shift matrix A_a is a circulant matrix so that every row x of A_a has only one 1 at the corresponding column indexed at $x \oplus a$ and zeros elsewhere. Similarly, the frequency-shift matrix B_a is a diagonal matrix with diagonal entries $(-1)^{xa^\top}$, where the m-dimensional binary vector x ranges over all $M = 2^m$ rows. A direct calculation reveals that $HA_a = B_aH$, and $HB_a = A_aH$.

Lemma 4.1: Let Φ_0 and v be as above, and let a and b be any two binary m-dimensional offsets. Then the vector $w_2 = [A_aB_bv; -B_aA_bv]$ is in the null space of Φ_0.

Proof: We have $\Phi_0 w_2 =$

$$\begin{aligned}&A_aB_bv - \frac{1}{\sqrt{M}}\mathrm{H}B_aA_bv = A_aB_bv - A_aB_b\frac{1}{\sqrt{M}}\mathrm{H}v \\ &= A_aB_b\left(\left[\mathrm{I}, \frac{1}{\sqrt{M}}\mathrm{H}\right][v; -v]\right) = 0.\end{aligned} \tag{6}$$

■

Lemma 4.2: Let

$$W = \{w_2 : w_2 = [A_aB_bv; -B_aA_bv], a, b \in \mathbb{F}_2^m\}.$$

Then W is an orthogonal basis forming the null space of Φ_0.

Proof: Since v is the indicator of a self dual code, it is $\sqrt{M}$-sparse. Moreover, there are exactly $\sqrt{M}$ offsets a that produce distinct shifts of v. The reason is that since $\mathcal{C}$ is linear, if c is any codeword of $\mathcal{C}$, then both A_av and $A_{a\oplus c}v$ provide the same vector whose non-zero entries correspond to indices of the form $z \oplus a$ where z ranges over all codewords. That is $A_av = A_{a\oplus c}v$. Similarly, since $\mathcal{C}$ is self-dual, for every pair of codewords z and c we have $(-1)^{b^\top z} = (-1)^{(b\oplus c)^\top z}$. This implies that there are $\sqrt{M}$ distinct choices b for the frequency shift of the vector v.

Now we show that these M vectors are orthogonal. To see this, let (a, b) and (a', b') be two distinct pairs of time and frequency shift offsets. Let $\xi_1 = B_aA_bv$ and $\xi_2 = B'_aA'_bv$. Then, it is sufficient to show that if ξ_1 and ξ_2 are distinct, then $\xi_1 \top \xi_2 = 0$.

Since v is the indicator vector of a linear self-dual code, then if $\mathrm{Supp}(\xi_1) \cap \mathrm{Supp}(\xi_2)$ contains some element y, then

the set $y \oplus \mathcal{C}$ is also in the intersection of the two supports. This set has $\sqrt{M}$ elements, and since both supports also have $\sqrt{M}$ elements, we must have $\mathrm{Supp}(\xi_1) = \mathrm{Supp}(\xi_2)$. As a result, if $A_b v \neq A'_b v$ then $\mathrm{Supp}(\xi_1) \cap \mathrm{Supp}(\xi_2) = \emptyset$, and therefore $B_a A_b v$ and $B'_a A'_b v$ are orthogonal.

On the other hand, if $A_b v = A'_b v$, then ξ_1 and ξ_2 are two distinct Walsh tones restricted to the set $b \oplus \mathcal{C}$. Now let $\tilde{c}$ be an element in $\mathcal{C}$ such that $\tilde{c}^\top (a \oplus a') = 1$. Then

$$(-1)^{\tilde{c}^\top (a\oplus a')} \left(\sum_{x: x\oplus b \in \mathcal{C}} (-1)^{x^\top (a \oplus a')} \right) \tag{7}$$
$$= \left(\sum_{x: x\oplus b \in \mathcal{C}} (-1)^{(x\oplus \tilde{c})^\top (a\oplus a')} \right) = \sum_{x: x\oplus b \in \mathcal{C}} (-1)^{x^\top (a\oplus a')}.$$

Therefore, we must have $\sum_{x: x\oplus b\in\mathcal{C}} (-1)^{x^\top (a\oplus a')} = 0$, which proves that ξ_1 and ξ_2 are orthogonal. A similar argument can be used to show that if $\xi_1 = A_a B_b v$ and $\xi_2 = A'_a B'_b v$, then, if ξ_1 and ξ_2 are distinct, then $\xi_1 \top \xi_2 = 0$.

