Linear codes with disjoint repair groups

Mary Wootters

February 28, 2016

Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code; we are often interested in the *locality* of $\mathcal{C}$. 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 $\mathcal{C} \subset \mathbb{F}^n$ be a linear code. We say that $\mathcal{C}$ 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 $\mathrm{Supp}(\lambda^{(r)})$ *and $\{i\}$ are disjoint for $r = 1, \ldots, t$, and*

(2) for all $c \in \mathcal{C}$, and for all $r = 1, \ldots, t$,

$$c_i = \sum_{j=1}^{n} \lambda_j^{(r)} c_j$$

That is, for any $i \in [s]$ and any $c \in \mathcal{C}$, 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 $s = \dim(\mathcal{C})$, 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 $\mathcal{C}$ which have t disjoint repair groups, with

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

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

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

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

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

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 $P_i : \mathbb{F}^\ell \to \mathbb{F}$ 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 $\mathbb{F}$. 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. □

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

References

[DGRS14] Alexandros G Dimakis, Anna Gal, Ankit Singh Rawat, and Zhao Song. Batch codes through dense graphs without short cycles. *arXiv preprint arXiv:1410.2920*, 2014.

[FVY15] Arman Fazeli, Alexander Vardy, and Eitan Yaakobi. Pir with low storage overhead: Coding instead of replication. *arXiv preprint arXiv:1505.06241*, 2015.