Forum of Mathematics, Sigma (2014), Vol. 2, e3, 23 pages
doi:10.1017/fms.2014.1

SYLVESTER–GALLAI TYPE THEOREMS FOR APPROXIMATE COLLINEARITY

ALBERT AI[1], ZEEV DVIR[2], SHUBHANGI SARAF[3] and AVI WIGDERSON[4]

[1] Department of Mathematics, University of California, Berkeley, USA
[2] Department of Computer Science and Department of Mathematics, Princeton University, USA;
email: zdvir@princeton.edu, zeev.dvir@gmail.com
[3] Department of Computer Science and Department of Mathematics, Rutgers University, USA
[4] School of Mathematics, Institute for Advanced Study, USA

Received 14 February 2013; accepted 17 January 2014

Abstract

We study questions in incidence geometry where the precise position of points is 'blurry' (for example due to noise, inaccuracy or error). Thus lines are replaced by narrow tubes, and more generally affine subspaces are replaced by their small neighborhood. We show that the presence of a sufficiently large number of approximately collinear triples in a set of points in $\mathbb{C}^d$ implies that the points are *close* to a low dimensional affine subspace. This can be viewed as a stable variant of the Sylvester–Gallai theorem and its extensions. Building on the recently found connection between Sylvester–Gallai type theorems and complex locally correctable codes (LCCs), we define the new notion of *stable* LCCs, in which the (local) correction procedure can also handle small perturbations in the Euclidean metric. We prove that such stable codes with constant query complexity do not exist. No impossibility results were known in any such local setting for more than two queries.

2010 Mathematics Subject Classification: primary 52C35; secondary 68Q99, 94B65

1. Introduction

The Sylvester–Gallai theorem is a statement about configurations of points in $\mathbb{R}^d$ in which there is a certain structure of collinear triples.

THEOREM 1.1 (Sylvester–Gallai). *Suppose $v_1, \ldots, v_n \in \mathbb{R}^d$ are such that for all $i \neq j \in [n]$ there is some $k \in [n] \setminus \{i, j\}$ for which v_i, v_j, v_k are on a line. Then all the points $v_1, \ldots, v_n$ are on a single line.*

This theorem takes *local* information about dependences between points and concludes *global* information about the entire configuration. For more on the

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

history and generalizations of this theorem we refer the reader to the survey [4]. A complex variant of this theorem was proved by Kelly:

THEOREM 1.2 [15]. *Suppose $v_1, \ldots, v_n \in \mathbb{C}^d$ are such that for all $i \neq j \in [n]$ there is some $k \in [n] \setminus \{i, j\}$ for which v_i, v_j, v_k are on a (complex) line. Then all the points $v_1, \ldots, v_n$ lie on a single (complex) plane.*

The global dimension bound given by Kelly's theorem is tight since, over the complex numbers, there are two-dimensional configurations of points satisfying the condition on triples.

In a recent work, Barak *et al.* [2] proved quantitative (or fractional) analogs of Kelly's theorem in which the condition 'for all $i \neq j \in [n]$' is relaxed and we have information only on a large subset of the pairs of points for which there exists a third collinear point (the sets of points satisfying the conditions of the theorem were called δ-SG configurations in [2]).

THEOREM 1.3 [2]. *Suppose $v_1, \ldots, v_n \in \mathbb{C}^d$ are such that for all $i \in [n]$ there exist at least $\delta(n-1)$ values of $j \in [n] \setminus \{i\}$ for which there is $k \in [n] \setminus \{i, j\}$ such that v_i, v_j, v_k are on a line. Then all the points $v_1, \ldots, v_n$ lie in an affine subspace of dimension $13/\delta^2$.*

A more recent work [8] improves the dimension upper bound obtained in the above theorem from $O(1/\delta^2)$ to the asymptotically tight $O(1/\delta)$ and also gives a new proof of Kelly's theorem (when $\delta = 1$ one gets an upper bound of 2 on the dimension).

In this work we consider configurations of points in which there are many triples that are 'almost' collinear, in the sense that there is a line close to all three points (in the usual Euclidean metric on $\mathbb{C}^d$). Equivalently, the points are contained in a narrow tube. Our goal is to prove stable analogs of the above theorems, where stable means that the conclusion of the theorem will not change significantly upon perturbing the point set slightly. Clearly, in such settings one can only hope to prove that there is a low dimensional subspace that *approximates* the set of points. There are many technical issues to discuss when defining approximate collinearity and there are some nontrivial examples showing that word-to-word generalizations of the above theorems do not hold in the approximate-collinearity setting (at least for some of the possible definitions). Nonetheless, we are able to prove several theorems of this flavor for configurations of points satisfying certain 'niceness' conditions. We also study stable variants of error correcting codes (over the reals) which are locally correctable, in which such approximately collinear tuples of points naturally arise from the correcting procedure.

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

In [2], a connection was made between the Sylvester–Gallai theorem and special kinds of error correcting codes called locally correctable codes (LCCs). In these codes, a receiver of a corrupted codeword can recover a single symbol of the codeword correctly, making only a small number of queries to the corrupted word. When studying linear LCCs over the real or complex numbers, one encounters the same kinds of difficulties in trying to convert local dependences into global dimension bounds. Building on this connection, and our ability to analyze 'approximate' linear dependences, we define the notion of *stable* LCCs and show that these do not exist for constant query complexity. Stable LCCs correspond to configurations of points with many approximately dependent small subsets and so our techniques can be used to analyze them.

We note here that understanding the possible intersection structure of tubes in high dimensional real space comes up in connection to other geometric problems, most notably the Euclidean Kakeya problem [18] (we do not, however, see a direct connection between our results and this difficult problem).

Our proof techniques extend those of [2, 8] and rely on high rank properties of sparse matrices whose support is a 'design'. In this work we go a step further and, instead of relying on rank alone, we need to bound the number of small singular values of such matrices.

Organization. In Section 2 we formally state our results for point configurations. The results are stated in several subsections, corresponding to different variants of the problem that we consider. In Section 3 we define stable LCCs and state our results in this scenario. The proofs are given in Sections 4–7.

Notation. We use big 'O' notation to suppress absolute constants only. For two complex vectors $u, v \in \mathbb{C}^d$ we denote their inner product by $\langle u, v\rangle = \sum_{i=1}^{d} u_i \cdot \overline{v_i}$ and use $\|v\| = \sqrt{\langle v, v\rangle}$ to denote the ℓ_2 norm. For an $m \times n$ matrix A, we denote by $\|A\|$ the norm of A as a vector of length mn (that is, the Frobenius norm). The *distance* between two points $u, v \in \mathbb{C}^d$ is defined to be $\|u - v\|$ and is denoted as $\mathsf{dist}(u, v)$. For a set $S \subset \mathbb{C}^d$ and a point $v \in \mathbb{C}^d$ we define $\mathsf{dist}(v, S) = \inf_{u\in S} \mathsf{dist}(u, v)$. We let $S^d \subset \mathbb{C}^{d+1}$ denote the d-dimensional unit sphere in complex $d + 1$-dimensional space. By fixing a basis we can identify each $v \in S^d$ with a $d + 1$-length complex vector of ℓ_2-norm equal to 1.

2. Point configurations

In this section we state our results concerning point configurations. The first section, Section 2.1, deals with the most natural setting—the affine setting—in which we consider sets of points in $\mathbb{C}^d$ with many almost-collinear triples. In

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Section 2.2 we consider the projective setting where the points are located on the sphere and collinearity is replaced with linear dependence. Section 2.3 states a more general theorem from which both the affine and the projective results follow.

2.1. The affine setting.

We begin with the definition of an ϵ-line.

DEFINITION 2.1 (line, line_ϵ). Let $u \neq v \in \mathbb{C}^d$. We define $\mathsf{line}(u, v) = \{\alpha u + (1-\alpha)v \mid \alpha \in \mathbb{C}\}$ to be the complex line passing through u, v. We define $\mathsf{line}_\epsilon(u, v) = \{w \in \mathbb{C}^d \mid \mathsf{dist}(w, \mathsf{line}(u, v)) \leqslant \epsilon\}$.

