Linear Locally Decodable Codes	**Fall 2016**

Lecture 9: LDC's and LCC's over $\mathbb{R}$ and $\mathbb{C}$

Lecturer: Zeev Dvir *Scribe: Kalina Petrova*

The "geometric" definitions of LDC's and LCC's, that is, the definitions that use matchings, extend to any field.

Definition 9.1. An r-LDC over $\mathbb{R}^k$ is a list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ such that $\forall i \in [k]$, there is a matching $M^i = \{T^i_1, \ldots, T^i_{m_i}\}$, where $\forall j \in [m_i], T^i_j \subseteq [n], |T^i_j| = r$ and $m_i \geq \frac{\delta n}{r}$, such that $\forall T^i_j \in M^i, \mathbf{e}_i \in \text{span}\{\mathbf{v}_g | g \in T^i_j\}$.

Definition 9.2. An r-LCC over $\mathbb{R}^k$ is a list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ such that $\forall i \in [n]$, there is a matching $M^i = \{T^i_1, \ldots, T^i_{m_i}\}$, where $\forall j \in [m_i], T^i_j \subseteq [n], |T^i_j| = r$ and $m_i \geq \frac{\delta n}{r}$, such that $\forall T^i_j \in M^i, \mathbf{v}_i \in \text{span}\{\mathbf{v}_g | g \in T^i_j\}$.

The main questions we will be interested in this lecture are what constructions and lower bounds generalize to infinite fields, and whether it is easier or harder to construct codes over $\mathbb{R}$ and $\mathbb{C}$ than over finite fields.

Here we discuss the known constructions. Firstly, notice that 2-query Hadamard Locally Decodable Code works over any field, because we can take $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, where $\forall i \in [n], \mathbf{v}_i \in \{0,1\}^k \subseteq \mathbb{R}^k$, and decoding works exactly the same way as before. Matching Vector codes work as LDC's over any field, finite or not, that has an element ω_m of order m, where the Matching Vector family used is over $\mathbb{Z}_m$. For instance, we can take $\mathbb{C}$ with $\omega_m = e^{\frac{2\pi i}{m}}$. Note that this can also work in $\mathbb{R}$ using twice as many queries by taking $\omega_m = e^{\frac{-2\pi}{2m}}$ which has order $2m$ since $e^{-2\pi} = 1^i = 1$.

When it comes to LCC's, Hadamard and Reed-Muller do not work, and actually no non-trivial LCC over $\mathbb{R}$ or $\mathbb{C}$ with $o(n)$ queries is known.

Conjecture 9.1. There are no constant query Locally Correctable Codes over $\mathbb{R}$ or $\mathbb{C}$.

This conjecture is known to be true for 2–query Locally Correctable Codes.

Theorem 9.1. [BDWY11] Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{C}^k$ are a 2-query Locally Correctable Code with error δ. Then $k = \dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) \leq poly(\frac{1}{\delta})$.

This implies that for any constant $\delta > 0$, there does not exist an infinite family of Locally Correctable Codes with error δ over the complex numbers. We will prove this theorem for the special case when there are no repetitions in the code, that is, no $\mathbf{v}_i$ is a multiple of some $\mathbf{v}_j$ with $i \neq j$.

