Sylvester-Gallai for Arrangements of Subspaces

Zeev Dvir[1] and Guangda Hu[2]

1 **Department of Computer Science and Department of Mathematics, Princeton University**
35 Olden Street, Princeton, NJ 08540-5233, USA
`zeev.dvir@gmail.com`

2 **Department of Computer Science, Princeton University**
35 Olden Street, Princeton, NJ 08540-5233, USA
`guangdah@cs.princeton.edu`

Abstract

In this work we study arrangements of k-dimensional subspaces $V_1, \ldots, V_n \subset \mathbb{C}^\ell$. Our main result shows that, if every pair V_a, V_b of subspaces is contained in a dependent triple (a triple V_a, V_b, V_c contained in a $2k$-dimensional space), then the entire arrangement must be contained in a subspace whose dimension depends only on k (and not on n). The theorem holds under the assumption that $V_a \cap V_b = \{0\}$ for every pair (otherwise it is false). This generalizes the Sylvester-Gallai theorem (or Kelly's theorem for complex numbers), which proves the $k = 1$ case. Our proof also handles arrangements in which we have many pairs (instead of all) appearing in dependent triples, generalizing the quantitative results of Barak et. al. [1].

One of the main ingredients in the proof is a strengthening of a theorem of Barthe [3] (from the $k = 1$ to $k > 1$ case) proving the existence of a linear map that makes the angles between pairs of subspaces large on average. Such a mapping can be found, unless there is an obstruction in the form of a low dimensional subspace intersecting many of the spaces in the arrangement (in which case one can use a different argument to prove the main theorem).

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Sylvester-Gallai, Locally Correctable Codes

Digital Object Identifier 10.4230/LIPIcs.SOCG.2015.29

1 Introduction

The Sylvester-Gallai (SG) theorem states that for n points $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^\ell$, if for every pair $\boldsymbol{v}_i, \boldsymbol{v}_j$ there is a third point $\boldsymbol{v}_k$ on the line passing through $\boldsymbol{v}_i, \boldsymbol{v}_j$, then all points must lie on a single line. This was first posed by Sylvester [14], and was solved by Melchior [13]. It was also conjectured independently by Erdös [9] and proved shortly after by Gallai. We refer the reader to the survey [4] for more information about the history and various generalizations of this theorem. The complex version of this theorem was proved by Kelly [11] (see also [8, 7] for alternative proofs) and states that if $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{C}^\ell$ and for every pair $\boldsymbol{v}_i, \boldsymbol{v}_j$ there is a third $\boldsymbol{v}_k$ on the same complex line, then all points are contained in some complex plane (over the complex numbers, there are planar examples and so this theorem is tight).

In [7] (based on earlier work in [1]), the following quantitative variant of the SG theorem was proved. For a set $S \subset \mathbb{C}^\ell$ we denote by $\dim(S)$ the smallest d such that S is contained in a d-dimensional subspace of $\mathbb{C}^\ell$.

▶ **Theorem 1.1** ([7]). *Given n points $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{C}^\ell$, if for every $i \in [n]$ there exists at least δn values of $j \in [n] \setminus \{i\}$ such that the line through $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ contains a third point $\boldsymbol{v}_k$, then* $\dim\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \leq 10/\delta$.

31st International Symposium on Computational Geometry (SoCG'15).
Editors: Lars Arge and János Pach; pp. 29–43
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

(The dependence on δ is asymptotically tight). From here on, we will work with homogeneous subspaces (passing through zero) instead of affine subspaces (lines/planes etc). The difference is not crucial to our results and the affine version can always be derived by intersecting with a generic hyperplane. In this setting, the above theorem will be stated for a set of one-dimensional subspaces, each spanned by some $\boldsymbol{v}_i$ (and no two $\boldsymbol{v}_i$'s being a multiple of each other) and collinearity of $\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k$ is replaced with the three vectors being linearly dependent (i.e., contained in a 2-dimensional subspace).

One natural high dimensional variant of the SG theorem, studied in [10, 1], replaces 3-wise dependencies with t-wise dependencies (e.g, every triple is in some coplanar four-tuple). In this work, we raise another natural high-dimensional variant in which the *points* themselves are replaced with k-dimensional subspaces. We consider such arrangements with many 3-wise dependencies (defined appropriately) and attempt to prove that the entire arrangement lies in some low dimensional space. We will consider arrangements $V_1, \ldots, V_n \subset \mathbb{C}^\ell$ in which each V_i is k-dimensional and with each pair satisfying $V_{i_1} \cap V_{i_2} = \{\mathbf{0}\}$. A dependency can then be defined as a triple $V_{i_1}, V_{i_2}, V_{i_3}$ of k-dimensional subspaces that are contained in a single $2k$-dimensional subspace. The pair-wise zero intersections guarantee that every pair of subspaces defines a unique $2k$-dimensional space (their span) and so, this definition of dependency behaves in a similar way to collinearity. For example, we have that if $V_{i_1}, V_{i_2}, V_{i_3}$ are dependent and $V_{i_2}, V_{i_3}, V_{i_4}$ are dependent then also $V_{i_1}, V_{i_2}, V_{i_4}$ are dependent. This would not hold if we allowed some pairs to have non zero intersections. In fact, if we allow non-zero intersection then we can construct an arrangement of two dimensional spaces with many dependent triples and with dimension as large as $\sqrt{n}$ (see below). We now state our main theorem, generalizing Theorem 1.1 (with slightly worse parameters) to the case $k > 1$. We use the standard $V + U$ notation to denote the subspace spanned by all vectors in $V \cup U$. We use big 'O' notation to hide absolute constants.

▶ **Theorem 1.2.** *Let $V_1, V_2, \ldots, V_n \subset \mathbb{C}^\ell$ be k-dimensional subspaces such that $V_i \cap V_{i'} = \{\mathbf{0}\}$ for all $i \neq i' \in [n]$. Suppose that, for every $i_1 \in [n]$ there exists at least δn values of $i_2 \in [n] \setminus \{i_1\}$ such that $V_{i_1} + V_{i_2}$ contains some V_{i_3} with $i_3 \notin \{i_1, i_2\}$. Then*

$$\dim(V_1 + V_2 + \cdots + V_n) = O(k^4/\delta^2).$$

The condition $V_i \cap V_{i'} = \{\mathbf{0}\}$ is needed due to the following example. Set $k = 2$ and $n = \ell(\ell-1)/2$ and let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_\ell\}$ be the standard basis of $\mathbb{R}^\ell$. Define the n spaces to be $V_{ij} = \operatorname{span}\{\boldsymbol{e}_i, \boldsymbol{e}_j\}$ with $1 \leq i < j \leq \ell$. Now, for each $(i,j) \neq (i',j')$ the sum $V_{ij} + V_{i'j'}$ will contain a third space (since the size of $\{i, j, i', j'\}$ is at least three). However, this arrangement has dimension $\ell > \sqrt{n}$.

The bound $O(k^4/\delta^2)$ is probably not tight and we conjecture that it could be improved to $O(k/\delta)$, possibly with a modification of our proof. One can always construct an arrangement with dimension $2k/\delta$ by partitioning the subspaces into $1/\delta$ groups, each contained in a single $2k$ dimensional space.

Overview of the proof: A preliminary observation is that it suffices to prove the theorem over $\mathbb{R}$. This is because an arrangement of k-dimensional complex subspaces can be translated into an arrangement of $2k$-dimensional real subspaces (this is proved at the end of Section 2). Hence, we will now focus on real arrangements.

The proof of the theorem is considerably simpler when the arrangement of subspaces $V_1, \ldots, V_n$ satisfies an extra 'robustness' condition, namely that every two spaces have an angle bounded away from zero. More formally, if for every two unit vectors $\boldsymbol{v}_1 \in V_{i_1}$ and $\boldsymbol{v}_2 \in V_{i_2}$ we have $|\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle| \leq 1 - \tau$ for some absolute constant $\tau > 0$. This condition implies that, when we have a dependency of the form $V_{i_3} \subset V_{i_1} + V_{i_2}$, every unit vector in V_{i_3} can be obtained as a linear combination *with bounded coefficients* (in absolute value) of unit vectors from V_{i_1}, V_{i_2}. Fixing an orthogonal basis for each subspace and using the conditions of the theorem, we are able to construct many local linear dependencies between the basis elements. We then show (using the bound on the coefficients in the linear combinations) that the space of linear dependencies between all basis vectors, considered as a subspace of $\mathbb{R}^{kn}$, contains the rows of an $nk \times nk$ matrix that has large entries on the diagonal and small entries off the diagonal. Since matrices of this form have high rank (by a simple spectral argument), we conclude that the original set of basis vectors must have small dimension.