The following definition will be used to replace the notion of dimension with a more stable definition.

DEFINITION 2.2 ($\dim_\epsilon$). For a set of points $V \subset \mathbb{C}^d$ and $\epsilon > 0$ we denote by $\dim_\epsilon(V)$ the minimal k such that there exists a k-dimensional subspace (the difference of 1 between affine and linear dimension will not be significant in this paper and so we use a linear subspace in the definition) $L \subset \mathbb{C}^d$ such that $\mathsf{dist}(v, L) \leqslant \epsilon$ for all $v \in V$.

To give an idea of the subtleties that arise when dealing with approximate collinearity, take an orthonormal basis $e_1, \ldots, e_d$ in $\mathbb{C}^d$ and consider the set $V = \{e_1, e'_1, \ldots, e_d, e'_d\}$ with $e'_i = (1+\epsilon)e_i$. Clearly, there is no low dimensional subspace that approximates this set of points, even though there are many pairs for which there is a third ϵ-collinear point (e'_i is ϵ-close to the line passing through e_i and any other third point). An obvious solution to this problem is to require that the minimal distance between each pair of points is bounded from below (say by 1), so that the condition of ϵ-collinearity is meaningful. We now describe another, less trivial, example which shows that this condition alone is not sufficient in general.

EXAMPLE 2.3. Let $e_1, \ldots, e_d$ be an orthonormal basis in $\mathbb{C}^d$. Let $v_i = Be_i$, $u_i = (B-1)e_i$ for all $i \in [d]$ and let $V = \{e_i, u_i, v_i \mid i \in [d]\}$ be a set of $n = 3d$ points. Then for all $i, j \in [d]$ we have $u_i \in \mathsf{line}_\epsilon(v_i, e_j)$ and $v_i \in \mathsf{line}_\epsilon(u_i, e_j)$ with $\epsilon = 1/B$. Thus, there are many ϵ-collinear triples in V (as in the conditions of Theorem 2 with $\delta = 1/3$). However, for any subspace L of dimension $o(n)$, the distance of at least one of the points v_i to L must be at least $\Omega(B)$ (this can be shown, for example, using Lemma 4.3).

In this example, we had $\epsilon = 1/B$, where B is roughly equal to the ratio between the smallest and the largest distance, or the 'aspect ratio' of V. We will prevent this scenario by requiring that ϵ will be sufficiently smaller than $1/B$, where B will be the aspect ratio. This motivates the following definition.

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

DEFINITION 2.4 (B-balanced). A set $V \subset \mathbb{C}^d$ is said to be B-balanced if $1 \leqslant \mathsf{dist}(v, v') \leqslant B$ for all $v \neq v' \in V$.

The following theorem gives the most easy to state version of our results.

THEOREM 1. *Let $n, d > 0$ be integers and let $B, \epsilon > 0$ be real numbers with $\epsilon < 1/16B$. Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be B-balanced and suppose that for every $i \neq j \in [n]$ there exists $k \in [n] \setminus \{i, j\}$ such that $v_k \in \mathsf{line}_\epsilon(v_i, v_j)$. Then, $\dim_{\epsilon'}(V) \leqslant O(B^6)$ with $\epsilon' \leqslant O(\epsilon B^{2.5})$.*

Observe that a corollary of this theorem is that the number of points, n, is bounded from above by a function of B. A priori, we did not have this bound since a B-balanced configuration in $\mathbb{C}^d$ can have an unbounded number of points when d grows.

Notice that our definition of ϵ-collinearity is not symmetric in that it depends on the order of the triple. As is shown in Lemma 5.2, this is not an issue for B-balanced configurations, as long as we are willing to replace ϵ with ϵB. For general (that is, nonbalanced) configurations the situation can be more complicated and it is possible that using a stronger collinearity condition (for example, requiring that any permutation of the triple satisfies our condition) is sufficient for obtaining a global dimension bound.

Theorem 1 will be a special case of the following, more general theorem, in which we only have the information of a subset of the pairs (i, j). Assuming that V has many ϵ-collinear triples (for each point), we derive an upper bound on $\dim_{\epsilon'}(V)$ for ϵ' which depends on the other parameters. We also derive a better bound on ϵ' when restricting to a subset of the points.

THEOREM 2. *Let $n, d > 0$ be integers. Let $B, \delta, \epsilon > 0$ be real numbers with $\epsilon < 1/16B$. Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be B-balanced and suppose that for every $i \in [n]$ there are at least $\delta(n-1)$ values of $j \in [n] \setminus \{i\}$ for which there exists $k \in [n] \setminus \{i, j\}$ such that $v_k \in \mathsf{line}_\epsilon(v_i, v_j)$. Then:*

1. $\dim_{\epsilon'}(V) \leqslant O(B^6/\delta^2)$ *with* $\epsilon' \leqslant O(\epsilon B^{2.5}/\delta^{0.5})$.
2. *There exists a subset $V' \subset V$ of size $\Omega(n)$ with $\dim_{\epsilon''}(V') \leqslant O(B^6/\delta^2)$ and $\epsilon'' \leqslant O(B\epsilon)$.*

In both of the above theorems, the parameter B appears in the resulting global dimension bound. We suspect that this dependence can be removed so that the bound on the dimension will be $O(1)$ in Theorem 1 and $O(1/\delta^2)$ (or even $O(1/\delta)$) in Theorem 2. The blowup in ϵ' compared to ϵ is also likely to be suboptimal.

A stronger definition of collinearity, for which Example 2.3 fails, is to require that each point in the triple is ϵ-close to the line spanned by the other two points.

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Let us call such triples *strongly* ϵ-collinear triples. It is easy to see that, in Example 2.3, the triples do not satisfy this stronger definition. Thus, it is possible that one could prove analogs of Theorem 2 for configurations that are not B-balanced using this stronger definition of approximate collinearity.

We conclude this discussion with yet another example showing that, even for the case $\delta = 1$ (that is, the original Sylvester–Gallai condition), the weak definition of ϵ-collinearity requires some balancedness condition (though potentially weaker).

EXAMPLE 2.5. Fix some large $B > 0$. Take an orthonormal basis $e_1, \ldots, e_d \in \mathbb{C}^d$ and define $V = \{0\} \cup \bigcup_{i\in[d]} \left\{B^{i-1}e_i, (B^{i-1}+1)e_i\right\}$. One can verify by induction that for every $u, v \in V$ there is a third point inside $\mathsf{line}_\epsilon(u, v)$ with $\epsilon \approx 1/B$. There is also no low dimensional subspace that approximates V (like for the previous examples).

2.2. The projective setting. Since the definition of ϵ-collinearity (that is, $v_k \in \mathsf{line}_\epsilon(v_i, v_j)$) is sensitive to scaling, a projective statement of Theorem 2, in which these scaling issues do not arise, seems natural. In this setting we consider points on a sphere, and lines are replaced by circles (two-dimensional subspaces intersected with S^d).

DEFINITION 2.6 (circ,circ_ϵ). Let $u, v \in S^d$. We define $\mathsf{circ}(u, v) = \mathsf{span}\{u, v\} \cap S^d$. We define $\mathsf{circ}_\epsilon(u, v) = \{w \in S^d \mid \mathsf{dist}(w, \mathsf{circ}(u, v)) \leqslant \epsilon\}$.

An instructive example in the projective case is the following:

EXAMPLE 2.7. Take V to be a maximal set in S^d with pairwise distances of at least $\mu > 0$ (so $n \approx (c/\mu)^d$ with c a constant). Since every point in S^d is of distance at most μ from one of the points in V (otherwise we could add it) we get that each set $\mathsf{circ}_\mu(v_i, v_j)$ contains at least $\Omega(1/\mu) > 2$ points from V. On the other hand, for any low dimensional subspace L (say, with dimension d' independent of n) almost all points in V will have distance at least $1/100$ from L.