As a result, we have M distinct linearly independent vectors of the form $w_2 = [A_a B_b v; -B_a A_b v]$ which are all in the null space of Φ_0. On the other hand, since Φ_0 is $M \times 2M$, its null space has dimension M, and therefore the M null space vectors above form a basis for the null space of Φ_0. ∎

Next we show that the same argument can be used to form an orthogonal basis for the null space of matrices of form $\Phi_1 = \frac{1}{\sqrt{M}} [\mathrm{DH}, \mathrm{H}]$, where D is a diagonal matrix with diagonal entries $\left[i^{x \mathrm{P} x^\top} \right]$, $x \in \mathbb{F}_2^m$.

Theorem 4.1: Let $\Phi_1 = \frac{1}{\sqrt{M}} [\mathrm{DH}, \mathrm{H}]$, where $\mathrm{D} = [i^{x \mathrm{P} x^\top}]$, and where P is a zero-diagonal binary symmetric matrix[1]. Let $\mathcal{C}$ be a self-dual code such that for any codeword c, $\mathrm{P}c$ is also a codeword in $\mathcal{C}$. Let v denote the indicator vector of the codewords of $\mathcal{C}$. Let

$$W2 = \left\{ w_2 : w_2 = \left[\frac{\mathrm{HD}^{-1}}{\sqrt{M}} A_a B_b v; B_a A_b v \right], a, b \in \mathbb{F}_2^m \right\}.$$

Then $W2$ forms an orthonormal basis for the null space of the matrix Φ_1; moreover, every element of $W2$ is $2\sqrt{M}$-sparse.

Proof: Let w be any vector in $W2$. We have

$$\Phi_1 w = \frac{1}{\sqrt{M}} [\mathrm{DH}, \mathrm{H}] \left[\frac{\mathrm{HD}^{-1}}{\sqrt{M}} A_a B_b v; B_a A_b v \right] \tag{8}$$
$$= [A_a B_b v - \frac{1}{\sqrt{M}} \mathrm{H} B_a A_b v] = 0.$$

Now we show that any vector $w_2 \in W2$ is $2\sqrt{M}$-sparse. First, note that since v is the indicator of a self-dual code, and the operators A_b and B_a do not change the sparsity level, $B_a A_b v$ is $\sqrt{M}$-sparse. Hence, we need to show that $\omega \doteq \mathrm{HD}^{-1} A_a B_b v$ is also $\sqrt{M}$-sparse. We have

$$\omega(x) = \sum_{y\in\mathcal{C}} (-1)^{b^\top y} i^{(y\oplus a)^\top P (y\oplus a)} (-1)^{(y\oplus a)^\top x} \tag{9}$$
$$= i^{a^\top P a} (-1)^{a^\top x} \sum_{y\in\mathcal{C}} i^{2(b+x+Pa)^\top y + y^\top P y}.$$

Since $\mathcal{C}$ is self-dual and y and Py are both codewords in $\mathcal{C}$, the mapping $y \to 2(b+x+aP)^\top y + y^\top P y$ is a linear map from $\mathcal{C}$ to $\mathbb{Z}_4$. Theorem 3.2 now guarantees that $w(x)$ is non-zero only for $\sqrt{M}$ choices of x. ∎

Remark 4.1: So far we have shown how to generate a basis for the null space of a $M \times 2M$ matrix of the form $[\mathrm{DH}, \mathrm{H}]$. This construction can be generalized to matrices of form $[\mathrm{D}_R \mathrm{H}, \cdots, \mathrm{D}_1 \mathrm{H}, \mathrm{H}]$ in a straightforward manner by first zero-padding each null space vector, so that it becomes $(R+1)M$-dimensional, and then collecting all these RM vectors. For instance, a basis for the null space of a $M \times 3M$ matrix $[\mathrm{D}_2\mathrm{H}, \mathrm{D}_1\mathrm{H}, \mathrm{H}]$ has the form

$$\begin{bmatrix} 0 & V'_2 \\ V'_1 & 0 \\ V_1 & V_2 \end{bmatrix}, \tag{10}$$

where $[V'_1; V_1]$ is a basis for the null space of $[\mathrm{D}_1\mathrm{H}, \mathrm{H}]$, and $[V'_2; V_2]$ is a basis for the null space of $[\mathrm{D}_2\mathrm{H}, \mathrm{H}]$.