To handle the general case, we show that, unless some low dimensional subspace W intersects many of the spaces V_i in the arrangement, we can find a change of basis that makes the angles between the spaces large on average (in which case, the previous argument works). This gives us the overall strategy of the proof: If such a W exists, we project W to zero and continue by induction. The loss in the overall dimension is bounded by the dimension of W, which can be chosen to be small enough. Otherwise (if such W does not exist) we apply the change of basis and use it to bound the dimension.

The change of basis is found by generalizing a theorem of Barthe [3] (see [6] for a more accessible treatment) from the $k = 1$ case (arrangement of points) to higher dimension. We state this result here since we believe it could be of independent interest. To state the theorem we must first introduce the following, somewhat technical, definition.

▶ **Definition 1.3** (admissible basis set, admissible basis vector)**.** Given a list of vector spaces $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ $(V_i \subseteq \mathbb{R}^\ell)$, a set $H \subseteq [n]$ is called a $\mathcal{V}$*-admissible basis set* if

$$\dim(\sum_{i\in H} V_i) = \sum_{i\in H} \dim(V_i) = \dim(\sum_{i\in[n]} V_i),$$

i.e. if every space with index in H has intersection $\{\mathbf{0}\}$ with the span of the other spaces with indices in H, and the spaces with indices in H span the entire space $\sum_{i\in[n]} V_i$.

A $\mathcal{V}$*-admissible basis vector* is any indicator vector $\mathbf{1}_H$ of some $\mathcal{V}$-admissible basis set H (where the i-th entry of $\mathbf{1}_H$ equals 1 if $i \in H$ and 0 otherwise).

The following theorem is proved in Section 3.

▶ **Theorem 1.4.** *Given a list of vector spaces $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ $(V_i \subseteq \mathbb{R}^\ell)$ with $V_1 + V_2 + \cdots + V_n = \mathbb{R}^\ell$ and a vector $\boldsymbol{p} \in \mathbb{R}^n$ in the convex hull of all $\mathcal{V}$-admissible basis vectors. Then there exists an invertible linear map $M : \mathbb{R}^\ell \mapsto \mathbb{R}^\ell$ such that*

$$\sum_{i=1}^{n} p_i \operatorname{Proj}_{M(V_i)} = I_{\ell\times\ell},$$

where $M(V_i)$ is the linear space obtained by applying M on V_i, and $\operatorname{Proj}_{M(V_i)}$ is the orthogonal projection matrix onto $M(V_i)$.

The connection to the explanation given in the proof overview is as follows: If there is no subspace W of low dimension that intersects many of the spaces $V_1, \ldots, V_n$ then, one can show that there exists a vector $\boldsymbol{p}$ in the convex hull of all $\mathcal{V}$-admissible basis vectors such that the entries of $\boldsymbol{p}$ are not too small. This is enough to show that the average angle between pairs of spaces is large since otherwise one can derive a contradiction to the inequality which says that the sum of orthogonal projections of any unit vector must be relatively small.

The proof of the one dimensional case in [3] proceeds by defining a strictly convex function $f(t_1, \ldots, t_m)$ on $\mathbb{R}^m$ and shows that the function is bounded. This means that there must exist a point in which all partial derivatives of f vanish. Solving the resulting equations gives an invertible matrix that defines the required change of basis. We follow a similar strategy, defining an appropriate bounded function $f(t_1, \ldots, t_m, R_1, \ldots, R_n)$ in more variables, where the extra variables $R_1, \ldots, R_n$ represent the action of the orthogonal group $\mathbf{O}(k)$ on each of the spaces. However, in our case, we cannot show that f is strictly convex and so a maximum might not exist. However, we are still able to show that there exists a point in which all partial derivatives are very small (smaller than any $\epsilon > 0$), which is sufficient for our purposes.

Connection to Locally Correctable Codes. A q-query Locally Correctable Code (LCC) over a field $\mathbb{F}$ is a d-dimensional subspace $C \subset \mathbb{F}^n$ that allows for 'local correction' of codewords (elements of C) in the following sense. Let $\boldsymbol{y} \in C$ and suppose we have query access to $\boldsymbol{y}'$ such that $\boldsymbol{y}_i = \boldsymbol{y}'_i$ for at least $(1-\delta)n$ indices $i \in [n]$ (think of $\boldsymbol{y}'$ as a noisy version of $\boldsymbol{y}$). Then, for every i, we can probabilistically pick q positions in $\boldsymbol{y}'$ and, from their (possibly incorrect values), recover the correct value of $\boldsymbol{y}_i$ with high probability (over the choice of queries). LCC's play an important role in theoretical computer science (mostly over finite fields but recently also over the reals, see [5]) and are still poorly understood. In particular, when q is constant greater than 2, there are exponential gaps between the dimension of explicit constructions and the proven upper bounds. In [2] it was observed that q-LCCs are essentially equivalent to configurations of points with many local dependencies[1]. A variant of Theorem 1.1 shows for example that the maximal dimension of a 2-LCC in $\mathbb{R}^n$ has dimension bounded by $(1/\delta)^{O(1)}$. Our results can be interpreted in this framework as dimension upper bounds for 2-query LCC's in which each coordinate is replaced by a 'block' of k coordinates. Our results then show that, even under this relaxation, the dimension still cannot increase with n. The case of 3-query LCC's over the reals is still wide open (some modest progress was made recently in [6]) and we hope that the methods developed in this work could lead to further progress on this tough problem.

Organization. In Section 2, we define the notion of (α, δ)-systems (which generalizes the SG condition) and reduce our k-dimensional Sylvester-Gallai theorem to a more general theorem, Theorem 2.6, on the dimension of (α, δ)-systems (this part also includes the reduction from complex to real arrangements). Then, in Section 3, we prove the generalization of Barthe's theorem (Theorem 1.4). Finally, in Section 4, we prove our main result regarding (α, δ)-systems. Due to the page limit, some of the proof are available in the full version of this paper.

Acknowledgements. We would like to thank Patrick Devlin for helpful discussions on strengthening Theorem 1.4.

2 Reduction to (α, δ)-systems

The notion of an (α, δ)-system is used to 'organize' the dependent triples in the arrangement in a more convenient form so that each space is in many triples and every pair of spaces is together only in a few dependent triples. We also allow dependent *pairs* as those might arise when we apply a linear map on the arrangement.

[1] One important difference is that LCC's give rise to configurations where each point can repeat more than once.

▶ **Definition 2.1** ((α,δ)-system)**.** Given a list of vector spaces $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$), we call a list of sets $\mathcal{S} = (S_1, S_2, \ldots, S_w)$ an (α,δ)*-system* of $\mathcal{V}$ ($\alpha \in \mathbb{Z}^+$, $\delta > 0$) if

1. Every S_j is a subset of $[n]$ of size either 3 or 2.
2. If S_j contains 3 elements i_1, i_2 and i_3, then $V_{i_1} \subseteq V_{i_2} + V_{i_3}$, $V_{i_2} \subseteq V_{i_1} + V_{i_3}$ and $V_{i_3} \subseteq V_{i_1} + V_{i_2}$. If S_j contains 2 elements i_1 and i_2, then $V_{i_1} = V_{i_2}$.
3. Every $i \in [n]$ is contained in at least δn sets of $\mathcal{S}$.
4. Every pair $\{i_1, i_2\}$ ($i_1 \neq i_2 \in [n]$) appears together in at most α sets of $\mathcal{S}$.

Note that we allow $\delta > 1$ in an (α,δ)-systems. This is different from the statement of the Sylvester-Gallai theorem where $\delta \in [0,1]$. We have the following 3 simple observations, which are proved in the full version of this paper.

▶ **Lemma 2.2.** *Let $\mathcal{S} = (S_1, S_2, \ldots, S_w)$ be an (α,δ)-system of some vector space list $\mathcal{V}$. Then $\delta n^2/3 \leq w \leq \alpha n^2/2$ and $\delta/\alpha \leq 3/2$.*

▶ **Lemma 2.3.** *Let $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$) be a list of vector spaces and $\mathcal{S} = (S_1, S_2, \ldots, S_w)$ be a list of sets. If $w \geq \delta n^2$ and $\mathcal{S}$ satisfies the first, second and fourth requirements in Definition 2.1, then there exists a sublist $\mathcal{V}'$ of $\mathcal{V}$ and a sublist $\mathcal{S}'$ of $\mathcal{S}$ such that $|\mathcal{V}'| \geq \delta n/(2\alpha)$ and $\mathcal{S}'$ is an $(\alpha, \delta/2)$-system of $\mathcal{V}'$.*

▶ **Lemma 2.4.** *Let $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$) be a list of vector spaces with an (α,δ)-system $\mathcal{S} = (S_1, S_2, \ldots, S_w)$. Then for any linear map $P : \mathbb{R}^\ell \mapsto \mathbb{R}^\ell$, $\mathcal{S}$ is also an (α,δ)-system of $\mathcal{V}' = (V_1', V_2', \ldots, V_n')$, where $V_i' = P(V_i)$.*

Theorem 1.2, will be derived from the following, more general statement, saying that the dimension d is small if there is a (α,δ)-system.