From this example we see that there needs to be some upper bound on ϵ as a function of the minimal distance in the set. We will use the following definition to replace B-balancedness.

DEFINITION 2.8 (μ-separated). A set $V \subset S^d$ is said to be μ-separated if for every $u \neq v \in V$ we have $\min\{\mathsf{dist}(u, v), \mathsf{dist}(u, -v)\} \geqslant \mu$.

We now state our theorem for points on a sphere.

THEOREM 3. *Let $n, d > 0$ be integers and let $\delta, \mu, \epsilon > 0$ be real numbers with $\epsilon < \mu^2/32$. Let $V = \{v_1, \ldots, v_n\} \subset S^d$ be μ-separated and suppose that for every*

$i \in [n]$ there are at least $\delta(n-1)$ values of $j \in [n] \setminus \{i\}$ for which there exists $k \in [n] \setminus \{i, j\}$ such that $v_k \in \mathsf{circ}_\epsilon(v_i, v_j)$. Then

1. $\dim_{\epsilon'}(V) \leqslant O(1/\delta^2\mu^6)$ *with* $\epsilon' \leqslant O(\epsilon/\delta^{0.5}\mu^{2.5})$.
2. *There exists a subset* $V' \subset V$ *of size* $\Omega(n)$ *with* $\dim_{\epsilon''}(V') \leqslant O(1/\delta^2\mu^6)$ *and* $\epsilon'' \leqslant O(\epsilon/\mu)$.

Notice that, comparing with Theorem 3, the parameter μ corresponds to $1/B$. However, the condition on $\epsilon < \mu^2/32$ is more restrictive in this case. We do not know whether this condition can be improved to $\epsilon \leqslant O(\mu)$. As is the case with Theorem 3, we do not expect the dependence in the dimension bound and in ϵ' to be tight.

2.3. The general statement.

Both Theorem 2 and Theorem 3 will follow from a more general statement requiring a set of points with a family of ϵ-dependent triples satisfying certain conditions.

DEFINITION 2.9 ((ϵ, μ)-dependent). We say that a triple of points $u, v, w \in \mathbb{C}^d$ is (ϵ, μ)-dependent if there exist complex numbers α, β, γ with $|\alpha|, |\beta|, |\gamma| \in [\mu, 1]$ such that

$$\|\alpha u + \beta v + \gamma w\| \leqslant \epsilon.$$

DEFINITION 2.10 ((p, g)-design). Let $T \subset \binom{[n]}{3}$ be a family of triples in $[n]$. We say that T is a (p, g)-design if:

1. For all $i \in [n]$ there are at least p triples in T that contain i.
2. For all $i \neq j \in [n]$ there are at most g triples in T containing both i and j.

The following theorem gives a low dimensional subspace that approximates all points in a configuration in which there is a design of triples that are (ϵ, μ)-dependent. Below we will also prove a slightly more refined statement (see Theorem 4.1) giving better distance from L for *many* points in the configuration.

THEOREM 4. *Let $n, d > 0$ be integers and $p, g, \delta, \mu, \epsilon > 0$ be real numbers. Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$, $T \subset \binom{[n]}{3}$ be such that T is a (p, g)-design, and for every $\{i, j, k\} \in T$ the triple v_i, v_j, v_k is (ϵ, μ)-dependent. Then,*

$$\dim_{\epsilon'}(V) \leqslant \frac{2n^2g^2}{p^2\mu^4}$$

with

$$\epsilon' \leqslant \frac{5\epsilon\sqrt{g|T|}}{p\mu^2}.$$

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

A setting of the parameters which will be most relevant to us is when $|T|$ is quadratic in n, p is linear in n and g and μ are constants. In this case we get a constant upper bound on the dimension $\dim_{\epsilon'}(V)$ with $\epsilon' = O(\epsilon)$.

The proof of Theorem 4 is given in the next section with the proofs of Theorems 2 and 3 in Sections 5 and 6 respectively. We give a high level overview of the proof below.

Proof overview. We place the points $v_1, \ldots, v_n$ as rows in a matrix A. We then use the triple family T to construct a matrix M such that:

- M is a $|T| \times n$ matrix whose support is determined by T. More precisely, the nonzero coordinates of the tth row of M, with $t \in T$, will be the three elements in t.
- The values of the entries of M will be, in absolute value, between μ and 1.
- The product $M \cdot A$ will have small Frobenius norm.

We then observe that the matrix $X = M^* M$ is diagonal dominant (its diagonal elements are much larger than its off-diagonal elements). This implies, using the Hoffman–Wielandt inequality, that M has only a few small singular values. From this we get that the columns of A must have small distance (on average) to the span of the small singular vectors of M and so can be approximated well by a low dimensional space. We then show that the same statement holds when one replaces the columns of A with the rows of A (a fact which generalizes the simple fact that the row rank is equal to the column rank). Using the bound on the average distance of rows we argue that there is a large subset that is approximated well by a low dimensional subspace. We then extend this to *all* points using interpolation.

3. Stable locally correctable codes

Before discussing local correction, we briefly mention the exciting recent developments regarding 'standard' (nonlocal) error correcting codes over the reals. Like in the analogous theory over finite fields, one would like to encode (typically via a linear transformation) a vector of entries from a given field $\mathbb{F}$ by a longer one, such that the original message can be decoded even when some entries of the codeword are corrupted. The breakthrough of 'compressed sensing' by Donoho and Candes-Tao, with subsequent developments [see for example **5**, **6**, **10**, **11**, **13**, **17**], has led to an understanding of codes over the reals that is almost as good as in the finite-field case. In particular, there are real-valued codes which achieve the gold-standard of coding theory of constant rate linear codes with efficient encoding and decoding algorithms from a linear number of errors of arbitrary magnitude. Moreover, these codes have *stable* versions which can recover a vector close to the original message even if small errors affect *all* coordinates of the encoding. Our local variant may be viewed as one local analog of such stable codes.

Informally, locally correctable codes (LCCs) are error correcting codes that allow the transmission of information over a noisy channel, so the symbols of the transmitted words have many local dependences between them. The most general definition requires that one can reconstruct (with high probability) *any* coordinate in a possibly corrupted codeword, using a small number of (randomly chosen) queries to the other coordinates. The noise model is adversarial, meaning that the corrupted positions are arbitrary (and not random) and one only has a bound on the total number of errors (which is usually assumed to be a small constant fraction). LCCs are closely related to codes of another type—locally decodable codes (LDCs)—whose study was initiated in a work of Katz and Trevisan [14]. We refer the interested reader to [20] for the relevant background on LDCs and LCCs and their applications in computer science.

The connection between LCCs and the Sylvester–Gallai theorem was first observed in [2]. When studying the special case of *linear* LCCs (that is, LCCs that are given by linear mappings over a field), one can easily show that LCCs are equivalent to point configurations with many linearly dependent small subsets. The general definition of linear LCCs is as follows (we fix the field to be $\mathbb{C}$ but the same definition works for any field). We use $w(v)$ to denote the number of nonzero elements in a vector $v \in \mathbb{C}^n$.

DEFINITION 3.1 (Linear LCC—first definition). A (q, δ)-LCC over $\mathbb{C}$ is a linear subspace $U \subset \mathbb{C}^m$ such that there exists a randomized decoding procedure $D : \mathbb{C}^m \times [m] \mapsto \mathbb{C}$ with the following properties:

1. For all $x \in U$, for all $i \in [m]$ and for all $v \in \mathbb{C}^m$ with $w(v) \leqslant \delta m$ we have that $D(x + v, i) = x_i$ with probability at least $3/4$ (the probability is taken only over the internal randomness of D).

2. For every $y \in \mathbb{C}^m$ and $i \in [m]$, the decoder $D(y, i)$ reads at most q positions in y.

The *dimension* of an LCC is simply its dimension as a subspace of $\mathbb{C}^m$.