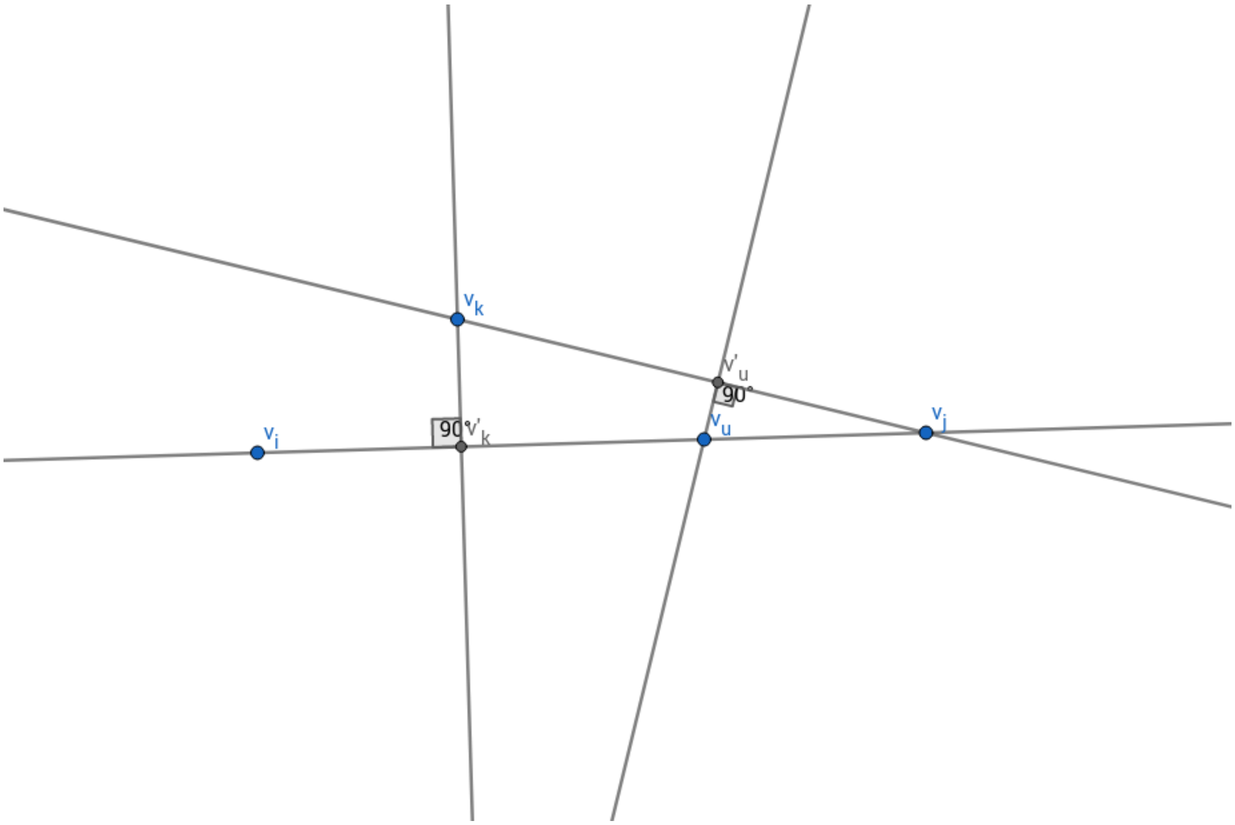

Figure 1: Proof of the Sylvester-Gallai Theorem

Proof. The proof goes through an equivalent way of looking at 2-query LCC's related to the Sylvester-Gallai theorem from discrete geometry.

Sylvester-Gallai configurations

Theorem 9.2. (Sylvester-Gallai) Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ be n distinct points. Suppose that $\forall i, j \in [n], i \neq j$, the line through $\mathbf{v}_i$ and $\mathbf{v}_j$ contains at least one other point from the set. Then all points must lie on a single line.

Proof. Suppose not all points are on a single line. For any $i, j \in [n], i \neq j$, denote by $L_{i,j}$ the line that passes through $\mathbf{v}_i$ and $\mathbf{v}_j$. Define $\alpha = \min_{i,j,k,i \neq j}\{dist(L_{i,j}, \mathbf{v}_k) | \mathbf{v}_k \notin L_{i,j}\}$. Since we assume not all points are on the same line, we know that α is well defined and $\alpha > 0$. Let i, j, k be the indices that minimize α as above. Let $\mathbf{v}_u \in L_{i,j}$ and $u \neq i, u \neq j$. We know such a point exists by assumption. At least two of the points $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_u$ lie on the same side of the perpendicular projection of $\mathbf{v}_k$ on $L_{i,j}$ (denote this projection by $\mathbf{v}'_k$). Without loss of generality, suppose that $\mathbf{v}_u$ and $\mathbf{v}_j$ lie on the same side of $\mathbf{v}'_k$ and $\mathbf{v}_u$ is closer to $\mathbf{v}'_k$ than $\mathbf{v}_j$ is (it is possible that $\mathbf{v}_u$ coincides with $\mathbf{v}'_k$). But then $\mathbf{v}_u$ is closer to $L_{j,k}$ than $\mathbf{v}_k$ is to $L_{i,j}$, which contradicts the minimality of α. This is because if $\mathbf{v}'_u$ is the perpendicular projection of $\mathbf{v}_u$ on $L_{j,k}$, then $\mathbf{v}_j\mathbf{v}_u\mathbf{v}'_u$ and $\mathbf{v}_j\mathbf{v}_k\mathbf{v}'_k$ are similar triangles, contained in one another.