▶ **Definition 2.5** (k-bounded)**.** A vector space $V \subseteq \mathbb{R}^\ell$ is k*-bounded* if $\dim V \leq k$.

▶ **Theorem 2.6.** *Let $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$) be a list of k-bounded vector spaces with an (α,δ)-system and $d = \dim(V_1 + V_2 + \cdots + V_n)$, then $d = O(\alpha^2 k^4/\delta^2)$.*

We can easily reduce the high dimensional Sylvester-Gallai problem in $\mathbb{C}^\ell$ (Theorem 1.2) to the setting of Theorem 2.6 in $\mathbb{R}^\ell$ as shown below.

Proof of Theroem 1.2 using Theorem 2.6. Let $B_j = \{\boldsymbol{v}_{j1}, \boldsymbol{v}_{j2}, \ldots, \boldsymbol{v}_{jk}\}$ be a basis of V_j. Define

$$V_j' = \operatorname{span}\left\{ \operatorname{Re}(\boldsymbol{v}_{j1}), \operatorname{Re}(\boldsymbol{v}_{j2}), \ldots, \operatorname{Re}(\boldsymbol{v}_{jk}), \operatorname{Im}(\boldsymbol{v}_{j1}), \operatorname{Im}(\boldsymbol{v}_{j2}), \ldots, \operatorname{Im}(\boldsymbol{v}_{jk}) \right\} \quad \forall j \in [n].$$

▶ **Claim 2.7.** *$V_j' = \{\operatorname{Re}(\boldsymbol{v}) : \boldsymbol{v} \in V_j\}$ for every $j \in [n]$.*

Proof. For every $\boldsymbol{v}' \in V_j'$, there exist $\lambda_1, \lambda_2, \ldots, \lambda_k, \mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}$ such that

$$\boldsymbol{v}' = \sum_{s=1}^{k} \Big(\lambda_s \operatorname{Re}(\boldsymbol{v}_{js}) + \mu_s \operatorname{Im}(\boldsymbol{v}_{js}) \Big) = \sum_{s=1}^{k} \Big(\lambda_s \operatorname{Re}(\boldsymbol{v}_{js}) + \mu_s \operatorname{Re}(-i\boldsymbol{v}_{js}) \Big)$$
$$= \operatorname{Re}\left(\sum_{s=1}^{k} (\lambda_s - i\mu_s)\boldsymbol{v}_{js} \right).$$

Since $\lambda_1, \lambda_2, \ldots, \lambda_k, \mu_1, \mu_2, \ldots, \mu_k$ can take all values in $\mathbb{R}$, we can see the claim is proved. ◀

▶ **Claim 2.8** ([1, Lemma 2.1])**.** *Given a set A with $r \geq 3$ elements, we can construct a family of $r^2 - r$ triples of elements in A with following properties: 1) Every triple contains three distinct elements; 2) Every element of A appears in exactly $3(r-1)$ triples; 3) Every pair of two distinct elements in A is contained together in at most 6 triples.*

We call a $2k$-dimensional subspace $U \subset \mathbb{C}^\ell$ *special* if it contains at least three of $V_1, V_2, \dots, V_n$. We define the *size* of a special space as the number of spaces among $V_1, V_2, \dots, V_n$ contained in it. For a special space with size r, we take the $r^2 - r$ triples of indices of the spaces in it with the properties in Claim 2.8. Let $\mathcal{S}$ be the family of all these triples. We claim that $\mathcal{S}$ is a $(6, 3\delta)$-system of $\mathcal{V} = (V_1', V_2', \dots, V_n')$.

For every triple $\{j_1, j_2, j_3\} \in \mathcal{S}$, we can see that $V_{j_1}, V_{j_2}, V_{j_3}$ are contained in the same $2k$-dimensional special space. And by $V_{j_1} \cap V_{j_2} = \{\mathbf{0}\}$, the space must be $V_{j_1} + V_{j_2}$ and hence $V_{j_3} \subseteq V_{j_1} + V_{j_2}$. By Claim 2.7,

$$V_{j_3}' = \{\mathrm{Re}(\boldsymbol{v}) : \boldsymbol{v} \in V_{j_3}\} \subseteq \{\mathrm{Re}(\boldsymbol{u}) + \mathrm{Re}(\boldsymbol{w}) : \boldsymbol{u} \in V_{j_1}, \boldsymbol{w} \in V_{j_2}\} = V_{j_1}' + V_{j_2}'.$$

Similarly, $V_{j_1}' \subseteq V_{j_2}' + V_{j_3}'$ and $V_{j_2}' \subseteq V_{j_1}' + V_{j_3}'$. One can see that every pair in $[n]$ appears in at most 6 triples because the corresponding two spaces are contained in at most one special space, and the pair appears at most 6 times in the triples constructed from this special space. For every $j \in [n]$, there are at least δn values of $j' \in [n] \setminus \{j\}$ such that there is a special space containing V_j and $V_{j'}$. This implies that the number of triples that j appears in is

$$\sum_{\substack{\text{special space } U \\ V_j \subseteq U}} 3\big(\mathrm{size}(U) - 1\big) = 3 \sum_{\substack{\text{special space } U \\ V_j \subseteq U}} \big|\{j' \neq j : V_{j'} \subseteq U\}\big| \geq 3\delta n.$$

Therefore $\mathcal{S}$ is a $(6, 3\delta)$-system of $\mathcal{V}$. By Theorem 2.6,

$$\dim(V_1' + V_2' + \dots + V_n') = O(6^2 (2k)^4 / (3\delta)^2) = O(k^4/\delta^2).$$

Note that

$$\begin{aligned} V_1 + V_2 + \dots + V_n &\subseteq \mathrm{span}\left\{\mathrm{Re}(\boldsymbol{v}_{js}), \mathrm{Im}(\boldsymbol{v}_{js})\right\}_{j \in [n], s \in [k]} && \text{(span with complex coefficients)}, \\ V_1' + V_2' + \dots + V_n' &= \mathrm{span}\left\{\mathrm{Re}(\boldsymbol{v}_{js}), \mathrm{Im}(\boldsymbol{v}_{js})\right\}_{j \in [n], s \in [k]} && \text{(span with real coefficients)}. \end{aligned}$$

We thus have $\dim(V_1 + V_2 + \dots + V_n) \leq \dim(V_1' + V_2' + \dots + V_n') = O(k^4/\delta^2)$. ◀

3 A generalization of Barthe's Theorem

We prove Theorem 1.4 in the following 3 subsections. In the fourth and last subsection, we state a convenient variant of the theorem (Theorem 3.8) that will be used later in the proof of our main result. The idea of the proof is similar to [3] (see also [6, Section 5]), which considers the maximum point of a function, and using the fact that all derivatives are 0 the result is proved. Here we consider a similar function f defined in Section 3.1. However, since our problem is more complicated, it is unclear whether we can find a maximum point at which all derivatives are 0. Instead we will show that there is a point with very small derivatives in Section 3.2, which is sufficient for our proof of the theorem in Section 3.3.

3.1 The function and basic properties

Let $k_1, k_2, \dots, k_n$ be the dimensions of $V_1, V_2, \dots, V_n$ respectively and $m = k_1 + k_2 + \dots + k_n$. Throughout our proof, we use pairs (i, j) with $i \in [n]$, $j \in [k_i]$ to denote the element of $[m]$ of position $\sum_{i' < i} k_{i'} + j$. We define a vector $\boldsymbol{\gamma} \in \mathbb{R}^m$ as

$$\gamma_{ij} = p_i \qquad \forall i \in [n], j \in [k_i].$$

For every $i \in [n]$, we fix $\{\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, \dots, \boldsymbol{v}_{ik_i}\}$ to be some basis of V_i (not necessarily orthonormal). A set $I \subseteq [m]$ is called a *good basis set* if

$$I = \bigcup_{i \in H} \{(i,1), (i,2), \dots, (i,k_i)\}$$

for some $\mathcal{V}$-admissible basis set H. We can see that for any good basis set I, the set $\{\boldsymbol{v}_{ij} : (i,j) \in I\}$ is a basis of $\mathbb{R}^\ell$. For a list of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_q$ ($q \in \mathbb{Z}^+$), we use $[\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_q]$ to denote the matrix consisting of columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_q$.

