Linear Locally Decodable Codes **Fall 2016**

Lecture 5: Matching Vector Codes

Lecturer: Zeev Dvir *Scribe: Kalina Petrova*

In this lecture we discuss Locally Decodable Codes that arise from families of Matching Vectors, to be defined below. We start with some examples that can help the reader better appreciate the useful properties of Matching Vector families.

Example 5.1. Let $\omega_m = e^{\frac{2\pi i}{m}}$, the primitive root of unity of order m in $\mathbb{C}$. Then we have that $\omega_m^k = 1$ iff k is a multiple of m (since $e^{2\pi i} = 1$). Consider, for some $u \in \{0, 1, \ldots, m-1\}$, the "wave function with frequency u":

$$\varphi_u : \mathbb{Z}_m \to \mathbb{C}$$
$$\varphi_u(x) = \omega_m^{ux}$$

Take any $c \in \mathbb{Z}_m$ and "direction" $v \in \mathbb{Z}_m$ and sum φ_u along the "line in direction v" $\{c + yv | y \in \mathbb{Z}_m\}$:

$$\sum_{y \in \mathbb{Z}_m} \varphi_u(c + yv) = \sum_{y \in \mathbb{Z}_m} \omega_m^{uc+uyv} = \begin{cases} \sum_{y \in \mathbb{Z}_m} \omega_m^{uc} = \omega_m^{uc} m & \text{if } uv = 0 \mod m \\ \omega_m^{uc} \frac{(\omega_m^{uv})^m - 1}{\omega_m^{uv} - 1} = 0 & \text{otherwise} \end{cases}$$

Example 5.2. Suppose we have two bits of data $A_0, A_1 \in \{0, 1\}$. We will show here one way to encode these bits as a linear combination of wave functions. Suppose $u_0, u_1, v_0, v_1 \in \mathbb{Z}_m$ and $u_i v_j = 0 \mod m$ if and only if $i = j$. Then we can encode A_0 and A_1 with the function $f : \mathbb{Z}_m \to \mathbb{C}$, defined in the following way: $f(x) = A_0 \varphi_{u_0}(x) + A_1 \varphi_{u_1}(x)$. Now if can query the function f for different values, we can find A_0 by summing f over any line of the form $\{c + yv_0 | y \in \mathbb{Z}\}$ for any fixed $c \in \mathbb{Z}_m$:

$$\sum_{y \in \mathbb{Z}_m} f(c + yv_0) =$$
$$\sum_{y \in \mathbb{Z}_m} A_0 \varphi_{u_0}(c + yv_0) + A_1 \varphi_{u_1}(c + yv_0) =$$
$$A_0 \omega_m^{u_0 c} m,$$

where the last equality follows from what we derived in Example 5.1.

Example 5.3. We are going to extend Example 5.1 to a higher dimension to see a property analogical to the one discussed above. Take $\mathbf{u} \in \mathbb{Z}_m^\ell$ (the frequency of the wave function), and let the wave function $\varphi_{\mathbf{u}} : \mathbb{Z}_m^\ell \to \mathbb{C}$ be defined as $\varphi_{\mathbf{u}}(\mathbf{x}) = \omega_m^{<\mathbf{x},\mathbf{u}>}$. Then if we sum $\varphi_{\mathbf{u}}$ over a line $\{\mathbf{c} + y\mathbf{v} | y \in \mathbb{Z}_m\}$, we get

$$\sum_{y \in \mathbb{Z}_m} \varphi_{\mathbf{u}}(\mathbf{c} + y\mathbf{v}) = \sum_{y \in \mathbb{Z}_m} \omega_m^{<\mathbf{c}+y\mathbf{v},\mathbf{u}>} =$$

$$\omega_m^{<\mathbf{c},\mathbf{u}>} \sum_{y \in \mathbb{Z}_m} \left(\omega_m^{<\mathbf{v},\mathbf{u}>}\right)^y = \begin{cases} \omega_m^{<\mathbf{c},\mathbf{u}>} m & \text{if } <\mathbf{v},\mathbf{u}>= 0 \mod m \\ \omega_m^{<\mathbf{c},\mathbf{u}>} \frac{(\omega_m^{<\mathbf{v},\mathbf{u}>})^m - 1}{\omega_m^{<\mathbf{v},\mathbf{u}>} - 1} = 0 & \text{otherwise} \end{cases}$$

Definition 5.1. A Matching-Vector (MV) family in $\mathbb{Z}_m^\ell$ of size k is given by two lists of vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_k) \in \mathbb{Z}_m^\ell$ such that $\forall i \in [k], <\mathbf{u}_i, \mathbf{v}_i>= 0 \mod m$ and $\forall i \neq j, <\mathbf{u}_i, \mathbf{v}_j> \neq 0 \mod m$.