It is shown in [2] that, without loss of generality. the decoding procedure is *linear*, in the sense that it first picks a set of at most q coordinates to read and then outputs a linear combination of them (with coefficients in $\mathbb{C}$). This linearity of the decoder implies that, for each coordinate in the code, there are many small subsets of the other coordinates that span it. Since each coordinate corresponds to a row

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

of the generating matrix of the code, we obtain a configuration of points with many dependent small subsets. We will make this formal in the next definition, which is equivalent to the first definition, if one replaces δ with the slightly less good bound of δ/q (when q is constant this change is negligible).

DEFINITION 3.2 (Linear LCC—second definition). We say that a finite set $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ is a (q, δ)-LCC if for every $i \in [n]$ and every set $S \subset [n]$ of size $|S| \leqslant \delta n$ there exists a set $J \subset [n]\backslash S$ with $|J| \leqslant q$ such that $v_i \in \mathsf{span}(v_j \mid j \in J)$.

The main open problem regarding LCCs is that of determining the maximum dimension (as a function of n) when we fix q, δ to be constants. Intuitively, the larger d is, the more 'information' we can transmit using the code (the *rate* of the code if d/n). While the case of $q = 2$ is understood quite well (d is at most logarithmic over finite fields and constant over characteristic zero [2, 3]), it is an open problem to determine the maximum dimension of a q-query LCC when $q > 2$. There are exponential gaps between the known lower and upper bound. For example, when $q = 3$, the best upper bound is $d \leqslant O(\sqrt{n})$ [16, 19] while the best constructions give polylogarithmic d over finite fields and constant d over characteristic zero. We refer the reader to the survey article [7] for more background on LCCs and for an overview of the known constructions.

Due to their roots in coding theory, LCCs were traditionally studied exclusively over finite fields. The study of LCCs over arbitrary fields was initiated in [2] and was motivated by its connection to the Sylvester–Gallai theorem. Further motivation comes from a work connecting LCCs with an approach for constructing *rigid matrices* over infinite fields [9]. We note here that for $q > 2$, the best upper bounds on the dimensions of LCCs are the same, no matter what the field is. This also motivates the study of LCCs over infinite fields as a potentially easier scenario to tackle first, before proceeding to codes over finite fields (where we have fewer techniques).

Our methods enable us to prove strong upper bounds on the dimension of codes that we call *stable LCCs*. Before discussing the relationship between stable and nonstable LCCs we give the formal definition.

DEFINITION 3.3 (span_B). Let $v, u_1, \ldots, u_m \in \mathbb{C}^d$. We say that $v \in \mathsf{span}_B(u_1, \ldots, u_m)$ if there exist $a_1, \ldots, a_m \in \mathbb{C}$ with $|a_i| \leqslant B$ for all i and $v = \sum_{i=1}^m a_i u_i$.

DEFINITION 3.4 (Stable LCC). We say that a finite set $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ is a (q, δ, B, ϵ)-*stable LCC* if for every $i \in [n]$ and every set $S \subset [n]$ of size $|S| \leqslant \delta n$ there exists a set $J \subset [n] \setminus S$ with $|J| \leqslant q$ such that $\mathsf{dist}(v_i, \mathsf{span}_B(v_j \mid j \in J)) \leqslant \epsilon$.

Notice that this definition is not comparable to Definition 3.2. On the one hand, we restrict the linear dependences to use only coefficients of bounded

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

magnitude. On the other hand, we allow the linear combinations to result in an 'approximate' vector, instead of the exact one. To see why the bound on the coefficients is natural (once you allow approximate recovery), notice that the decoder can handle small perturbations *even in the 'correct positions'*. Stated in the scenario of Definition 3.1, suppose that in a received codeword at most δ fraction of the positions are completely changed (to arbitrary values) and, in addition, all other coordinates are perturbed by some small α in Euclidean distance. Then, the decoder can still recover (approximately) the value of a given codeword coordinate by reading at most q other positions, as long as $\alpha \ll \epsilon/qB$. Since each of the read coordinates is multiplied by a coefficient that can be as large as B and the errors sum over q positions, we get at most $\alpha \cdot qB$ resulting error in the output of the decoder. (One can potentially define stable LCCs in this sense (as in Definition 3.1) and then prove (similarly to [2]) that, up to constants, it is equivalent to Definition 3.4 (we did not verify the details).)

The next simple claim shows that Definition 3.4 is also stable in the sense that, perturbing the elements in a stable LCC gives another stable LCC (with slightly less good parameters).

CLAIM 3.5. Let $V = \{v_1, \dots, v_n\} \subset \mathbb{C}^d$ be a (q, δ, B, ϵ)-*stable LCC* and let $V = \{v'_1, \dots, v'_n\} \subset \mathbb{C}^d$ be such that $\mathsf{dist}(v_i, v'_i) \leqslant \alpha$ for all $i \in [n]$. Then V' is a $(q, \delta, B, \epsilon')$-*stable LCC* with $\epsilon' \leqslant \epsilon + (qB + 1)\alpha$.

Proof. Take some $v_i \in V$ and a set $J \subset [n]$ of size $|J| \leqslant q$ such that $\mathsf{dist}(v_i, \mathsf{span}_B(v_j \mid j \in J)) \leqslant \epsilon$. Then, there exist coefficients b_j, $j \in J$, with $|b_j| \leqslant B$ and such that

$$\left\| v_i - \sum_{j \in J} b_j v_j \right\| \leqslant \epsilon.$$

Replacing v_i with v'_i we get that

$$\left\| v'_i - \sum_{j \in J} b_j v'_j \right\| \leqslant \epsilon + \|v_i - v'_i\| + \sum_{j \in J} b_j \|v_j - v'_j\| \leqslant \epsilon + (qB + 1)\alpha. \qquad \square$$

Notice that, if we did not have the bound on the coefficients in the span, the small perturbations would have resulted in large errors in the linear combinations. Intuitively, if u is not in $\mathsf{span}_B(u_1, \dots, u_m)$ then a small perturbation to the u_i may result in u being very far from $\mathsf{span}(u_1, \dots, u_m)$. This explains the need for two separate stability parameters, ϵ and B.

Our main result regarding stable LCCs is the following theorem:

THEOREM 5. *Let* $V = \{v_1, \dots, v_n\} \subset \mathbb{C}^d$ *be a* (q, δ, B, ϵ)*-stable LCC. Then,*

$$\dim_{\epsilon'}(V) \leqslant O((qB/\delta)^4)$$

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

with

$$\epsilon' = O(q^2 B\epsilon/\delta^{1.5}).$$

In particular, when q is a constant and B and δ are fixed, the upper bound on $\dim_{\epsilon'}$ can be interpreted as saying that there *do not exist* stable q-query LCCs, where 'do not exist' means that the amount of information that one can transmit is constant, regardless of the codeword length. The proof of Theorem 5, which follows the same lines as the proof of the Sylvester–Gallai type theorems, works also for the more general setting where V is allowed to be an ordered multiset (that is, when different v_i can repeat several times).

If one sets $\epsilon = 0$, the definition of stable LCC changes into a definition of an LCC with bounded coefficients. That is, the linear dependences are required to be exact (as in the usual definition of an LCC) and, in addition, need to use bounded coefficients. Applying Theorem 5 to this special case, one gets $\epsilon' = 0$ and so obtains the stronger conclusion that the set V is actually *contained* in a low dimensional space. Stated more formally, we have:

COROLLARY 3.6. *Let* $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ *be a* $(q, \delta, B, 0)$*-stable LCC. Then,*

$$\dim(V) \leqslant O((qB/\delta)^4).$$

4. Proof of Theorem 4

We will derive Theorem 4 from the following, more refined, statement.

THEOREM 4.1. *Under the same conditions as in Theorem 4, there exists a subspace* $L \subset \mathbb{C}^d$ *with*

$$\dim(L) \leqslant \frac{2n^2 g^2}{p^2 \mu^4}$$

and such that

$$\sum_{i=1}^{n} \mathit{dist}(v_i, L)^2 \leqslant \frac{4|T|\epsilon^2}{\mu^2 p}.$$

Proof. First, observe that, for convenience, we can take $d = n$, so that the vectors v_i are in $\mathbb{C}^n$. The case $d > n$ is not interesting since we can restrict our attention to the span of the n vectors. The case $d < n$ can be similarly handled by padding each vector with zeros.