Let $\mathbf{O}(s)$ be the group of $s \times s$ orthogonal matrices. The function $f : \mathbb{R}^m \times \mathbf{O}(k_1) \times \mathbf{O}(k_2) \times \cdots \times \mathbf{O}(k_n) \mapsto \mathbb{R}$ is defined as

$$f(\boldsymbol{t}, R_1, \dots, R_n) = \langle \boldsymbol{\gamma}, \boldsymbol{t} \rangle - \ln \det \left(\sum_{i \in [n], j \in [k_i]} e^{t_{ij}} \boldsymbol{x}_{ij} \boldsymbol{x}_{ij}^T \right),$$

where, for every $i \in [n]$, the vectors $\boldsymbol{x}_{ij}$ are given by

$$[\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{ik_i}] = [\boldsymbol{v}_{i1}, \dots, \boldsymbol{v}_{ik_i}] R_i.$$

We note that here for every $i \in [n]$, $j \in [k_i]$, $\boldsymbol{x}_{ij}$ is a function of R_i and $\{\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{ik_i}\}$ is another basis of V_i.

The next lemma shows that the function f is bounded over its domain. The proof is similar to Proposition 3 in [3]. The proofs are given in the full version of this paper.

▶ **Lemma 3.1.** *There is a constant $C \in \mathbb{R}$ such that $f(\boldsymbol{t}, R_1, \dots, R_n) \le C$ for all $\boldsymbol{t} \in \mathbb{R}^m$ and $R_i \in \mathbf{O}(k_i)$ ($i \in [n]$).*

3.2 Finding a point with small derivatives

We first define some notation. Let

$$X = \sum_{i \in [n], j \in [k_i]} e^{t_{ij}} \boldsymbol{x}_{ij} \boldsymbol{x}_{ij}^T$$

be a matrix valued function of $\boldsymbol{t}, R_1, R_2, \dots, R_n$. Then

$$f(\boldsymbol{t}, R_1, \dots, R_n) = \langle \boldsymbol{\gamma}, \boldsymbol{t} \rangle - \ln \det(X).$$

Note that X is always a positive definite matrix, since for any $\boldsymbol{w} \neq \boldsymbol{0}$,

$$\boldsymbol{w}^T X \boldsymbol{w} = \sum_{i \in [n], j \in [k_i]} e^{t_{ij}} \langle \boldsymbol{x}_{ij}, \boldsymbol{w} \rangle^2 > 0,$$

when $\boldsymbol{x}_{11}, \dots, \dots, \boldsymbol{x}_{nk_n}$ span the entire space (implied by $V_1 + V_2 + \cdots + V_n = \mathbb{R}^\ell$). Define M to be the $\ell \times \ell$ full rank matrix satisfying $M^T M = X^{-1}$. We note that M is also a function of $\boldsymbol{t}, R_1, R_2, \dots, R_n$.

In a later part of the proof we will show that the linear map obtained from M satisfies the requirement in Theorem 1.4 when $\boldsymbol{t}$, $R_1, R_2, \dots, R_n$ take appropriate values. We first find an appropriate value of $(R_1, R_2, \dots, R_n) = (R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \dots, R_n^*(\boldsymbol{t}))$ for every $\boldsymbol{t} \in \mathbb{R}^m$, and then find some $\boldsymbol{t}^*$ with specific properties.

▶ **Lemma 3.2.** *For every $\boldsymbol{t} \in \mathbb{R}^m$, there exists $\big(R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \dots, R_n^*(\boldsymbol{t})\big)$ satisfying*

1. $f\big(\boldsymbol{t}, R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \dots, R_n^*(\boldsymbol{t})\big) = \max_{R_1, R_2, \dots, R_n} \big\{ f(\boldsymbol{t}, R_1, R_2, \dots, R_n) \big\}$.

2. *For every* $i \in [n]$*, if* $t_{ij} = t_{ij'}$ *for some* $j \neq j' \in [k_i]$*, then*

$$\langle M\boldsymbol{x}_{ij}, M\boldsymbol{x}_{ij'}\rangle = 0,$$

where $[\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ik_i}] = [\boldsymbol{v}_{i1}, \ldots, \boldsymbol{v}_{ik_i}]R_i^*(\boldsymbol{t})$*.*

Proof. The first condition can be satisfied by the compactness of $\mathbf{O}(k_1) \times \mathbf{O}(k_2) \times \cdots \times \mathbf{O}(k_n)$. We will show how to change $(R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \ldots, R_n^*(\boldsymbol{t}))$, which already satisfies the first condition, so that it satisfies the second condition while preserving the first condition.

Fix an $i \in [n]$ and partition the indices of $(t_{i1}, t_{i2}, \ldots, t_{ik_i})$ into equivalence classes $J_1, J_2, \ldots, J_b \subseteq [k_i]$ such that for j, j' in the same class $t_{ij} = t_{ij'}$ and for j, j' in different classes $t_{ij} \neq t_{ij'}$. We use t_{J_r} to denote the value of t_{ij} for $j \in J_r$, and L_{J_r} to denote the matrix consisting of all columns $\boldsymbol{x}_{ij}$ with $j \in J_r$. The terms in X that depend on R_i are

$$\sum_{r\in[b]} \left(e^{t_{J_r}} \sum_{j \in J_r} \boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^T \right) = \sum_{r\in[b]} \left(e^{t_{J_r}} \cdot L_{J_r} L_{J_r}^T \right) = \sum_{r\in[b]} \left(e^{t_{J_r}} \cdot L_{J_r} Q_r Q_r^T L_{J_r}^T \right),$$

where Q_r can be taken to be any $|J_r| \times |J_r|$ orthogonal matrix. This means that if we change $R_i^*(\boldsymbol{t})$ to $R_i^*(\boldsymbol{t})\,\mathrm{diag}(Q_1, \ldots, Q_b)$ (here $\mathrm{diag}(Q_1, \ldots, Q_b)$ denotes the matrix in which the submatrix with row and column indices J_r is Q_r), or equivalently change L_{J_r} to $L_{J_r}Q_r$ for every $r \in [b]$, the matrix X does not change, hence M and f do not change, and the first condition is preserved as f is still the maximum for the fixed $\boldsymbol{t}$.

For every $r \in [b]$, we can find a Q_r such that the columns of $ML_{J_r}Q_r$ are orthogonal (consider the singular value decomposition of ML_{J_r}). Change $R_i^*(\boldsymbol{t})$ to $R_i^*(\boldsymbol{t})\,\mathrm{diag}(Q_1, \ldots, Q_b)$ and the second condition is satisfied while preserving the first condition. Doing this for every i we can obtain an $(R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \ldots, R_n^*(\boldsymbol{t}))$ satisfying both conditions. ◀

From now on we use $R_1^*(\boldsymbol{t}), R_2^*(\boldsymbol{t}), \ldots, R_n^*(\boldsymbol{t})$ to denote the matrices satisfying the conditions in Lemma 3.2.

▶ **Lemma 3.3.** *For any* $\varepsilon > 0$*, there exists* $\boldsymbol{t}^* \in \mathbb{R}^m$ *such that for every* $i \in [n], j \in [k_i]$*.*

$$\left| \frac{\partial f}{\partial t_{ij}} \Big(\boldsymbol{t}^*, R_1^*(\boldsymbol{t}^*), R_2^*(\boldsymbol{t}^*), \ldots, R_n^*(\boldsymbol{t}^*) \Big) \right| \leq \varepsilon.$$

This lemma follows immediately from the following, more general lemma, proved in the full version of this paper.

▶ **Lemma 3.4.** *Let* $\mathcal{A} \subseteq \mathbb{R}^h$ *(*$h \in \mathbb{Z}^+$*) be a compact set. Let* $f : \mathbb{R}^m \times \mathcal{A} \mapsto \mathbb{R}$ *and* $y^* : \mathbb{R}^m \mapsto \mathcal{A}$ *be functions satisfying the following properties:*

1. $f(\boldsymbol{x}, y)$ *is bounded and continuous on* $\mathbb{R}^m \times \mathcal{A}$*.*
2. *For every* $\boldsymbol{x} \in \mathbb{R}^m$*,* $f(\boldsymbol{x}, y^*(\boldsymbol{x})) = \max_{y \in \mathcal{A}}\{f(\boldsymbol{x}, y)\}$*.*
3. *For every fixed* $y \in \mathcal{A}$*,* $f(\boldsymbol{x}, y)$ *as a function of* $\boldsymbol{x}$ *is differentiable on* $\mathbb{R}^m$*.*

Then, for every $\varepsilon > 0$*, there exists an* $\boldsymbol{x}^* \in \mathbb{R}^m$ *such that for every* $i \in [m]$*,*

$$\left| \frac{\partial f}{\partial x_i} \Big(\boldsymbol{x}^*, y^*(\boldsymbol{x}^*) \Big) \right| \leq \varepsilon.$$

3.3 Proof of Theorem 1.4

Fix some $\varepsilon > 0$. We apply Lemma 3.3 and obtain a $\boldsymbol{t}^*$. In the remaining proof we will use X, M and $\boldsymbol{x}_{ij}$ $(i \in [n], j \in [k_i])$ to denote their values when $\boldsymbol{t} = \boldsymbol{t}^*$ and $R_i = R_i^*(\boldsymbol{t}^*)$ $(i \in [n])$.