Theorem 5.1 ([Yek12], [Efr09])**.** If there is an MV-family in $\mathbb{Z}_m^\ell$ of size k, then there exists an (m, δ, ϵ)-LDC $E : \mathbb{F}_q^k \to \mathbb{F}_q^{m^\ell}$, where $\delta = \frac{1}{4m}, \epsilon = \frac{1}{4}$, and $\mathbb{F}_q$ is any field such that $q - 1$ is divisible by m.

Proof. In this proof, we are going to use the following fact from Number Theory.

Claim 5.1 ([Maz03])**.** If $q - 1$ is divisible by m, then $\mathbb{F}_q$ contains an m-th root of unity ω (that is, $\omega^m = 1$ and $\forall m' \in [m-1], \omega^{m'} \neq 1$).

Let $(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_k)$ be a Matching Vector family in $\mathbb{Z}_m^\ell$. To construct the Locally Decodable Code, for any message $\mathbf{a} \in \mathbb{F}_q^k$, define the function $F_{\mathbf{a}}(\mathbf{x}) = \sum_{i=1}^k \mathbf{a}_i \omega^{<\mathbf{x},\mathbf{u}_i>}$. Then define $E(\mathbf{a}) = \left(F_{\mathbf{a}}(\mathbf{x})\right)_{\mathbf{x} \in \mathbb{Z}_m^\ell}$. Notice that E is a linear map with a $m^\ell \times k$ generating matrix G of the following form: if the rows of the matrix are indexed by the elements of $\mathbb{Z}_m^\ell$, and the columns are indexed from 1 to k, then $G_{\mathbf{x},i} = \omega^{<\mathbf{u}_i,\mathbf{x}>}$, where $\mathbf{x} \in \mathbb{Z}_m^\ell$ and $i \in [k]$.

We are going to use the m-th root of unity to construct a local decoder for E. Given $E(\mathbf{a})$, to recover $\mathbf{a}_j$, we pick a uniform $\mathbf{c} \in \mathbb{Z}_m^\ell$. Next, we query the following m positions in $E(\mathbf{a})$: $\left(F_{\mathbf{a}}(\mathbf{c} + y\mathbf{v}_j)\right))_{y=0,1,\ldots,m-1}$. We sum over the results of these queries and we get

$$\sum_{y=0}^{m-1} F_{\mathbf{a}}(\mathbf{c}+y\mathbf{v}_j) =$$
$$\sum_{y=0}^{m-1}\sum_{i=1}^{k} \mathbf{a}_i \omega^{<\mathbf{c}+y\mathbf{v}_j,\mathbf{u}_i>} =$$
$$\sum_{i=1}^{k} \mathbf{a}_i \omega^{\mathbf{c},\mathbf{u}_i} \sum_{y=0}^{m-1} \left(\omega^{<\mathbf{v}_j,\mathbf{u}_i>}\right)^y =$$
$$\mathbf{a}_j \omega^{<\mathbf{c},\mathbf{u}_j>} m,$$

where the last equation follows from what we showed in Example 5.3. Now $\omega^{<\mathbf{c},\mathbf{u}_j>}m$ is a non-zero number that we can calculate (provided we take $m \neq 0$ in $\mathbb{F}_q$), so having computed $\sum_{y=0}^{m-1} F_{\mathbf{a}}(\mathbf{c}+y\mathbf{v}_j)$, we can divide it by $\omega^{<\mathbf{c},\mathbf{u}_j>}m$ to get an estimate for $\mathbf{a}_j$ (which will be correct if there were no errors in the m coordinates of $E(\mathbf{a})$ that we queried). Since each query is uniformly chosen over all coordinates of the codeword, and each of them has $\delta = \frac{1}{4m}$ probability of error, by the Union Bound the probability that there's at least one error is no more than $\frac{1}{4}$.

□

Question 5.1. How big can a Matching Vector family in $\mathbb{Z}_m^\ell$ be?

Theorem 5.2. If p is prime, then any Matching Vector family in $\mathbb{Z}_p^\ell$ has size at most $\ell^{p-1}+1$.

Proof. We will work with matrices over $\mathbb{F}_p$. We are going to use the following claim.