Let $m = |T|$. We use T to construct an $m \times n$ matrix M such that there is a one-to-one correspondence between rows of M and elements of T. By our assumptions, for each triple $t = \{i, j, k\} \in T$ there are complex numbers α, β, γ such that $\|\alpha v_i + \beta v_j + \gamma v_k\| \leqslant \epsilon$ and such that $\mu \leqslant |\alpha|, |\beta|, |\gamma| \leqslant 1$. Let s_t

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

denote the row vector in $\mathbb{C}^n$ with the value α in position i, the value β in position j, the value γ in position k and zeros everywhere else. We define M to be the matrix with rows s_t where t goes over all triples in T (in some order).

Next, let A be a complex $n \times n$ matrix whose ith row is the vector v_i. Then, from our definition of the rows of M, we have that the rows of the $m \times n$ matrix

$$E = MA \tag{4.1}$$

all have norm at most ϵ.

The next claim summarizes some of the properties of M that we will use. All three items follow immediately from the fact that T is a (p, g)-design and the bounds on the entries of M.

CLAIM 4.2. Let M be as above and let $M_j \in \mathbb{C}^m$, $j \in [n]$ denote the jth column of M. Then:

1. Each entry of M has absolute value at least μ and at most 1.
2. For each $j \in [n]$, $\|M_j\|^2 \geqslant p\mu^2$.
3. For each $j \neq j' \in [n]$, $\left|\langle M_j, M_{j'}\rangle\right| \leqslant g$.

The main technical ingredient in the proof is the following simple observation regarding the eigenvalues of *diagonal dominant* matrices, that is, matrices in which the diagonal elements are much larger than the off-diagonal elements. This lemma can be viewed as an extension of a folklore result regarding the *rank* of such matrices [see for example 1]. The proof is a simple application of the Hoffman–Wielandt inequality.

LEMMA 4.3. *Let $X = (X_{ij})_{i,j\in[n]}$ be an $n \times n$ complex Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Suppose that for all $i \in [n]$ we have $X_{ii} \geqslant K$, where K is some positive real number. Then,*

$$|\{i \in [n] \mid \lambda_i \leqslant K/4\}| \leqslant \frac{2}{K^2}\sum_{i\neq j}|X_{ij}|^2.$$

Proof. Let D be an $n \times n$ diagonal matrix with $D_{ii} = X_{ii}$ for all $i \in [n]$. Clearly, the eigenvalues of D are $D_{11}, \ldots, D_{nn}$. The Hoffman–Wielandt inequality [12] states that, under some ordering of the eigenvalues of X (without loss of generality, the one that we have chosen) we have

$$\sum_{i\in[n]}|\lambda_i - D_{ii}|^2 \leqslant \|X - D\|^2 = \sum_{i\neq j}|X_{ij}|^2.$$

Using the fact that all D_{ii} are at least K, we get the required bound. □

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Let $\sigma_1, \ldots, \sigma_n$ be the singular values of the matrix M (recall that these are the square roots of the eigenvalues of the PSD matrix M^*M). Let $r_1, \ldots, r_n$ be the corresponding right singular vectors (that is, the corresponding eigenvectors of M^*M). We thus have:

1. $r_1, \ldots, r_n$ form an orthonormal basis of $\mathbb{C}^n$.
2. For each $j \in [n]$, $\|Mr_j\| = \sigma_j$.
3. The vectors $Mr_1, \ldots, Mr_n$ are orthogonal (that is, $\langle Mr_i, Mr_j\rangle = 0$ for $i \neq j$).

Let

$$J = \{j \in [n] \,|\, \sigma_j \leqslant \mu\sqrt{p}/2\}$$

and let

$$L = \mathsf{span}\{r_j \mid j \in J\}.$$

We will now show that L is of small dimension and that most columns of A are close to L. We start by bounding the dimension of L.

Claim 4.4. Let L be as above. Then $|J| = \dim(L) \leqslant ((2n^2g^2)/(p^2\mu^4))$.

Proof. Consider the $n \times n$ matrix $X = M^*M$ with eigenvalues $\sigma_1^2, \ldots, \sigma_n^2$. By Claim 4.2 the diagonal elements of X are all lower bounded by $p\mu^2$ and the off-diagonal elements of X are all upper bounded by g in absolute value. Using Lemma 4.3, and these bounds on the entries of X, we get that

$$\left|\left\{i \in [n] \mid \sigma_i^2 \leqslant p\mu^2/4\right\}\right| \leqslant \frac{2n^2g^2}{p^2\mu^4}.$$

Taking square roots completes the proof. □

Let $u_1, \ldots, u_n$ denote the columns of A. We can write each u_j in the orthonormal basis $r_1, \ldots, r_n$ in a unique way as

$$u_j = \sum_{k=1}^{n} \alpha_{jk} r_k.$$

Observe that

$$\mathsf{dist}(u_j, L)^2 = \sum_{k \notin J} |\alpha_{jk}|^2. \tag{4.2}$$

Denote the rows of the matrix $E = MA$ by e_i, $i \in [m]$, such that $\|e_i\| \leqslant \epsilon$ for all $i \in [m]$. Let $f_1, \ldots, f_n$ be the columns of E and observe that

$$\sum_{j\in[n]} \|f_j\|^2 = \sum_{i\in[m]} \|e_i\|^2 \leqslant m\epsilon^2. \tag{4.3}$$

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

The next claim bounds the sum of distances of the vectors u_j to the subspace L.

CLAIM 4.5. With the above notation, we have

$$\sum_{j=1}^{n} \mathrm{dist}(u_j, L)^2 \leqslant \frac{4m\epsilon^2}{\mu^2 p}.$$

Proof. Using (2) and (3), the orthogonality of the Mr_j and the fact that $\sigma_j > ((\mu\sqrt{p})/2)$ for all $j \notin J$, we have

$$\begin{aligned} m\epsilon^2 &\geqslant \sum_{j\in[n]} \|f_j\|^2 = \sum_{j\in[n]} \|Mu_j\|^2 \\ &= \sum_{j\in[n]} \left\| \sum_{k\in[n]} \alpha_{jk} Mr_k \right\|^2 \\ &= \sum_{j\in[n]} \sum_{k\in[n]} |\alpha_{jk}|^2 \sigma_k^2 \\ &\geqslant \frac{\mu^2 p}{4} \sum_{j\in[n]} \sum_{k\notin J} |\alpha_{jk}|^2 \\ &= \frac{\mu^2 p}{4} \sum_{j\in[n]} \mathrm{dist}(u_j, L)^2. \end{aligned}$$

This proves the claim. □

We now use Claim 4.5 to deduce that many *rows* of A are close to a low dimensional subspace.

CLAIM 4.6. There exists a subspace $L' \subset \mathbb{C}^n$ with $\dim(L') \leqslant ((2n^2g^2)/(p^2\mu^4))$ and such that

$$\sum_{j=1}^{n} \mathrm{dist}(v_j, L')^2 \leqslant \frac{4m\epsilon^2}{\mu^2 p}.$$

Proof. Let Y be an $n \times n$ matrix such that the jth column of Y is the element of L closest to u_j. If we let L' be the span of the *rows* of Y we have $\dim(L') \leqslant \dim(L)$ and, using Claim 4.5,

$$\sum_{j\in[n]} \mathrm{dist}(v_j, L')^2 \leqslant \|Y - A\|^2 = \sum_{j\in[n]} \mathrm{dist}(u_j, L)^2 \leqslant \frac{4m\epsilon^2}{\mu^2 p}.$$ □

This claim completes the proof of Theorem 4.1. □

Proof of Theorem 4 using Theorem 4.1. From Theorem 4.1 we can get a large subset of V that is ϵ'-close to a low dimensional subspace L. To derive

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

the conclusion of Theorem 4, we will show that the rest of the points in V are also close to L, though with a slightly less good bound on the distance. This will follow from showing that, for every point $v \in V$, there are two points $u, w \in V$ that are close to L and such that v is close to the line passing through them. This will imply that v is also close to L. The details follow.