▶ **Lemma 3.5.** *$\langle M\boldsymbol{x}_{ij}, M\boldsymbol{x}_{ij'}\rangle = 0$ for every $i \in [n]$ and $j \neq j' \in [k_i]$.*

Proof. We fix $i_0 \in [n], j_0 \neq j_0' \in [k_{i_0}]$ and prove $\langle M\boldsymbol{x}_{i_0j_0}, M\boldsymbol{x}_{i_0j_0'}\rangle = 0$. If $t^*_{i_0j_0} = t^*_{i_0j_0'}$, this is guaranteed by Lemma 3.2. We only consider the case that $t^*_{i_0j_0} \neq t^*_{i_0j_0'}$.

Let $\theta \in \mathbb{R}$ be a variable, and define $\boldsymbol{x}'_{ij}$ for $i \in [n]$, $j \in [k_i]$ as follows.

$$\boldsymbol{x}'_{ij} = \begin{cases} \cos\theta \cdot \boldsymbol{x}_{i_0j_0} - \sin\theta \cdot \boldsymbol{x}_{i_0j_0'} & (i,j) = (i_0, j_0), \\ \sin\theta \cdot \boldsymbol{x}_{i_0j_0} + \cos\theta \cdot \boldsymbol{x}_{i_0j_0'} & (i,j) = (i_0, j_0'), \\ \boldsymbol{x}_{ij} & \text{otherwise.} \end{cases}$$

We consider the following function $h : \mathbb{R} \mapsto \mathbb{R}$,

$$h(\theta) = \langle \boldsymbol{\gamma}, \boldsymbol{t}^*\rangle - \ln\det\left(\sum_{i\in[n],j\in[k_i]} e^{t^*_{ij}} \boldsymbol{x}'_{ij}{\boldsymbol{x}'_{ij}}^T\right).$$

▶ **Claim 3.6.** *$h(\theta)$ has a maximum at $\theta = 0$.*

Proof. Let $R(\theta)$ be the $k_{i_0} \times k_{i_0}$ orthogonal matrix obtained from the identity matrix by changing the (j_0, j_0), (j_0', j_0') entries to $\cos\theta$, the (j_0, j_0') entry to $\sin\theta$, and the (j_0', j_0) entry to $-\sin\theta$. We can see $R(0)$ is the identity matrix and

$$[\boldsymbol{x}'_{i_01}, \ldots, \boldsymbol{x}'_{i_0k_{i_0}}] = [\boldsymbol{x}_{i_01}, \ldots, \boldsymbol{x}_{i_0k_{i_0}}]R(\theta).$$

Therefore for all $\theta \in \mathbb{R}$.

$$\begin{aligned} h(\theta) &= f\Big(\boldsymbol{t}^*, R_1^*(\boldsymbol{t}^*), \ldots, R^*_{i_0-1}(\boldsymbol{t}^*), R^*_{i_0}(\boldsymbol{t}^*) \cdot R(\theta), R^*_{i_0+1}(\boldsymbol{t}^*), \ldots, R^*_n(\boldsymbol{t}^*)\Big) \\ &\leq f\Big(\boldsymbol{t}^*, R_1^*(\boldsymbol{t}^*), \ldots, R^*_{i_0-1}(\boldsymbol{t}^*), R^*_{i_0}(\boldsymbol{t}^*), R^*_{i_0+1}(\boldsymbol{t}^*), \ldots, R^*_n(\boldsymbol{t}^*)\Big) \\ &= h(0). \end{aligned}$$

Thus the claim is proved. ◀

Using $\frac{d}{ds}\ln\det(A) = \operatorname{tr}(A^{-1}\frac{d}{ds}A)$ for an invertible matrix A (Theorem 4 in [12, Chapter 9]), we can calculate the derivative of h.

$$\begin{aligned} \frac{dh}{d\theta}(0) =& -\operatorname{tr}\left[X^{-1}\Big(e^{t^*_{i_0j_0}} \frac{d}{d\theta}\Big|_{\theta=0} \boldsymbol{x}'_{i_0j_0}{\boldsymbol{x}'_{i_0j_0}}^T + e^{t^*_{i_0j_0'}} \frac{d}{d\theta}\Big|_{\theta=0} \boldsymbol{x}'_{i_0j_0'}{\boldsymbol{x}'_{i_0j_0'}}^T\Big)\right] \\ =& -\operatorname{tr}\left[X^{-1}\Big(e^{t^*_{i_0j_0}} \frac{d}{d\theta}\Big|_{\theta=0} (\cos\theta\cdot\boldsymbol{x}_{i_0j_0} - \sin\theta\cdot\boldsymbol{x}_{i_0j_0'})(\cos\theta\cdot\boldsymbol{x}_{i_0j_0} - \sin\theta\cdot\boldsymbol{x}_{i_0j_0'})^T\right. \\ &\left. + e^{t^*_{i_0j_0'}} \frac{d}{d\theta}\Big|_{\theta=0} (\sin\theta\cdot\boldsymbol{x}_{i_0j_0} + \cos\theta\cdot\boldsymbol{x}_{i_0j_0'})(\sin\theta\cdot\boldsymbol{x}_{i_0j_0} + \cos\theta\cdot\boldsymbol{x}_{i_0j_0'})^T\Big)\right] \\ =& -e^{t^*_{i_0j_0}}\operatorname{tr}\left[\frac{d}{d\theta}\Big|_{\theta=0} (\cos\theta\cdot M\boldsymbol{x}_{i_0j_0} - \sin\theta\cdot M\boldsymbol{x}_{i_0j_0'})(\cos\theta\cdot M\boldsymbol{x}_{i_0j_0} - \sin\theta\cdot M\boldsymbol{x}_{i_0j_0'})^T\right] \\ & - e^{t^*_{i_0j_0'}}\operatorname{tr}\left[\frac{d}{d\theta}\Big|_{\theta=0} (\sin\theta\cdot M\boldsymbol{x}_{i_0j_0} + \cos\theta\cdot M\boldsymbol{x}_{i_0j_0'})(\sin\theta\cdot M\boldsymbol{x}_{i_0j_0} + \cos\theta\cdot M\boldsymbol{x}_{i_0j_0'})^T\right] \\ =& -e^{t^*_{i_0j_0}}\big[-2\cdot\langle M\boldsymbol{x}_{i_0j_0}, M\boldsymbol{x}_{i_0j_0'}\rangle\big] - e^{t^*_{i_0j_0'}}\big[2\cdot\langle M\boldsymbol{x}_{i_0j_0}, M\boldsymbol{x}_{i_0j_0'}\rangle\big] \\ =& 2(e^{t^*_{i_0j_0}} - e^{t^*_{i_0j_0'}})\cdot\langle M\boldsymbol{x}_{i_0j_0}, M\boldsymbol{x}_{i_0j_0'}\rangle. \end{aligned}$$

Since $h(0)$ is the maximum, we have $\frac{dh}{d\theta}(0) = 0$. By $t^*_{i_0j_0} \neq t^*_{i_0j_0'}$, the above equation implies $\langle M\boldsymbol{x}_{i_0j_0}, M\boldsymbol{x}_{i_0j_0'}\rangle = 0$. ◀

Finally we are able to prove Theorem 1.4.