V. Extended Delsarte-Goethals Frames

In this section, we use the results of the previous sections and design an $M \times N$ sensing matrix with $M \ll N$ for which there exists constants c_1, c_2, and c_3, such that (i) it is possible to uniquely recover every k-sparse vectors as long as $k \leq c_1\sqrt{M}$; (ii) there exists a k-sparse vector α with $k = c_2\sqrt{M}$ such that no reconstruction algorithm is able to uniquely recover it from the measurement vector $f = \Phi\alpha$; and (iii) it is possible to uniquely recover most k-sparse vectors as long as $k \leq c_3 M / \log N$.

The extended Delsarte-Goethals frame (EDGF) is constructed by concatenating the $M \times M$ identity matrix to a DGF.

Theorem 5.1: Let r be a constant integer, and let $m > 2r+1$ be an integer. Then:

I. The EDGF has $M = 2^m$ rows, and $N = M^{r+2} + M$ columns

II. The EDGF is a tight frame with redundancy $M^{r+1}+1$, and coherence $\frac{2^r}{\sqrt{M}}$.

Proof: The proof of Part I follows from the construction of the EDGF. To prove Part II. observe that the EDGF is a union of orthonormal bases, and therefore it is a tight-frame with redundancy $\frac{N}{M}$. The coherence bound follows from the fact that the inner product between two distinct columns of a DGF is bounded by $\frac{2^r}{\sqrt{M}}$, and that the inner product between a column of a DGF, and a column of the identity matrix is bounded by $\frac{1}{\sqrt{M}}$. ∎

Since the matrix $\Phi_0 = [\mathrm{I}, \frac{1}{\sqrt{M}}\mathrm{H}]$ is a submatrix of any EDGF, it follows from Theorem 3.1 that the spark of any

[1]The zero-diagonal assumption is only for simplifying the calculations, and is easy satisfied by using the Gray map [12].

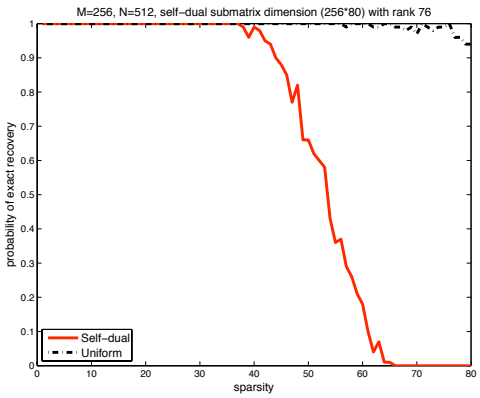

Fig. 1: Probability of exact recovery of the Basis Pursuit algorithm in recovering sparse vectors using a 256×512 submatrix of the $DG(8,0)$ frame. The matrix has a 256×80 rank-deficient submatrix with rank 76.

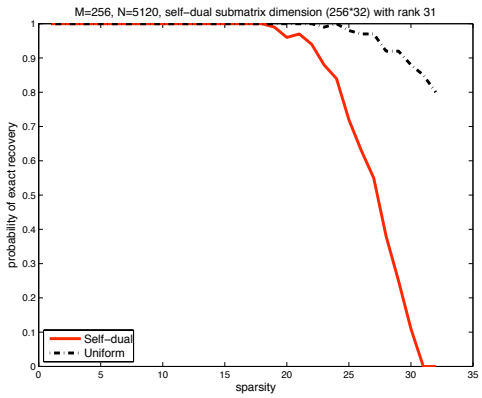

Fig. 2: Probability of exact recovery of the Basis Pursuit algorithm in recovering sparse vectors using a 256×5120 submatrix of the extended $DG(8,0)$ frame. The matrix has a 256×32 rank-deficient submatrix with rank 31.

EDGF is at most $2\sqrt{M}$. As a result, there can be two distinct $\sqrt{M}$-sparse vectors mapped to the same low-dimemsional measurement vector. On the other hand, Theorem 2.1 states that it is possible to uniquely recover every k-sparse vector with sparsity $k \leq c_1\sqrt{M}$. In contrast, it follows from Theorem 2.2 that if α is a k-sparse vector with uniformly random support and random sign, then as long as $k \leq c_3 M/\log N$ (which can be much larger than $c_1\sqrt{M}$), α is uniquely recoverable from $f = \Phi\alpha$ with overwhelming probability.

