Linear codes with disjoint repair groups

Mary Wootters

February 28, 2016

Let $C \subseteq \mathbb{F}^n$ be a linear code; we are often interested in the locality of $C$. This can be quantified in several ways, but one way which has recently been fruitful in several applications is by measuring the number of disjoint repair groups for any given symbol.

**Definition 1.** Let $C \subseteq \mathbb{F}^n$ be a linear code. We say that $C$ has the $t$-disjoint-repair-group property for $s$ symbols ($(t,s)$-DRGP) if the following holds. For all $i \in [s]$, there are vectors $\lambda^{(1)}, \ldots, \lambda^{(t)} \in \mathbb{F}^n$ so that:

1. The sets $\text{Supp}(\lambda^{(r)})$ and $\{i\}$ are disjoint for $r = 1, \ldots, t$, and
2. for all $c \in C$, and for all $r = 1, \ldots, t$,
   $$c_i = \sum_{j=1}^{n} \lambda^{(r)}_j c_j$$

That is, for any $i \in [s]$ and any $c \in C$, $c_i$ can be recovered in $t$ different ways (other than looking at $c_i$ itself), each of which relies on a disjoint set of indices. Notice that if $s = \text{dim}(C)$, this gives recovery of the systematic symbols, and for $s = n$, this gives recovery for all symbols.

When $t = \Omega(n)$, this property is enough to give a constant-query LDC/LCC. This property is also useful in distributed storage, and is related to batch codes, as in [DGRS14]. When $t$ is small, this property has been shown to be useful for PIR codes, which reduce the amount of storage overhead required in a PIR scheme [FVY15].

In [FVY15], constructions of codes with the $(t,s)$-DRGP are given, for $t \leq \sqrt{n}$. For constant $t$, these constructions give codes $C$ which have $t$ disjoint repair groups, with

$$\text{dim}(C) \geq n - O(\sqrt{n}).$$

It is asked in that work whether this bound is tight. In this note, we show that it is for $t = 2$.

**Lemma 1.** Let $C \subseteq \mathbb{F}^n$ be a linear code with the $(2,s)$-DRGP, and let $\ell = n - \text{dim}(C)$ be the redundancy. Then

$$2s \leq (\ell + 1) \cdot \ell.$$

**Proof.** Let $C$ and $\ell$ be as in the statement. Consider the dual code $C^\perp$. This is a linear code of dimension $\ell$ and length $n$; this implies that there is some set $\Omega = \{\omega_1, \ldots, \omega_n\} \subseteq \mathbb{F}^\ell$ so that

$$C^\perp = \{\langle \alpha, \omega_1 \rangle, \langle \alpha, \omega_2 \rangle, \ldots, \langle \alpha, \omega_n \rangle : \alpha \in \mathbb{F}^\ell \}.$$ 

In this language, the $(2,s)$-DRGP is that for all $i \in [s]$, there exist some $\alpha_i, \beta_i \in \mathbb{F}^\ell$ so that

- $\langle \alpha_i, \omega_i \rangle, \langle \beta_i, \omega_i \rangle \neq 0$, and
- for all $j \neq i$, $\langle \alpha_i, \omega_j \rangle \cdot \langle \beta_i, \omega_j \rangle = 0$. 

For $i \in [s]$, define polynomials $P_i : F^\ell \to F$ by
\[
P_i(X_1, \ldots, X_\ell) = \left( \sum_{j=1}^\ell \alpha_i[j] X_j \right) \cdot \left( \sum_{j=1}^\ell \beta_i[j] X_j \right).
\]
Now the above implies that
\[
P_i(\omega_i) = \langle \alpha_i, \omega_i \rangle \langle \beta_i, \omega_i \rangle \neq 0
\]
and for all $j \neq i$,
\[
P_i(\omega_j) = 0.
\]
Thus, the $P_i$’s are linearly independent over $F$. However, they are spanned by the monomials of degree exactly two in $X_1, \ldots, X_\ell$. But there are $\binom{\ell+1}{2}$ of these, and so
\[
s \leq \binom{\ell+1}{2},
\]
aka
\[
2s \leq (\ell + 1) \cdot \ell,
\]
as claimed. \qed

Notice that when $s = \dim(C)$, this gives the systematic result, and when $s = n$, this gives the non-systematic lower bound.

References