Proof of Theorem 1.4. With a slight abuse of notation, we also use M to denote the linear map defined by the matrix M. We show that M satisfies the requirement in Theorem 1.4. Let $\boldsymbol{u}_{ij} = M\boldsymbol{x}_{ij}/\|M\boldsymbol{x}_{ij}\|$ ($i \in [n]$, $j \in [k_i]$). Then $\{\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2}, \ldots, \boldsymbol{u}_{ik_i}\}$ is an orthonormal basis of $M(V_i)$, and

$$\mathrm{Proj}_{M(V_i)} = [\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2}, \ldots, \boldsymbol{u}_{ik_i}] \begin{bmatrix} \boldsymbol{u}_{i1}^T \\ \vdots \\ \boldsymbol{u}_{ik_i}^T \end{bmatrix} = \sum_{j=1}^{k_i} \boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T. \tag{1}$$

We define

$$\varepsilon_{ij} = \frac{\partial f}{\partial t_{ij}}\Big(\boldsymbol{t}^*, R_1^*(\boldsymbol{t}^*), R_2^*(\boldsymbol{t}^*), \ldots, R_n^*(\boldsymbol{t}^*)\Big) \in [-\varepsilon, \varepsilon].$$

Again using $\frac{d}{ds}\ln\det(A) = \mathrm{tr}(A^{-1}\frac{d}{ds}A)$ for an invertible matrix A, we have

$$\varepsilon_{ij} = p_i - \mathrm{tr}\left(X^{-1}e^{t_{ij}^*}\boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^T\right) = p_i - e^{t_{ij}^*}\cdot \mathrm{tr}\left(M\boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^T M^T\right) = p_i - e^{t_{ij}^*}\cdot\|M\boldsymbol{x}_{ij}\|^2.$$

By the definition of X and M,

$$M^{-1}(M^T)^{-1} = X = \sum_{i\in[n], j\in[k_i]} e^{t_{ij}^*}\boldsymbol{x}_{ij}\boldsymbol{x}_{ij}^T \quad\Longrightarrow\quad \sum_{i\in[n], j\in[k_i]} e^{t_{ij}^*}(M\boldsymbol{x}_{ij})(M\boldsymbol{x}_{ij})^T = I_{\ell\times\ell}.$$

Therefore

$$\sum_{i\in[n], j\in[k_i]} (p_i - \varepsilon_{ij})\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T = \sum_{i\in[n], j\in[k_i]} e^{t_{ij}^*}\|M\boldsymbol{x}_{ij}\|^2\left(\frac{M\boldsymbol{x}_{ij}}{\|M\boldsymbol{x}_{ij}\|}\right)\left(\frac{M\boldsymbol{x}_{ij}}{\|M\boldsymbol{x}_{ij}\|}\right)^T = I_{\ell\times\ell}.$$

By (1),

$$\left\|\sum_{i=1}^n p_i \,\mathrm{Proj}_{M(V_i)} - I_{\ell\times\ell}\right\| = \left\|\sum_{i\in[n], j\in[k_i]} \varepsilon_{ij}\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T\right\| \le \varepsilon \sum_{i\in[n], j\in[k_i]} \|\boldsymbol{u}_{ij}\boldsymbol{u}_{ij}^T\| \le \varepsilon m.$$

Let $\overline{M} = M/\|M\|$, we can see that $\overline{M}(V_i)$ and $M(V_i)$ are the same linear space, hence

$$\left\|\sum_{i=1}^n p_i \,\mathrm{Proj}_{\overline{M}(V_i)} - I_{\ell\times\ell}\right\| \le \varepsilon m.$$

Take $\varepsilon \to 0$, noting that $\overline{M}$ is contained in a compact set, there must exist a matrix M^* such that

$$\sum_{i=1}^n p_i \,\mathrm{Proj}_{M^*(V_i)} = I_{\ell\times\ell}.$$

It remains to show that M^* is invertible. Assume it is not invertible, then there is a nonzero vector $\boldsymbol{w}$ orthogonal to the range of M^*. We have $\mathrm{Proj}_{M^*(V_i)}(\boldsymbol{w}) = \boldsymbol{0}$ for every $i \in [n]$. This contradicts the fact that the sum of $p_i \,\mathrm{Proj}_{M^*(V_i)}$ is the identity matrix. Therefore M^* is invertible. Thus Theorem 1.4 is proved. ◀

3.4 A convenient form of Theorem 1.4

We give Theorem 3.8 which is implied by Theorem 1.4 and is the form that will be used in our proof. Before stating the theorem, we need to define *admissible sets* and *admissible vectors* as Definition 3.7, which have weaker requirements than admissible basis sets and admissible basis vectors (Definition 1.3) as they are not required to span the entire arrangement.

▶ **Definition 3.7** (admissible set, admissible vector)**.** Given a list of vector spaces $\mathcal{V} = (V_1, V_2, \dots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$), a set $H \subseteq [n]$ is called a $\mathcal{V}$*-admissible set* if $\dim(\sum_{i\in H} V_i) = \sum_{i\in H} \dim(V_i)$, i.e. if every space with index in H has intersection $\{\mathbf{0}\}$ with the span of the other spaces with indices in H. A $\mathcal{V}$*-admissible vector* is any indicator vector $\mathbf{1}_H$ of some $\mathcal{V}$-admissible set H.

▶ **Theorem 3.8.** *Given a list of vector spaces $\mathcal{V} = (V_1, V_2, \dots, V_n)$ ($V_i \subseteq \mathbb{R}^\ell$) and a vector $\boldsymbol{p} \in \mathbb{R}^n$ in the convex hull of all $\mathcal{V}$-admissible vectors. Then there exists an invertible linear map $M : \mathbb{R}^\ell \mapsto \mathbb{R}^\ell$ such that for any unit vector $\boldsymbol{w} \in \mathbb{R}^\ell$,*

$$\sum_{i=1}^{n} p_i \| \operatorname{Proj}_{M(V_i)}(\boldsymbol{w})\|^2 \leq 1,$$

where $\operatorname{Proj}_{M(V_i)}(\boldsymbol{w})$ is the projection of $\boldsymbol{w}$ onto $M(V_i)$.

The simple derivation of Theorem 3.8 from Theorem 1.4 is included in the full version of this paper.

4 Proof of the main Theorem

Theorem 2.6 will follow from the following theorem using a simple recursive argument, provided in the full version of this paper.

▶ **Theorem 4.1.** *Let $\mathcal{V} = (V_1, V_2, \dots, V_n)$ ($V_i \in \mathbb{R}^\ell$) be a list of k-bounded vector spaces with an (α, δ)-system and $d = \dim(V_1 + V_2 + \cdots + V_n)$, then for any $\beta \in (0, 1)$, at least one of these two cases holds:*
1. *$d \leq 40\alpha k^3/(\beta\delta)$,*
2. *There is a sublist of $q \geq \delta n/(20\alpha)$ spaces $(V_{i_1}, V_{i_2}, \dots, V_{i_q})$ such that there are nonzero vectors $\boldsymbol{z}_1 \in V_{i_1}, \boldsymbol{z}_2 \in V_{i_2}, \dots, \boldsymbol{z}_q \in V_{i_q}$ with*

$$\dim(\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_q) \leq \beta d.$$

4.1 Proof of Theorem 4.1 – a special case

In this subsection, we consider the case that all vector spaces are 'well separated'.

▶ **Definition 4.2.** Two vector spaces $V, V' \subseteq \mathbb{R}^\ell$ are τ*-separated* if $|\langle \boldsymbol{u}, \boldsymbol{u}'\rangle| \leq 1 - \tau$ for any two unit vectors $\boldsymbol{u} \in V$ and $\boldsymbol{u}' \in V'$.

We will use the following two simple lemmas about τ-separated spaces (both are proved in the full version of this paper.)

▶ **Lemma 4.3.** *Given two vector spaces $V, V' \subseteq \mathbb{R}^\ell$ that are τ-separated and let $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_{k_1}\}$ and $B' = \{\boldsymbol{u}'_1, \boldsymbol{u}'_2, \dots, \boldsymbol{u}'_{k_2}\}$ be orthonormal bases for V, V' respectively. For any unit vector $\boldsymbol{u} \in V + V'$, if we write $\boldsymbol{u}$ as*

$$\boldsymbol{u} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \cdots + \lambda_{k_1} \boldsymbol{u}_{k_1} + \mu_1 \boldsymbol{u}'_1 + \mu_2 \boldsymbol{u}'_2 + \cdots + \mu_{k_2} \boldsymbol{u}'_{k_2},$$

then the coefficients satisfy $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{k_1}^2 + \mu_1^2 + \mu_2^2 + \cdots + \mu_{k_2}^2 \leq \frac{1}{\tau}$.

▶ **Lemma 4.4.** *Given two vector spaces $V, V' \subseteq \mathbb{R}^\ell$ and let $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k_1}\}$ be an orthonormal basis of V. If V and V' are not τ-separated, there must exist $j \in [k_1]$ such that $\|\operatorname{Proj}_{V'}(\boldsymbol{u}_j)\|^2 \geq (1-\tau)^2/k_1$, where $\operatorname{Proj}_{V'}(\boldsymbol{u}_j)$ is the projection of $\boldsymbol{u}_j$ onto V'.*

We will need the following lower bound for the rank of a diagonal dominating matrix. The proof is included in the full version of this paper.

▶ **Lemma 4.5.** *Let $D = (d_{ij})$ be a complex $m \times m$ matrix and L, K be positive real numbers. If $d_{ii} = L$ for every $i \in [m]$ and $\sum_{i \neq j} |d_{ij}|^2 \leq K$, then $\operatorname{rank}(D) \geq m - K/L^2$.*

The following theorem handles the 'well separated case' of Theorem 4.1.