$\square$

Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ support some family $\mathcal{F}$ of linearly dependent triples. We can find $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^k$ whose dimension is at least $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) - 1$ so that all triples in $\mathcal{F}$ are collinear. We do this in the following way. Let $H \subseteq \mathbb{R}^k$ be a hyperplane that does not pass through the origin but passes through the lines $O\mathbf{v}_1, \ldots, O\mathbf{v}_n$, where O is the origin. Then for each $i \in [n]$, take $\mathbf{u}_i = c\mathbf{v}_i$ for some $c \in \mathbb{R}$ such that $\mathbf{u}_i \in H$. Then any previously linearly dependent triple now becomes collinear, because the three points in question previously defined a plane that the origin lied on, and now this plane is projected on H as a line. This means that if $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ are a 2-query LCC with error δ, then there are vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^k$ with the same matchings $M^1, \ldots, M^n$ as in the 2-LCC such that in matching M^i all pairs $\{\mathbf{u}_j, \mathbf{u}_k\}$ are collinear with $\mathbf{u}_i$. Thus, from now on we can work with $\mathbf{u}_1, \ldots, \mathbf{u}_n$ instead of $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

This leads to the following generalization of the Sylvester-Gallai Theorem.

Definition 9.3. A δ-Sylvester-Gallai configuration is a set of n distinct points $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ such that $\forall i \in [n]$, there exists a set $J_i \subseteq [n] \setminus \{i\}$ with $|J_i| \geq \delta(n-1)$ such that $\forall j \in J_i$, $L_{i,j}$ contains a third point from the set.

Theorem 9.3. [DSW14]

If $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^k$ is a δ-Sylvester-Gallai configuration, then we have that $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) = O(\frac{1}{\delta})$.

Theorem 9.3 also holds over the complex number and in its form over $\mathbb{C}^k$ it immediately implies Theorem 9.4 without repetitions. This is because if $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{C}^k$ define a 2-query LCC with error δ, then there are vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{C}^k$ with $\dim(\text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}) \geq \dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) - 1$ and such that $\forall i \in [n], \exists M^i$ with $|M^i| \geq \frac{\delta n}{2}$ and $\forall T^i_j = \{\mathbf{u}_k, \mathbf{u}_l\} \in M^i$, $\mathbf{u}_i, \mathbf{u}_k$, and $\mathbf{u}_l$ are collinear. Then $\mathbf{u}_1, \ldots, \mathbf{u}_n$ form a δ-Sylvester-Gallai configuration, so by Theorem 9.3 $\dim(\text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}) = O(\frac{1}{\delta})$, which implies Theorem 9.4.

$\square$

We will devote the rest of these notes to proving Theorem 9.3 over the real numbers.

Proof of Theorem 9.3. Given $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathbb{R}^k$ as in Theorem 9.3, we denote by A the $n \times k$ matrix with rows $\mathbf{v}_1, \ldots, \mathbf{v}_n$. If $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k$ are collinear, there are non-zero coefficients $a, b, c \in \mathbb{R}$ such that $a\mathbf{v}_i + b\mathbf{v}_j + c\mathbf{v}_k = 0$. Consider the row vector ω such that $\omega_i = a, \omega_j = b, \omega_k = c$ and $\forall g \in [n]$ with $g \neq i, g \neq j, g \neq k, \omega_g = 0$. We have that $\omega A = 0$ (that is, ωA is a row vector of length k), because $a\mathbf{v}_i + b\mathbf{v}_j + c\mathbf{v}_k = 0$. For any collinear triple $T = \{i, j, k\}$, let ω^T be the row vector of length n that corresponds to $a\mathbf{v}_i + b\mathbf{v}_j + c\mathbf{v}_k = 0$ as described above. Given a family $\mathcal{T}$ of collinear triples with