Claim 5.2. Suppose A is a $k \times k$ matrix over $\mathbb{F}_p$ with entries $A = (a_{i,j})_{i,j=1}^k$. Let $B = (b_{i,j} = a_{i,j}^t)_{i,j=1}^k$ be the matrix obtained from A by raising each entry to the t-th power. Then $rank(B) \leq rank(A)^t$.

Proof. For any X, Y such that X is a $m \times n$ matrix and Y is a $p \times q$ matrix, let $X \otimes Y$ be the Kronecker product of X and Y, a $mp \times nq$ matrix with entries $(X \otimes Y)_{p(r-1)+v,q(s-1)+w} = x_{r,s}y_{v,w}$. Let $T = A^{\otimes t} = A \otimes A \otimes \cdots \otimes A$, where in the last expression A occurs t times. Now T is a $k^t \times k^t$ matrix. It can be shown that for any two matrices X and Y, $rank(X \otimes Y) = rank(X)rank(Y)$ [Lau05]. Thus, $rank(T) = rank(A)^t$. Since B is a submatrix of T, $rank(B) \leq rank(T)$, so we conclude that $rank(B) \leq rank(A)^t$.

□

Consider a Matching Vector family $(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_k)$ in $\mathbb{Z}_p^\ell$. Let A be the $k \times k$ matrix with entries $a_{i,j} =< \mathbf{u}_i, \mathbf{v}_j >$. Then $rank(A) \leq \ell$ since $A = DG$, where $D_{i,j} = \mathbf{u}_{i,j}$ and

$G_{i,j} = \mathbf{v}_{j,i}$ (D is a $k \times \ell$ matrix and so $rank(D) \leq \ell$, therefore $rank(DG) \leq \ell$). Let $B = (b_{i,j} = a_{i,j}^p)$, then by Claim 5.2, $rank(B) \leq rank(A)^{p-1} \leq \ell^{p-1}$. Notice that A has zero diagonal entries and non-zero entries off the diagonal, and so $\forall i \in [k], b_{i,i} = 0$ and $\forall i, j \in [k], i \neq j, b_{i,j} = a_{i,j}^{p-1} = 1$ by Fermat's Little Theorem since $a_{i,j} \neq 0 \mod p$. Let C be the $k \times k$ matrix with $\forall i, j \in [k], c_{i,j} = 1$, and let I_k be the identity $k \times k$ matrix. We have that $C - B = I_k$. Therefore, $rank(C) + rank(B) \geq rank(I_k)$. This is because the columns of I_k are in the span of the set of vectors that consists of all columns of C and all columns of B. Next, notice that $rank(C) = 1, rank(I_k) = k$, so $rank(B) \geq k - 1$. Using the fact that $rank(B) \leq \ell^{p-1}$, which we established above, we get that $k - 1 \leq \ell^{p-1}$.

□

Corollary 5.1. A Matching Vector code in $\mathbb{Z}_p^\ell$ with p prime yields a Locally Decodable Code that is no better than the Low-Degree Extension code ($E : \mathbb{F}_q^k \to \mathbb{F}_q^{p^\ell}$, with $k = \ell^p$ and p queries).

Theorem 5.3 ([Gro00])**.** There exists a Matching Vector family over $\mathbb{Z}_6^\ell$ of size $\ell^{\frac{C \log \ell}{\log^2 \log \ell}}$, where $C = \frac{1}{81}$.

The proof of this theorem is in the next lecture.

Corollary 5.2. There is a 6-query Locally Decodable Code $E : \mathbb{F}_q^k \to \mathbb{F}_q^{6^\ell}$ with $k = \ell^{\frac{\log \ell}{\log \log \ell}}$. Notice that since $\forall \epsilon > 0, n = 6^\ell < 2^{k^\epsilon}$, this is a sub-exponential Locally Decodable Code.

References

[Efr09] Klim Efremenko. 3-query locally decodable codes of subexponential length. In *STOC '09 Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 39–44, 2009.

[Gro00] Vince Grolmusz. Superpolynomial size set-systems with restricted intersections mod 6 and explicit ramsey graphs. *Combinatorica*, 20:71–86, 2000.

[Lau05] Alan J. Laub. *Matrix Analysis for Scientists and Engineers*. SIAM: Society for Industrial and Applied Mathematics, 2005.

[Maz03] Marc Moreno Maza. Advanced computer algebra: From Newton to Hensel. `http://www.csd.uwo.ca/~moreno//AM583/Lectures/Newton2Hensel.html/node9.html`, 2003.

[Yek12] Sergey Yekhanin. Locally decodable codes. *Foundations and Trends® in Theoretical Computer Science*, 6(3):139–255, 2012.