First, apply Theorem 4.1 to get a subspace L such that

$$\dim(L) \leqslant \frac{2n^2g^2}{p^2\mu^4}$$

and such that

$$\sum_{i=1}^{n} \mathsf{dist}(v_i, L)^2 \leqslant \frac{4m\epsilon^2}{\mu^2 p}.$$

Let

$$I = \left\{ i \in [n] \;\middle|\; \mathsf{dist}(v_i, L)^2 > \frac{4gm\epsilon^2}{\mu^2 p^2} \right\}$$

and observe that $|I| < p/g$. Our final step is to argue that the points v_i, $i \in I$, are also close to L' since they are close to the span of two points v_j, v_k with $j, k \notin I$ (using the design properties of T).

Claim 4.7. For each $i \in I$ there are indices $j, k \in [n] \setminus I$ such that $\{i, j, k\} \in T$.

Proof. Fix some $i \in I$. If the claim is false then every triple in T that contains i must have some other element in I. By a pigeonhole argument, there must be an element $j \in I \setminus \{i\}$ and at least $p/|I| > g$ triples containing both i and j, contradicting the design property of T. □

We will need the following simple lemma:

Lemma 4.8. *Let $u, v, w \in \mathbb{C}^d$ be an (ϵ, μ)-dependent triple. Let $L \subset \mathbb{C}^d$ be a subspace with $\mathsf{dist}(v, L), \mathsf{dist}(u, L) \leqslant \rho$ for some $\rho > 0$. Then $\mathsf{dist}(w, L) \leqslant (\epsilon + 2\rho)/\mu$.*

Proof. Let α, β, γ be such that $|\alpha|, |\beta|, |\gamma| \in [\mu, 1]$ and $\|\alpha u + \beta v + \gamma w\| \leqslant \epsilon$. Let $v', u' \in L$ be such that $\|v - v'\|, \|u - u'\| \leqslant \rho$. Then

$$\begin{aligned}
\mathsf{dist}(w, L) &\leqslant \|w + (\alpha/\gamma)v' + (\beta/\gamma)u'\| \\
&\leqslant \|w + (\alpha/\gamma)v + (\beta/\gamma)u\| + \|(\alpha/\gamma)v - (\alpha/\gamma)v'\| \\
&\quad + \|(\beta/\gamma)u - (\beta/\gamma)u'\| \\
&\leqslant \epsilon/|\gamma| + |\alpha/\gamma|\rho + |\beta/\gamma|\rho \\
&\leqslant (\epsilon + 2\rho)/\mu.
\end{aligned}$$
□

Combining Claim 4.7 with Lemma 4.8 we have that each $v_i, i \in [n]$ is ϵ' close to L with $\epsilon' \leqslant (\epsilon + 2\rho)/\mu$, where $\rho = ((2\epsilon\sqrt{gm})/(p\mu))$. Simplifying, we get

$$\epsilon' \leqslant \frac{5\epsilon\sqrt{gm}}{p\mu^2}$$

as was required. This completes the proof of Theorem 4. □

5. Proof of Theorem 2

We start with some preliminary lemmas.

LEMMA 5.1. *Let $\{u, v, w\} \in \mathbb{C}^d$ be B-balanced. If $w \in \mathsf{line}_\epsilon(u, v)$ with $\epsilon < 1/2$ then the triple u, v, w is $(\epsilon, 1/4B)$-dependent. Furthermore, there exists a complex α with $|\alpha| \geqslant 1/4B$ such that $\|w - \alpha u - (1-\alpha)v\| \leqslant \epsilon$.*

Proof. By shifting w to zero we can assume that both u and v have norm bounded by B. By definition, there exists $\alpha \in \mathbb{C}$ such that $\|w - \alpha u - (1-\alpha)v\| \leqslant \epsilon$ and so we only need to show that $|\alpha| \geqslant 1/4B$ (the same argument will apply to $1-\alpha$ by symmetry). Observe that

$$\begin{aligned} 1 &\leqslant \|w - v\| \\ &\leqslant \|w - \alpha u - (1-\alpha)v\| + \|\alpha u\| + \|\alpha v\| \\ &\leqslant \epsilon + 2\alpha B, \end{aligned}$$

which proves the lemma. □

LEMMA 5.2. *Let $\{u, v, w\} \in \mathbb{C}^d$ be B-balanced and let $0 < \epsilon \leqslant 1/2$ be a real number such that $w \in \mathsf{line}_\epsilon(u, v)$. Then $v \in \mathsf{line}_{\epsilon'}(w, u)$ with $\epsilon' = 4\epsilon B$.*

Proof. By Lemma 5.1 there exists a complex α with $|\alpha| \geqslant 1/4B$ such that

$$\|w - \alpha v - (1-\alpha)u\| \leqslant \epsilon.$$

Then

$$\|v - (1/\alpha)w + (1/\alpha - 1)v\| \leqslant \epsilon/\alpha \leqslant 4\epsilon B.$$

This completes the proof. □

LEMMA 5.3. *Let $u, v \in \mathbb{C}^d$ be two distinct points. Let k be the maximum size of a B-balanced set contained in $\mathsf{line}_\epsilon(u, v)$. If $\epsilon < 1/4$, then $k \leqslant 5B$.*

Proof. Suppose $k > 5B$ and let $V = \{v_1, \ldots, v_k\}$ be a B-balanced set contained in $\mathsf{line}_\epsilon(u, v)$. For each v_i let $u_i \in \mathsf{line}(u, v)$ be a point of distance at most ϵ from it. Since the k points $u_1, \ldots, u_k$ are all on a line segment of length at most $2B$, we can apply a pigeonhole argument to conclude that there must be $i \neq j$ with $\mathsf{dist}(u_i, u_j) \leqslant 2B/(k-1)$. This implies $\mathsf{dist}(v_i, v_j) \leqslant 2\epsilon + 2B/(k-1) < 1$, which is a contradiction. □

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Proof of Theorem 2. We define $T \subset \binom{[n]}{3}$ to be the set of triples $\{i, j, k\} \subset [n]$ (with three distinct indices) for which $v_k \in \mathsf{line}_\epsilon(v_i, v_j)$. By Lemma 5.1 we have that for each triple $\{i, j, k\}$ in T, the corresponding triple $v_i, v_j, v_k \in \mathbb{C}^d$ is $(\epsilon, 1/4B)$-dependent.

CLAIM 5.4. T as defined above is a (p, g) design with $p = \delta(n-1)$ and $g < 5B$.

Proof. By the conditions of the theorem, each v_i is contained in at least $\delta(n-1)$ triples that are in T and so the bound on p holds. To prove the bound on g, fix $i \neq j \in [n]$. If the triple $\{i, j, k\}$ appears in T. Then either $v_k \in \mathsf{line}_\epsilon(v_i, v_j)$, $v_i \in \mathsf{line}_\epsilon(v_j, v_k)$ or $v_j \in \mathsf{line}_\epsilon(v_i, v_k)$. In all three cases, we have, using Lemma 5.2, that $v_k \in \mathsf{line}_{\epsilon'}(v_i, v_j)$ with $\epsilon' = 4\epsilon B$. Since $\epsilon < 1/16B$ we have $\epsilon' < 1/4$ and we can apply Lemma 5.3 to conclude that there could be at most $5B$ such triples. □

Observe that we can discard some of the triples in T such that $|T| \leqslant \delta n^2$ and such that T is still a (p, g)-design (simply keep for each i only $\delta(n-1)$-dependent triples).

Plugging the bounds obtained in the above claims and the bound $|T| \leqslant \delta n^2$ into Theorem 4, we get a subspace L with $\dim(L) \leqslant O(B^6/\delta^2)$ and such that $\mathsf{dist}(v_i, L) \leqslant O(\epsilon B^{2.5}/\sqrt{\delta})$ for all $i \in [n]$. The second part of the theorem follows from applying Theorem 4.1.