$|\mathcal{T}| \le \binom{\lfloor n \rfloor}{3}$, we define the *dependency matrix* of V with respect to $\mathcal{T}$ as the matrix $M^{\mathcal{T}}$ with $|\mathcal{T}|$ rows such that row T of $M^{\mathcal{T}}$, where $T \in \mathcal{T}$, is ω^T. Since as discussed above, $\forall T \in \mathcal{T}, \omega^T A = 0$, then $M^{\mathcal{T}} A = 0$, that is, $M^{\mathcal{T}} A$ is a $|\mathcal{T}| \times k$ matrix in which all entries are 0. This means that columns of A are in the kernel of $M^{\mathcal{T}}$. Hence, if $\text{rank}(M^{\mathcal{T}}) \ge n-d$, then by the Rank-nullity theorem, $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) = \text{rank}(A) \le d$.

The idea of the proof is as follows. We will pick a family $\mathcal{T}$ of collinear triples such that M has high rank. We will require the following properties from $\mathcal{T}$:

1. Each $\mathbf{v}_i$ is in many ($\ge \delta n$) triples from $\mathcal{T}$.
2. Every pair $\mathbf{v}_i, \mathbf{v}_j$ is together in at most $O(1)$ triples of $\mathcal{T}$.

These properties imply that A is a *design matrix.*

Definition 9.4. An $m \times n$ matrix M is a (q, s, t)-*design matrix* if

1. Each row of A has at most q non-zero entries.
2. Each column of A has at least s non-zero entries.
3. The supports of any two columns intersect in at most t positions. That is, $\forall j, j' \in [n], j \ne j'$, there are at most t values of $i \in [m]$ such that $A_{i,j} A_{i,j'} \ne 0$.

Notice that if $\mathcal{T}$ satisfies the properties above, then $M^{\mathcal{T}}$ is a $(3, \delta n, O(1))$-design matrix. This is because by construction of $M^{\mathcal{T}}$, each row has exactly 3 non-zero entries; each $\mathbf{v}_i$ being in at least δn triples implies that each column of $M^{\mathcal{T}}$ has at least δn non-zero entries; every pair $\mathbf{v}_i, \mathbf{v}_j$ being together in at most $O(1)$ triples of $\mathcal{T}$ implies that for each $i, j \in [n], i \ne j$, the number of indices $k \in |\mathcal{T}|$ such that $A_{k,i} \ne 0$ and $A_{k,j} \ne 0$ is $O(1)$.

To finish the proof of Theorem 9.3, we use the following Theorem.

Theorem 9.4. [BDWY11] If M is an $m \times n$ (q, s, t)-design matrix over $\mathbb{R}$ or over $\mathbb{C}$, then $\text{rank}(M) \ge n - \left(\frac{qtn}{2s}\right)^2$.

This result was improved in [DSW14] to $\text{rank}(M) \ge n - \frac{q(q-1)tn}{s}$.

Since we have that $M^{\mathcal{T}}$ is a $(3, \delta n, O(1))$-design matrix, we get from Theorem 9.4 that $\text{rank}(M^{\mathcal{T}}) \ge n - \left(\frac{3O(1)n}{2\delta n}\right)^2 = n - O(\frac{1}{\delta^2})$. This implies that $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) \le O(\frac{1}{\delta^2})$ by the Rank-nullity theorem. Using the improved result from [DSW14], we get the tighter bound $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}) \le O(\frac{1}{\delta})$ claimed in Theorem 9.3.

We still have to prove that a family $\mathcal{T}$ of collinear triples with the properties listed above exists, and we will also prove Theorem 9.4.