▶ **Theorem 4.6.** *Let $\mathcal{V} = (V_1, V_2, \ldots, V_n)$ ($V_i \in \mathbb{R}^\ell$) be a list of k-bounded vector spaces with an (α, δ)-system $\mathcal{S} = (S_1, S_2, \ldots, S_w)$ and $d = \dim(V_1 + V_2 + \cdots + V_n)$. If for every $j \in [w]$ and $\{i_1, i_2\} \subseteq S_j$, V_{i_1} and V_{i_2} are τ-separated, then $d \leq \alpha k/(\tau\delta)$.*

Proof. Let $k_1, k_2, \ldots, k_n$ be the dimensions of $V_1, V_2, \ldots, V_n$, and $m = k_1 + k_2 + \cdots + k_n$. For every $i \in [n]$, fix $B_i = \{\boldsymbol{u}_{i1}, \boldsymbol{u}_{i2}, \ldots, \boldsymbol{u}_{ik_i}\}$ to be some orthonormal basis of V_i. We use A to denote the $m \times \ell$ matrix whose rows are $\boldsymbol{u}_{11}^T, \ldots, \ldots, \boldsymbol{u}_{nk_n}^T$. We will bound $d = \operatorname{rank}(A)$ by constructing a high rank $m \times m$ matrix D satisfying $DA = 0$.

For $s \in [m]$, we use $\psi(s) \in [n]$ to denote the number satisfying

$$k_1 + k_2 + \cdots + k_{\psi(s)-1} + 1 \leq s \leq k_1 + k_2 + \cdots + k_{\psi(s)-1} + k_{\psi(s)}.$$

In other words, the s-th row of A is a vector in $B_{\psi(s)}$.

▶ **Claim 4.7.** *For every $s \in [m]$, there is a vector $\boldsymbol{y}_s \in \mathbb{R}^m$ satisfying $\boldsymbol{y}_s^T A = \boldsymbol{0}^T$, $y_{ss} = \lceil \delta n \rceil$, and $\sum_{t \neq s} y_{st}^2 \leq \alpha \lceil \delta n \rceil / \tau$.*

Proof. Say the s-th row of A is $\boldsymbol{u}^T$, where $\boldsymbol{u} \in B_{\psi(s)}$. Let $J \subseteq [w]$ be a set of size $|J| = \lceil \delta n \rceil$ such that for every $j \in J$, S_j contains $\psi(s)$. We construct a vector $\boldsymbol{c}_j$ for every $j \in J$ as follows.

- If S_j contains 3 elements $\{\psi(s), i, i'\}$, we have $\lambda_1, \lambda_2, \ldots, \lambda_{k_i}, \mu_1, \mu_2, \ldots, \mu_{k_{i'}} \in \mathbb{R}$ such that

$$\boldsymbol{u} - \lambda_1 \boldsymbol{u}_{i1} - \lambda_2 \boldsymbol{u}_{i2} - \cdots - \lambda_{k_i} \boldsymbol{u}_{ik_i} - \mu_1 \boldsymbol{u}_{i'1} - \mu_2 \boldsymbol{u}_{i'2} - \cdots - \mu_{k_{i'}} \boldsymbol{u}_{i'k_{i'}} = \boldsymbol{0}.$$

  We can obtain from this equation a vector $\boldsymbol{c}_j$ such that $\boldsymbol{c}_j^T A = \boldsymbol{0}^T$, $c_{js} = 1$, and by Lemma 4.3

$$\sum_{t \neq s} c_{jt}^2 = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{k_i}^2 + \mu_1^2 + \mu_2^2 + \cdots + \mu_{k_{i'}}^2 \leq \frac{1}{\tau}.$$

- If S_j contains 2 elements $\{\psi(s), i\}$, there exist $\lambda_1, \lambda_2, \ldots, \lambda_{k_i}$ with $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{k_i}^2 = 1$ such that

$$\boldsymbol{u} - \lambda_1 \boldsymbol{u}_{i1} - \lambda_2 \boldsymbol{u}_{i2} - \cdots - \lambda_{k_i} \boldsymbol{u}_{ik_i} = \boldsymbol{0}.$$

  We can obtain from this equation a vector $\boldsymbol{c}_j$ such that $\boldsymbol{c}_j^T A = \boldsymbol{0}^T$, $c_{js} = 1$, and

$$\sum_{t \neq s} c_{jt}^2 = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{k_i}^2 = 1 \leq 1/\tau.$$

In either case we obtain a $\boldsymbol{c}_j$ such that $\boldsymbol{c}_j^T A = \boldsymbol{0}^T$, $c_{js} = 1$ and $\sum_{t \neq s} c_{jt}^2 \leq 1/\tau$. We define

$$\boldsymbol{y}_s = \sum_{j \in J} \boldsymbol{c}_j.$$

We have $\boldsymbol{y}_s^T A = \mathbf{0}^T$ and $y_{ss} = \lceil \delta n \rceil$. We consider $\sum_{t \neq s} y_{st}^2$. From the above construction of $\boldsymbol{c}_j$, we can see $c_{jt} \neq 0$ $(t \neq s)$ only when $\psi(t) \neq \psi(s)$ and $\{\psi(s), \psi(t)\} \subseteq S_j$. Hence for every $t \neq s$, there are at most α nonzero values in $\{c_{jt}\}_{j \in J}$. It follows that

$$\sum_{t \neq s} y_{st}^2 = \sum_{t \neq s} \left(\sum_{j \in J} c_{jt} \right)^2 \leq \alpha \sum_{t \neq s} \left(\sum_{j \in J} c_{jt}^2 \right) = \alpha \sum_{j \in J} \left(\sum_{t \neq s} c_{jt}^2 \right) \leq \frac{\alpha \lceil \delta n \rceil}{\tau}.$$

Thus the claim is proved. ◀

Define D to be the matrix consists of rows $\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_m^T$. Then every entry on the diagonal of D is $\lceil \delta n \rceil$, and the sum of squares of all entries off the diagonal is at most $\alpha \lceil \delta n \rceil m / \tau$. Apply Lemma 4.5 on D, and we have

$$\operatorname{rank}(D) \geq m - \frac{\alpha \lceil \delta n \rceil m / \tau}{\lceil \delta n \rceil^2} = m - \frac{\alpha m}{\tau \lceil \delta n \rceil} \geq m - \frac{\alpha k}{\tau \delta}.$$

By $DA = 0$, the rank of A is $d \leq \alpha k / (\tau \delta)$. ◀

4.2 Proof of Theorem 4.1 – general case

Now we prove Theorem 4.1. We assume that the first case of Theorem 4.1 does not hold, i.e. $d > 40 \alpha k^3 / (\beta \delta)$. We will show the second case holds.

▶ **Lemma 4.8.** *If the second case of Theorem 4.1 does not hold, i.e. for any sublist of $q \geq \delta n / (20 \alpha)$ spaces $(V_{i_1}, V_{i_2}, \ldots, V_{i_q})$ and nonzero vectors $\boldsymbol{z}_1 \in V_{i_1}, \boldsymbol{z}_2 \in V_{i_2}, \ldots, \boldsymbol{z}_q \in V_{i_q}$,*

$$\dim(\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_q) > \beta d,$$

then there exists a distribution $\mathcal{D}$ on $\mathcal{V}$-admissible sets and an $I \subseteq [n]$ with $|I| \geq (1 - \delta/(10\alpha))n$ such that for every $i \in I$,

$$\Pr_{H \sim \mathcal{D}}[i \in H] \geq \frac{\beta d}{kn}.$$

Proof. Fix $q = \lceil \delta n / (20 \alpha) \rceil$. By assumption $d > 40 \alpha k^3 / (\beta \delta)$, we have $n \geq d/k > 10 \alpha / \delta$. It follows that $q < \delta n / (10 \alpha)$. We can also see $\delta n / (10 \alpha) < n$ by $\delta / \alpha \leq 3/2$ (Lemma 2.2).

We will find a distribution using the following claim.