6. Proof of Theorem 3

We first prove some preliminary lemmas.

LEMMA 6.1. *Suppose $u, v \in S^d$ are such that* $\min\{\mathsf{dist}(u, v), \mathsf{dist}(u, -v)\} = \mu$. *Then, for all complex β, $\mathsf{dist}(u, \beta v) \geqslant \mu/4$.*

Proof. Suppose without loss of generality that $\mathsf{dist}(u, v) = \mu \leqslant \sqrt{2}$. We have

$$\mu = \sqrt{\langle u - v, u - v\rangle} = \sqrt{2 - 2\langle u, v\rangle},$$

which gives $\langle u, v\rangle = 1 - \mu^2/2$. Since $\mathsf{dist}(u, \gamma v)$ is minimized for $\gamma = \langle u, v\rangle$ we have $\mathsf{dist}(u, \beta v) \geqslant \mathsf{dist}(u, (1 - \mu^2/2)v) = \|u - v + (\mu^2/2)v\| \geqslant \|u - v\| - \|(\mu^2/2)v\| \geqslant \mu - \mu^2/2 \geqslant \mu/4$ (for $\mu \leqslant \sqrt{2}$). □

LEMMA 6.2. *Let $u, v, w \in S^d$ be distinct and let $\epsilon, \mu > 0$ be real numbers such that $\epsilon < \mu/8$. Suppose that $\|w - \alpha u - \beta v\| \leqslant \epsilon$ for some complex numbers α, β. If* $\min\{\mathsf{dist}(w, v), \mathsf{dist}(w, -v)\} \geqslant \mu$, *then $|\alpha| > \mu/8$.*

Proof. By the triangle inequality,

$$\|w - \beta v\| \leqslant \|\alpha u\| + \epsilon = |\alpha| + \epsilon.$$

Using Lemma 6.1 we have $\mathsf{dist}(w, \beta v) \geqslant \mu/4$ which gives $|\alpha| \geqslant \mu/4 - \epsilon \geqslant \mu/8$. □

LEMMA 6.3. *Let $u, v, w \in S^d$ be μ-separated and suppose $\epsilon < \mu/8$. Suppose $w \in \mathsf{circ}_\epsilon(u, v)$. Then, there exist complex numbers α, β, γ with $\|\alpha u + \beta v + \gamma w\| \leqslant \epsilon$ and such that $\mu/8 \leqslant |\alpha|, |\beta|, |\gamma| \leqslant 1$.*

Proof. By the assumption, there are α', β' with $\|w - \alpha' u - \beta' v\| \leqslant \epsilon$. If $|\alpha'|$ and $|\beta'|$ are at most 1 then we are done using Lemma 6.2. If not, suppose that $|\alpha'| = \max\{|\alpha'|, |\beta'|\} > 1$ and divide the equation by α' to obtain $\|(1/\alpha')w - u - (\beta'/\alpha')v\| \leqslant \epsilon/|\alpha'| < \epsilon$. Now, all three coefficients are at most 1 in absolute value and, using Lemma 6.2, we have the lower bound $\mu/8$ on $|1/\alpha'|, |\beta'/\alpha'|$. □

LEMMA 6.4. *Let $u, v, w \in S^d$ be distinct. Let $\epsilon, \mu > 0$ be real numbers such that $\epsilon < \mu/8$. Suppose $w \in \mathsf{circ}_\epsilon(u, v)$ and $\min\{\mathsf{dist}(w, v), \mathsf{dist}(w, -v)\} \geqslant \mu$. Then $u \in \mathsf{circ}_{\epsilon'}(w, v)$ with $\epsilon' = 8\epsilon/\mu$.*

Proof. By our assumption, there exist complex numbers α, β such that

$$\|w - \alpha u - \beta v\| \leqslant \epsilon.$$

By Lemma 6.2 we have $|\alpha| > \mu/8$ and so

$$\|u - (1/\alpha)w + (\beta/\alpha)v\| \leqslant 8\epsilon/\mu.$$

This implies that $u \in \mathsf{circ}_{\epsilon'}(w, v)$ as was required. □

LEMMA 6.5. *Let $u, v \in S^d$ be two distinct points. Let k be the maximum size of a μ-separated set contained in $\mathsf{circ}_\epsilon(u, v)$. If $\epsilon < \mu/4$, then $k \leqslant 8/\mu$.*

Proof. Suppose that $k > 8/\mu$ and let $V = \{v_1, \ldots, v_k\}$ be a μ-separated set contained in $\mathsf{circ}_\epsilon(u, v)$. For each v_i let $u_i \in \mathsf{circ}(u, v)$ be a point of distance at most ϵ from it. By a pigeonhole argument, there must be $i \neq j$ with $\min\{\mathsf{dist}(u_i, u_j), \mathsf{dist}(u_i, -u_j)\} \leqslant \pi/k \leqslant \mu/2$. This implies that $\min\{\mathsf{dist}(v_i, v_j), \mathsf{dist}(v_i, -v_j)\} \leqslant 2\epsilon + \mu/2 < \mu$, which is a contradiction. □

Proof of Theorem 3. To reduce to Theorem 4 we will define $T \subset \binom{[n]}{3}$ to be the set of triples $\{i, j, k\} \subset [n]$ for which $v_k \in \mathsf{circ}_\epsilon(v_i, v_j)$.

CLAIM 6.6. Let $\{i, j, k\} \in T$. Then the triple $v_i, v_j, v_k \in \mathbb{C}^d$ is $(\epsilon, \mu/8)$-dependent.

Proof. This is immediate from Lemma 6.3. □

CLAIM 6.7. T as defined above is a (p, g) design with $p = \delta(n-1)$ and $g < 8/\mu$.

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Proof. By the conditions of the theorem, each v_i is contained in at least $\delta(n-1)$ triples that are in T and so the bound on p holds. To prove the bound on g, fix $i \neq j \in [n]$. If the triple $\{i, j, k\}$ appears in T, then $v_k \in \mathsf{circ}_\epsilon(v_i, v_j)$, $v_i \in \mathsf{circ}_\epsilon(v_j, v_k)$ or $v_j \in \mathsf{circ}_\epsilon(v_i, v_k)$. In all three cases, we have, using Lemma 6.4, that $v_k \in \mathsf{circ}_{\epsilon'}(v_i, v_j)$ with $\epsilon' = 8\epsilon/\mu$. Since $\epsilon < \mu^2/32$ we have $\epsilon' < \mu/4$ and we can apply Lemma 6.5 to conclude that there could be at most $8/\mu$ such triples. □

Plugging the bounds obtained in the above claims and the bound $|T| \leqslant \delta n^2$ (which can be obtained by discarding some of the triples in T, as before) into Theorem 4 and into Theorem 4.1 completes the proof. □

7. Proof of Theorem 5

Since the proof follows the same lines as the proof of Theorem 4, we will assume familiarity with the proof of that theorem and only give details where the proofs differ.

We will use the following definition:

DEFINITION 7.1 (LCC matrix). Let M be an $nk \times n$ matrix over $\mathbb{C}$ and let $M_1, \ldots, M_n$ be $k \times n$ matrices such that M is the concatenation of the blocks $M_1, \ldots, M_n$ placed on top of each other (so M_ℓ contains the rows of M numbered $k(\ell-1)+1, \ldots, k\ell$). We say that M is a (k, q)-LCC matrix if, for each $i \in [n]$, the block M_i satisfies the following conditions:

- Each row of M_i has support size at most $q+1$.
- All rows in M_i have the value 1 in position i.
- The supports of two distinct rows in M_i intersect only in position i.

Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be a (q, δ, B, ϵ)-stable LCC and assume without loss of generality that $d = n$ (that is, pad the vectors v_i with zeros so that we can think of them as vectors in $\mathbb{C}^n$). Let A be the $n \times n$ matrix with rows v_i.