Proof of Theorem 9.4. We first consider a special case, which is easier to handle. Suppose all non-zero entries of M are bounded as follows: $\forall i,j$ such that $M_{i,j} \neq 0$, we have $\frac{1}{100} \leq |M_{i,j}| \leq 1$. Geometrically, this corresponds to not having unbalanced collinear triples $\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z$ in which, say, $\mathbf{v}_x$ and $\mathbf{v}_y$ are much closer to each other than each of them is to $\mathbf{v}_z$. This is because the non-zero entries correspond to the coefficients in the linear combinations between the collinear triples. Consider the $n \times n$ matrix $B = M^T M$ in which $B_{i,j} = \langle C_i, C_j \rangle$, where C_j is the jth column of M. Notice that:

- $\forall j \in [n], B_{j,j} = \langle C_j, C_j \rangle \geq \Omega(s)$ because C_j has at least s non-zero entries and each one has absolute value at least $\frac{1}{100}$.
- $\forall i \neq j, |B_{i,j}| \leq t$ because C_i and C_j intersect in at most t positions and each entry is at most 1 in absolute value.

This means that B has large diagonal entries and small off-diagonal entries. We can think of B as a scaled permutation of the identity matrix. Such matrices have high rank, as can be seen from the following lemma.

Lemma 9.1. Let B be an $n \times n$ symmetric matrix such that $\forall j \in [n], B_{j,j} \geq L$ and $\forall i,j \in [n], i \neq j, B_{i,j} \leq c$. Then $\text{rank}(B) \geq n - \frac{n^2 c}{L^2}$.

Proof. Without loss of generality, we can assume that $\forall j \in [n], B_{j,j} = L$ (otherwise we can scale down the columns in which this is not true). B has n eigenvalues (with possible multiplicities) $\lambda_1, \ldots, \lambda_n$. Let $r = \text{rank}(B)$ and suppose $\lambda_1, \ldots, \lambda_r$ are non-zero and $\lambda_{r+1}, \ldots, \lambda_n$ are all zeros. So $B_{1,1} + \cdots + B_{n,n} = tr(B) = \lambda_1 + \cdots + \lambda_r$. Now

$$(nL)^2 \leq tr(R)^2 = (\sum_{i=1}^{r} \lambda_i)^2 \leq r \sum_{i=1}^{r} \lambda_i^2$$

$$= r \cdot tr(B^2) = r \cdot tr(B^T B) = r \sum_{i,j} B_{i,j}^2 \leq r(nL^2 + n^2 c),$$

where we used the Cauchy-Schwarz inequality to show $(\sum_{i=1}^{r} \lambda_i)^2 \leq r \sum_{i=1}^{r}$. Thus we have that $r \geq \frac{n^2 L^2}{nL^2 + n^2 c} = \frac{n}{1 + \frac{nc}{L^2}} \geq n(1 - \frac{nc}{L^2}) = n - \frac{n^2 c}{L^2}$. $\square$

Applying Lemma 9.1 to $B = M^T M$, for which we know $B_{j,j} \geq \Omega(s)$ and $B_{i,j} \leq t$, we get that $\text{rank}(B) = \text{rank}(M) \geq n - O(\frac{n^2 t}{s^2})$.

We now have to reduce the general case, in which we know nothing about the matrix M, to the special case in which the entries are bounded as above. We use the technique of *matrix-scaling*, as follows. We multiply each row/column of M by the same non-zero real number. When we do this, $\text{rank}(M)$ does not change and M is still a (q, s, t)-design matrix. In this way we can try to balance the entries to get them to be as in the special case we considered.

Definition 9.5. An $m \times n$ matrix M' is a *scaling* of M if $\exists \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in \mathbb{R}$ with $\forall i \in [m], \alpha_i \neq 0, \forall i \in [n], \beta_i \neq 0$ such that $\forall i \in [m], j \in [n], M'_{i,j} = \alpha_i \beta_j M_{i,j}$.

Definition 9.6. An $m \times n$ matrix M is called l_2-*doubly stochastic*, which we will abbreviate as DS, if $\forall i \in [m], \sum_{j=1}^{n} |M_{i,j}|^2 = 1$ and $\forall j \in [n], \sum_{i=1}^{m} |M_{i,j}|^2 = \frac{m}{n}$.