▶ **Claim 4.9.** *For a subset $E \subseteq [n]$ of size greater than q, we can find a $\mathcal{V}$-admissible set $H \subseteq E$ with size at least $\beta d / k$.*

Proof. Initially let $H = \emptyset$. In each step we pick an $i_0 \in E$ with $V_{i_0} \bigcap \sum_{i \in H} V_i = \{\mathbf{0}\}$, and add i_0 to H. If such an i_0 does not exist, the procedure terminates. If $|H| < \beta d / k$, then for every $i_0 \in E$, V_{i_0} has a nonzero vector contained in the space $\sum_{i \in H} V_i$, which has dimension at most βd. This contradicts the condition that the second case of Theorem 4.1 does not hold. Hence $|H| \geq \beta d / k$, and the claim is proved. ◀

We repeatedly find a $\mathcal{V}$-admissible sets $H_1, H_2, \ldots$ such that $H_i \subseteq [n] \setminus (H_1 \cup \cdots \cup H_{i-1})$ and $|H_i| \geq \beta d / k$ using the above claim. We can find at most

$$\frac{n - q}{\beta d / k} \leq \frac{nk}{\beta d}$$

such $\mathcal{V}$-admissible sets in total. Let I be the union of these $\mathcal{V}$-admissible sets. We have $|I| \geq n - q \geq (1 - \delta/(10\alpha))n$. Let $\mathcal{D}$ be the uniform distribution on these $\mathcal{V}$-admissible sets. We can see that the probability $\Pr_{H \sim \mathcal{D}}[i \in H] \geq \beta d / (kn)$ for every $i \in I$. Thus the lemma is proved. ◀

Assume the second case of Theorem 4.1 does not hold and apply Lemma 4.8. For $i \in [n]$, we use k_i to denote the dimension of V_i, and p_i to denote $\Pr_{H \sim \mathcal{D}}[i \in H]$. Then $p_i \geq \beta d/(kn)$ for every $i \in I$.

▶ **Lemma 4.10.** *The vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ is in the convex hull of $\mathcal{V}$-admissible vectors.*

Proof. For every $\mathcal{V}$-admissible set H, we use q_H to denote the probability that H is picked according to $\mathcal{D}$, and $\mathbf{1}_H$ to denote the $\mathcal{V}$-admissible vector corresponding to H. Then,

$$\boldsymbol{p} = (p_1, p_2, \ldots, p_n) = \sum_{\mathcal{V}\text{-admissible } H} q_H \mathbf{1}_H$$

and p_i is exactly the probability that $i \in H$. ◀

We apply Theorem 3.8 with the $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$, and obtain an invertible linear map $M : \mathbb{R}^\ell \mapsto \mathbb{R}^\ell$ such that for any unit vector $\boldsymbol{w} \in \mathbb{R}^\ell$,

$$\sum_{i=1}^{n} p_i \| \operatorname{Proj}_{V_i'}(\boldsymbol{w})\|^2 \leq 1,$$

where V_i' denotes $M(V_i)$. Since $p_i \geq \beta d/(kn)$ for every $i \in I$, we have

$$\sum_{i \in I} \| \operatorname{Proj}_{V_i'}(\boldsymbol{w})\|^2 \leq \frac{kn}{\beta d}. \tag{2}$$

We will reduce the problem to the special case discussed in the previous subsection. We say a pair $\{i_1, i_2\} \subseteq [n]$ is *bad* if V_{i_1}', V_{i_2}' are not $\frac{1}{2}$-separated. Let $\mathcal{S} = (S_1, S_2, \ldots, S_w)$ be the (α, δ)-system of $\mathcal{V}$. By Lemma 2.4, $\mathcal{S}$ is also an (α, δ)-system of $\mathcal{V}' = (V_1', V_2', \ldots, V_n')$. We estimate the number of sets among $S_1, S_2, \ldots, S_w$ containing a bad pair.

▶ **Lemma 4.11.** *For every $i_0 \in I$, there are at most $\delta n/(10\alpha)$ values of $i \in I$ such that V_{i_0}' and V_i' are not $\frac{1}{2}$-separated.*

Proof. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k_{i_0}}\}$ be an orthonormal basis of V_{i_0}'. For any i that V_{i_0}' and V_i' are not $\frac{1}{2}$-separated, by Lemma 4.4, there must be $j \in [k_{i_0}]$ such that

$$\| \operatorname{Proj}_{V_i'}(\boldsymbol{u}_j)\|^2 \geq \frac{1}{4k_{i_0}} \geq \frac{1}{4k}.$$

For every $j_0 \in [k_{i_0}]$, we set $\boldsymbol{w} = \boldsymbol{u}_{j_0}$ in inequality (2). The number of i's such that $\| \operatorname{Proj}_{V_i'}(\boldsymbol{u}_{j_0})\| \geq 1/(4k)$ is at most

$$\frac{kn}{\beta d} \bigg/ \frac{1}{4k} = \frac{4k^2 n}{\beta d}.$$

Since there are $k_{i_0} \leq k$ values of $j_0 \in [k_{i_0}]$, the number of i's that V_{i_0}' and V_i' are not $\frac{1}{2}$-separated is at most

$$k \cdot \frac{4k^2 n}{\beta d} \leq \frac{4k^3 n}{\beta d} \leq \frac{\delta n}{10\alpha}.$$

In the last inequality we used the assumption $d > 40\alpha k^3/(\beta\delta)$. ◀

The number of bad pairs is at most

$$|[n] \setminus I| \cdot n + |I| \cdot \frac{\delta n}{10\alpha} \leq \frac{\delta n^2}{10\alpha} + \frac{\delta n^2}{10\alpha} = \frac{\delta n^2}{5\alpha}.$$

We remove all S_j's that contains a bad pair and use $\mathcal{S}'$ to denote the list of the remaining sets. Since each pair appears at most α times, we have removed at most $\delta n^2/5$ sets. Originally we have at least $\delta n^2/3$ sets by Lemma 2.2. Now we have at least $\delta n^2/3 - \delta n^2/5 \geq \delta n^2/10$ sets. By Lemma 2.3, there is a sublist $\mathcal{V}'' = (V'_{i_1}, V'_{i_2}, \ldots, V'_{i_q})$ $(q \geq \delta n/(20\alpha))$ of $\mathcal{V}'$ and a sublist $\mathcal{S}''$ of $\mathcal{S}'$ such that $\mathcal{S}''$ is an $(\alpha, \delta/20)$-system of $\mathcal{V}''$.

Since we have removed all bad pairs, $\mathcal{V}''$ and $\mathcal{S}''$ must satisfy the conditions of Theorem 4.6. By Theorem 4.6,

$$\dim(V'_{i_1} + V'_{i_2} + \cdots + V'_{i_q}) \leq \frac{\alpha k}{\frac{1}{2} \cdot \delta/20} = \frac{40\alpha k}{\delta} \leq \beta d.$$

In the last inequality we used the assumption $d > 40\alpha k^3/(\beta\delta)$. Recall that the linear map M is invertible. So the space $V_{i_1} + V_{i_2} + \cdots + V_{i_q}$ has the same dimension as $V'_{i_1} + V'_{i_2} + \cdots + V'_{i_q}$. Therefore there are $q \geq \delta n/(20\alpha)$ spaces $V_{i_1}, V_{i_2}, \ldots, V_{i_q}$ within dimension βd. The second case of Theorem 4.1 holds.

In summary, under the assumption $d > 40\alpha k^3/(\beta\delta)$ we have shown the second case of Theorem 1.4 is always satisfied. Therefore Theorem 4.1 is proved. ◀

References

1. Boaz Barak, Zeev Dvir, Avi Wigderson, and Amir Yehudayoff. Fractional Sylvester-Gallai theorems. *Proceedings of the National Academy of Sciences*, 110(48):19213–19219, 2013.
2. Boaz Barak, Zeev Dvir, Amir Yehudayoff, and Avi Wigderson. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC'11, pages 519–528, 2011.
3. Franck Barthe. On a reverse form of the Brascamp-Lieb inequality. *Inventiones mathematicae*, 134(2):335–361, 1998.
4. P. Borwein and W. O. J. Moser. A survey of Sylvester's problem and its generalizations. *Aequationes Mathematicae*, 40(1):111–135, 1990.
5. Zeev Dvir. On matrix rigidity and locally self-correctable codes. *computational complexity*, 20(2):367–388, 2011.
6. Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Breaking the quadratic barrier for 3-LCC's over the reals. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC'14, pages 784–793, 2014.
7. 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*, 2, 10 2014.
8. Noam Elkies, Lou M Pretorius, and Konrad J Swanepoel. Sylvester-Gallai theorems for complex numbers and quaternions. *Discrete & Computational Geometry*, 35(3):361–373, 2006.
9. P. Erdös, Richard Bellman, H. S. Wall, James Singer, and V. Thébault. Problems for solution: 4065-4069. *The American Mathematical Monthly*, 50(1):65–66, 1943.
10. Sten Hansen. A generalization of a theorem of Sylvester on the lines determined by a finite point set. *Mathematica Scandinavica*, 16:175–180, 1965.
11. L. M. Kelly. A resolution of the Sylvester-Gallai problem of J.-P. Serre. *Discrete & Computational Geometry*, 1(1):101–104, 1986.
12. Peter D. Lax. *Linear algebra and its applications*. Pure and Applied Mathematics. Wiley-Interscience, 2007.
13. E. Melchior. Uber vielseite der projektive ebene. *Deutsche Math.*, 5:461–475, 1940.
14. J. J. Sylvester. Mathematical question 11851. *Educational Times*, 59:98, 1893.