VI. Experimental Results

The experiments of this section compare the performance of the Basis Pursuit algorithm [6] in recovering sparse vectors with uniformly random supports versus recovering sparse vectors supported on a rank-deficient submatrix obtained using a self-dual code. We provide the comparisons for the DGFs (Fig. 1), as well as for the EDGFs (Fig. 2).

In Fig. 1 we used a 256×512 matrix of the form $\Phi = \frac{1}{\sqrt{M}}[\mathrm{H}, \mathrm{D}_1\mathrm{H}]$, where D_1 is a diagonal matrix with $d_{\mathrm{D}_1}(x) = i^{x^\top \mathrm{P} x}$, and P is the 8×8 zero-diagonal binary symmetric matrix obtained from applying the Gray map [12] to a 7×7 binary $DG(7,0)$ matrix. We also used the self-dual Hamming code $\mathcal{C}$ of length 8 in order to find a rank-deficient submatrix of Φ.

A simple calculation reveals that only 4 codewords are in $\mathcal{C}_0$ (defined by eq. (3)). Therefore, $\dim(\mathcal{C}_0) = 2$, and Theorem 3.2 predicts that $\mathrm{HD}_1^{-1}v$ must have $256/4 = 64$ non-zero entries. A direct calculation reveals that $\mathrm{HD}_1^{-1}v$ indeed has 64 non-zero entries. Therefore, the null space of $\frac{1}{\sqrt{M}}[\mathrm{H}, \mathrm{D}_1\mathrm{H}]$ has a 80-sparse element. That is, there exists, a 256×80 submatrix of Φ that is rank-deficient. As illustrated in Fig. 1, the Basis Pursuit Algorithm fails to recover some k-sparse vectors with sparsity level $k > 40$ which are supported on this rank-deficient submatrix. However, it is still possible to recover most k-sparse vectors with uniformly random support over the 512 columns, even for sparsity level $k = 80$. Fig. 2 shows the result of the same experiment using an EDGF.

VII. Conclusion

We have determined a natural basis for the null space of an extended Delsarte-Goethals frame and shown that this null space contains a vector that is $2\sqrt{M}$-sparse. We have demonstrated that this family of measurement matrices meets the lower bound of $k = O(\sqrt{M})$ on worst-case CS as well as the order optimal upper bound of $k = O(M/\log(N))$.

References

[1] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics, Vol. 59 (8) , pp. 1207-1223.*, 2006.

[2] D. Donoho. Compressed Sensing. *IEEE Transactions on Information Theory, Vol. 52 (4), pp. 1289-1306*, April 2006.

[3] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation, Vol 28 (3), pp. 253-263*, December 2008.

[4] W. Bajwa, R. Calderbank, and S. Jafarpour. Model Selection: Two fundamental measures of coherence and their algorithmic significance. *Proceedings of IEEE Symposium on Information Theory (ISIT)*, 2010.

[5] R. Calderbank, S. Howard, and S. Jafarpour. Construction of a large class of matrices satisfying a statistical isometry property. *IEEE Journal of Selected Topics in Signal Processing, Special Issue on Compressive Sensing, Vol. 4(2), pp. 358-374*, 2010.

[6] E. Candès and J. Plan. Near-ideal model selection by ℓ_1 minimization. *Annals of Statistics, Vol. 37, pp. 2145-2177*, 2009.

[7] L. Applebaum, S. Howard, S. Searle, and R. Calderbank. Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery. *Applied and Computational Harmonic Analysis, Vol. 26 (2), pp. 283-290*, March 2009.

[8] W. Bajwa, J. Haupt, G. Raz, S. Wright, and R. Nowak. Toeplitz-structured compressed sensing matrices. *Statistical Signal Processing. IEEE/SP 14th Workshop on*, pages 294–298, August 2007.

[9] D. L. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization. *Proc. Nat. Acad. Sci.*, 100(5):2197–2202, 2003.

[10] J. Tropp. On The Conditioning of Random Subdictionaries. *Applied and Computational Harmonic Analysis, Vol. 25, pp. 124*, 2008.

[11] R. Calderbank and S. Jafarpour. Reed-Muller sensing matrices and the LASSO. *International Conference on Sequences and their Applications (SETA), pp. 442-463*, 2010.

[12] A.R. Calderbank, P.J. Cameron, W.M. Kantor, and J.J. Seidel. $\mathbb{Z}_4$–Kerdock Codes, orthogonal spreads and extremal euclidean line sets. *Proceedings of London Math. Society, vol. 75, pp. 436-480*, 1997.