Definition 9.7. An $m \times n$ matrix M is *scalable* if $\forall \varepsilon > 0$, there exists a scaling M' of M that is ε-close to being l_2-doubly stochastic. That is, there exists an l_2-doubly stochastic matrix R such that $\|M' - R\|_2 \leq \varepsilon$.

Example 9.1. Consider the following matrix.

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

M is scalable because $\forall \varepsilon > 0$, we can take $\alpha_1 = \varepsilon, \alpha_2 = 1, \beta_1 = \frac{1}{\varepsilon}, \beta_2 = 1$, which gives us the scaling

$$M' = \begin{bmatrix} 1 & \varepsilon \\ 0 & 1 \end{bmatrix}$$

which is ε-close to the l_2-doubly stochastic matrix

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We use the following theorem due to Sinkhorn.

Theorem 9.5. *M is scalable if and only if $\forall a \times b$ sub-matrix N of M with $\forall i \in [a], j \in [b], N_{i,j} = 0$, we have $\frac{a}{m} + \frac{b}{n} \leq 1$.*

Example 9.2. If $m = n$, then we get that M is scalable if and only if for any zero $a \times b$ submatrix of M, we have $a + b \leq n$. This condition is equivalent to the one in Hall's Theorem, which says that there exists a perfect matching in a bipartite graph with bipartite sets X and Y if and only if $\forall W \subseteq X, |W| \leq |N_G(W)|$, where $N_G(W)$ is the set of all neighbour of vertices in W. In our case, the matrix would be the adjacency matrix of the graph, and a zero $a \times b$ submatrix of it would mean that there is $W \subseteq X$ with $|W| > |N_G(W)|$ as the vertices that correspond to the a rows which participate in the submatrix can only have as neighbours the columns outside that submatrix, and there are $n - b < a$ of those.

Example 9.3. We will use the special case in which the $m \times n$ matrix R has n "blocks" $R_1, \ldots, R_n$, each of size $s \times n$, where $m = ns$, and the j-th column in R_j has all non-zero entries. Notice that the condition from above is satisfied, so R is scalable. The reason why the condition is satisfied is the following. Suppose we have a zero $a \times b$ submatrix. If it uses rows from i different blocks of R, then it has at most $n - i$ columns, because for each used block, it cannot use the column with the same number. Thus we get $b \leq n - i$ and $a \leq ni$, so $\frac{a}{ns} + \frac{b}{n} \leq \frac{is}{ns} + \frac{n-i}{n} = 1$.

We now go back to reducing the general case of an $m \times n$ (q, s, t)-design matrix to an $m \times n$ (q, s, t)-design matrix such that $\forall i, j$ such that $M_{i,j} \neq 0$, we have $\frac{1}{100} \leq |M_{i,j}| \leq 1$. Let M be an $m \times n$ (q, s, t)-design matrix. Construct a scalable matrix R with ns rows and n columns such that R_1 corresponds to the submatrix consisting of the first s rows of R, R_2 to rows $s+1$ through $2s$, and R_i to rows $(i-1)s+1$ through is. To construct R, we use rows from M greedily. That is, to build R_j, we use the s rows of M that have non-zeros in position j.

Claim 9.1. *R is a (q, s, tq)-design matrix and* $\text{rank}(R) \leq \text{rank}(M)$.

Proof of Claim 9.1. The rows of R are also rows of M, so it holds that each row of R has at most q non-zero entries. Furthermore, each column i of R has s non-zeros – the ones in R_i. Finally, consider some columns j and j' with $j \neq j'$. There are at most t values $i \in [m]$ such that $M_{i,j}M_{i,j'} \neq 0$, and for each such row i, there are at most q non-zero entries the i-th row of M, which means that this row can participate at most q times in R, so there are at most qt values $i' \in [ns]$ such that $R_{i',j}R_{i',j'} \neq 0$. To see that $\text{rank}(R) \leq \text{rank}(M)$, notice that $\text{nullity}(R) \geq \text{nullity}(M)$ because any vector in the kernel of M is also in the kernel of R. $\square$