CLAIM 7.2. There exists a (k, q)-LCC matrix M with dimensions $nk \times n$ and with $k = \Omega(\delta n/q)$ such that all entries of M have absolute values at most B and such that

$$\|MA\|^2 \leqslant n^2\epsilon^2.$$

Proof. We will show how to construct the $k \times n$ block M_i of M (see Definition 7.1) row by row. Using the definition of stable LCC, there exists a family Q_i of $k = \Omega(\delta n/q)$ disjoint q-tuples of elements of V such that, for each q-tuple $J \in Q_i$,

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

we have $\mathsf{dist}(v_i, \mathsf{span}_B(J)) \leqslant \epsilon$. Each of these q-tuples, J, defines a row vector w_J with 1 in the ith position, B-bounded entries in positions indexed by J, and zeros everywhere else in the following manner: suppose that $v_i = \sum_{j \in J} b_j v_j + e$ with $|b_j| \leqslant B$ for all $j \in J$ and $\|e\| \leqslant \epsilon$; then we define w_j to have 1 in position i and values $-b_j$ in positions $j \in J$ (with zeros in all other positions). Then, we have $\|w_J A\| = \|e\| \leqslant \epsilon$. Taking all these row vectors to construct M_i we get the required bound on $\|MA\|^2$. □

Let $E = MA$, such that $\|E\|^2 \leqslant n^2\epsilon^2$. We now construct another $nk \times n$ matrix R such that $R^T M$ will be diagonal dominant. R will be comprised of n blocks, $R_1, \ldots, R_n$, each of dimensions $k \times n$, such that R_i has entries 1 in the ith column and zeros everywhere else. Notice that the ith row of $R^T M$ is the sum of the rows in the block M_i of M.

Let $\hat{M} = R^T M$ and $\hat{E} = R^T E$, such that $\hat{E} = \hat{M} A$. An application of the Cauchy–Schwarz inequality shows that

$$\|R^T E\|^2 \leqslant n\|E\|^2 \leqslant n^3\epsilon^2.$$

Observe that the diagonal elements of $\hat{M}$ are all equal to k and that the off-diagonal elements of $\hat{M}$ are all of absolute value at most B (since the supports of rows in M_i are disjoint except for the ith coordinate).

We proceed with analyzing the spectrum of $\hat{M}$. Let $r_1, \ldots, r_n$ be the right singular vectors and $\sigma_1, \ldots, \sigma_n$ the corresponding singular values. If we take $X = \hat{M}^* \hat{M}$, then the diagonal elements of X are all at least $K^2 \geqslant k^2$ and the off-diagonal elements can be bounded by $2kB + nB^2 \leqslant O(nB^2)$. If we define

$$L = \mathsf{span}\{r_j \mid \sigma_j < K/2\}$$

we get, using Lemma 4.3, that

$$\dim(L) \leqslant O(n^4 B^4/K^4) = O((qB/\delta)^4).$$

As in the proof of Theorem 4, we consider the columns $u_1, \ldots, u_n$ of A and obtain the bound

$$\sum_{j=1}^{n} \mathsf{dist}(u_j, L)^2 \leqslant 4\|\hat{E}\|^2/K^2 = O(n^3\epsilon^2/K^2).$$

This means that there is a subspace L' with the same dimension as L such that

$$\sum_{i=1}^{n} \mathsf{dist}(v_j, L')^2 \leqslant O(n^3\epsilon^2/K^2).$$

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

Thus, there is a set $V' \subset V$ of size $n' \geqslant (1-\delta/2)n$ such that for all $v' \in V'$ we have $\mathsf{dist}(v', L')^2 \leqslant O(n^2\epsilon^2/\delta K^2) = O(q^2\epsilon^2/\delta^3)$. To finish the proof we observe, using the definition of a stable LCC, that for every $v \in V$ there is a q-tuple $J \subset V'$ with $\mathsf{dist}(v_i, \mathsf{span}_B(J)) \leqslant \epsilon$. Using the bound on the distances of elements of V' to L' and the bound B on the coefficients in the linear combinations in $\mathsf{span}_B(J)$, we get that $\mathsf{dist}(v, L') \leqslant \epsilon + O(qB \cdot (q\epsilon/\delta^{1.5})) = O(q^2B\epsilon/\delta^{1.5})$. This completes the proof of Theorem 5.

Acknowledgements

Z. Dvir was partially supported by NSF grants CCF-1217416, CCF-0832797 and by the Sloan fellowship. S. Saraf and A. Wigderson were partially supported by NSF grant CCF-0832797.

References

[1] N. Alon, 'Perturbed identity matrices have high rank: Proof and applications', *Combin. Probab. Comput.* **18** (1-2) (2009), 3–15.

[2] B. Barak, Z. Dvir, A. Wigderson and A. Yehudayoff, 'Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes', *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11 (ACM, 2011) 519–528.

[3] A. Bhattacharyya, Z. Dvir, A. Shpilka and S. Saraf, 'Tight lower bounds for 2-query lccs over finite fields', *Proc. of FOCS 2011*, 2011, 638–647.

[4] P. Borwein and O. J. Moser, 'A survey of Sylvester's problem and its generalizations', *Aequationes Math.* **40** (1990).

[5] E. Candes and T. Tao, 'Decoding by linear programming', *IEEE Trans. Inform. Theory* **51** (2005), 4203–4215.

[6] D. Donoho, 'Compressed sensing', *IEEE Trans. Inf. Theory* **52** (2006), 1289–1306.

[7] Z. Dvir, 'Incidence theorems and their applications', *Found. Trends Theor. Comput. Sci.* **6** (2012), 257–393.

[8] Z. Dvir, S. Saraf and A. Wigderson, 'Improved rank bounds for design matrices and a new proof of Kelly's theorem', *Forum of Math. Sigma* (2013), (to appear).

[9] Z. Dvir, 'On matrix rigidity and locally self-correctable codes', *Comput. Complexity* **20** (2011), 367–388.

[10] C. Dwork, F. McSherry and K. Talwar, 'The price of privacy and the limits of LP decoding', *Proceedings of the thirty-ninth annual ACM symposium on theory of computing*, STOC '07 (ACM, 2007), 85–94.

[11] V. Guruswami, J. R. Lee and A. Wigderson, 'Expander codes over reals, euclidean sections, and compressed sensing', *Proceedings of the 47th annual Allerton conference on Communication, control, and computing*, Allerton'09 (Piscataway, NJ, USA, IEEE Press, 2009) 1231–1234.

[12] A. J. Hoffman and H. W. Wielandt, 'The variation of the spectrum of a normal matrix', *Duke Math. J.* **20** (1953), 37–39.

[13] B. S. Kashin and V. N. Temlyakov, 'A remark on compressed sensing', 2007. Available at: http://www.dsp.ece.rice.edu/cs/KT2007.pdf.

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press

[14] J. Katz and L. Trevisan, 'On the efficiency of local decoding procedures for error-correcting codes', *32nd ACM Symposium on Theory of Computing (STOC)*, 2000, 80–86.

[15] L. M. Kelly, 'A resolution of the Sylvester–Gallai problem of J. P. Serre', *Discrete Comput. Geom.* **1** (1986), 101–104.

[16] I. Kerenidis and R. de Wolf, 'Exponential lower bound for 2-query locally decodable codes via a quantum argument', *J. Comput. System Sci.* **69** (2004), 395–420.

[17] M. Rudelson and R. Vershynin, 'Geometric approach to error correcting codes and reconstruction of signals', *Int. Math. Res. Not.* **64** (2005), 4019–4041.

[18] T. Tao, 'From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE', *Not. Am. Math. Soc.* **48** (2001), 294–303.

[19] D. Woodruff, 'New lower bounds for general locally decodable codes', *Electronic Colloquium on Computational Complexity (ECCC)*, TR07-006, 2007.

[20] S. Yekhanin, 'Locally decodable codes', *Foundations and trends in theoretical computer science*, 2011, to appear. (Preliminary version available for download at http://research.microsoft.com/en-us/um/people/yekhanin/Papers/LDC_now.pdf).

https://doi.org/10.1017/fms.2014.1 Published online by Cambridge University Press