Let R' be a scaling of R that is l_2-doubly stochastic, where we ignore ε as we can let it go to 0. Such a scaling exists because as shown above, R is scalable. Then $B = R'^T R'$ is an $n \times n$ matrix with the following properties. First of all, $\forall i \in [n], B_{i,i} \geq \frac{m}{n} = \frac{ns}{n} = s$ because $B_{i,i} = \sum_{j=1}^m R'^2_{i,j} = \frac{m}{n}$ since R' is l_2-doubly stochastic. Furthermore, $\forall i, j \in [n]$ with $i \neq j$, $B_{i,j} \leq qt$ because $B_{i,j} = \sum_{k=1}^m R'_{k,i}R'_{k,j}$, there are at most tq values of k such that $R'_{k,i}R'_{k,j} \neq 0$ and as R' is l_2-doubly stochastic, no entry can have absolute value bigger than 1. Then by Lemma 9.1, we have that $\text{rank}(B) \geq n - \frac{n^2qt}{s^2}$. Thus, $\text{rank}(M) \geq \text{rank}(R') = \text{rank}(B) \geq n - \frac{n^2qt}{s^2}$.

$\square$

The only thing left to prove is that there exists a family $\mathcal{T}$ of collinear triples with the following properties:

1. Each $\mathbf{v}_i$ is in many $(\geq \delta n)$ triples from $\mathcal{T}$.
2. Every pair $\mathbf{v}_i, \mathbf{v}_j$ is together in at most $O(1)$ triples of $\mathcal{T}$.

To construct $\mathcal{T}$, we use the following claim.

Claim 9.2. *$\forall l \geq 3$, there exists a family of triples (multiset) $T_l \subseteq \binom{[l]}{3}$ such that*

1. *$\forall i$, there are at least $3(l-1)$ triples in T_l containing i.*
2. *$\forall i \neq j$, there are at most a constant number of triples in T_l containing both i and j.*

To build $\mathcal{T}$, apply Claim 9.2 on each line l with at least three points and take the union of all such families T_l. For each $\mathbf{v}_i$, suppose that it lies on k lines with at least three points on each of them and suppose the number of points on each of these lines is respectively $l_1, \ldots, l_k$. Since we are considering a δ-Sylvester-Gallai configuration, we have that $\sum_{j=1}^{k}(l_j - 1) \geq \delta(n-1)$. On the other hand, for each line $j \in [k]$ with at least three points on it that $\mathbf{v}_i$ lies on, we have that $\mathbf{v}_i$ participates in at least $3(l_j - 1)$ triples coming from this line. Thus we have that $\mathbf{v}_i$ participates in at least $\sum_{j=1}^{k} 3(l_j - 1) \geq 3\delta(n-1)$ triples. On the other hand, any pair $\mathbf{v}_i, \mathbf{v}_j$ can only participate together in triples coming from the line that they form, and they participate in at most $O(1)$ of those.

□

Exercise 9.1. Suppose there is a r-query LCC $E : \mathbb{Q}^k \to \mathbb{Q}^n$. Show that there exists an r-query LCC $E : \mathbb{F}_p^q \to \mathbb{F}_p^n$ for some prime p.

Exercise 9.2. Let M be an $m \times n$ (r, s, t)-design matrix over some field $\mathbb{F}$ with $s = \delta n$ and $t = O(1)$. Show that $V = \ker(M) \subseteq \mathbb{F}^n$ is an r-query LCC with error $\Omega(\delta)$.

Exercise 9.3. Show that one cannot improve the $\frac{1}{\delta}$ dimension bound of δ-SG configurations. Show that there exists a δ-SG configuration such that no line contains more than $o(n)$ points.

References

[BDWY11] Boaz Barak, Zeev Dvir, Avi Wigderson, and Amir Yehudayoff. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. *ACM Symposium on Theory of Computing*, 2011.

[DSW14] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Improved rank bounds for design matrices and a new proof of kelly's theorem. *Forum of Mathematics, Sigma*, 2014.