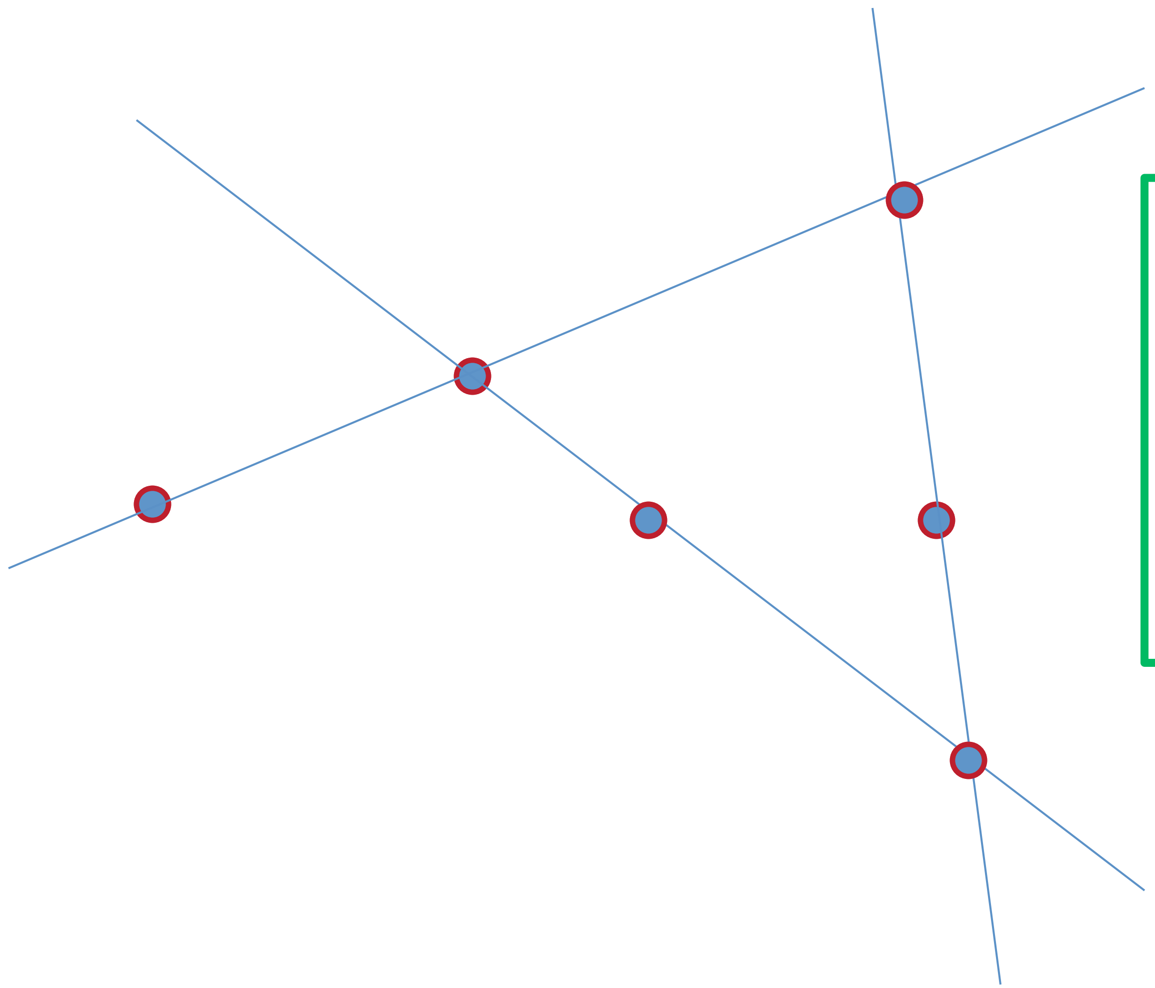
Incidence geometry and applications

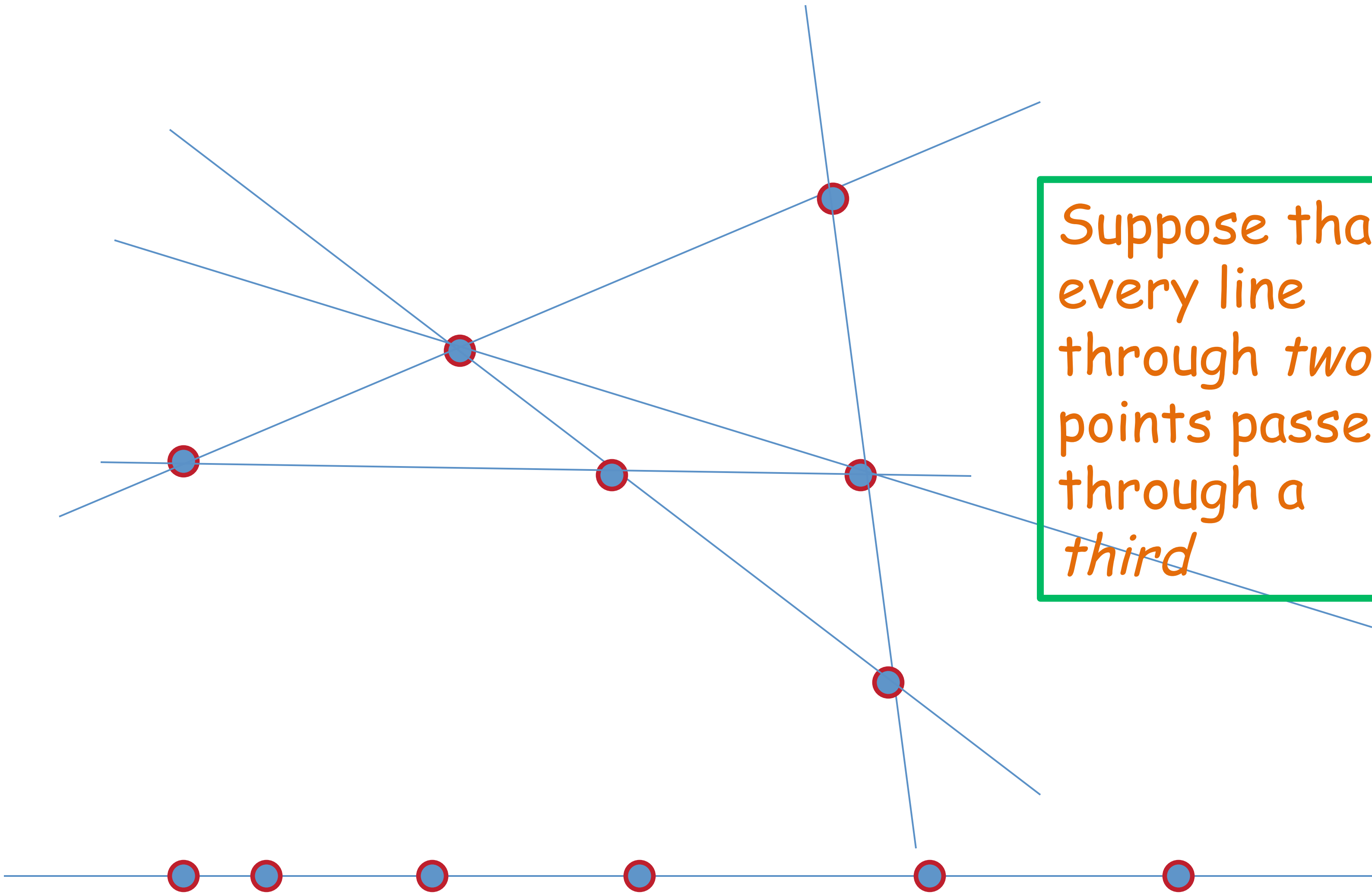
Shubhangi Saraf
Rutgers University

Sylvester-Gallai Theorem (1893)



Suppose that every line through *two* points passes through a *third*

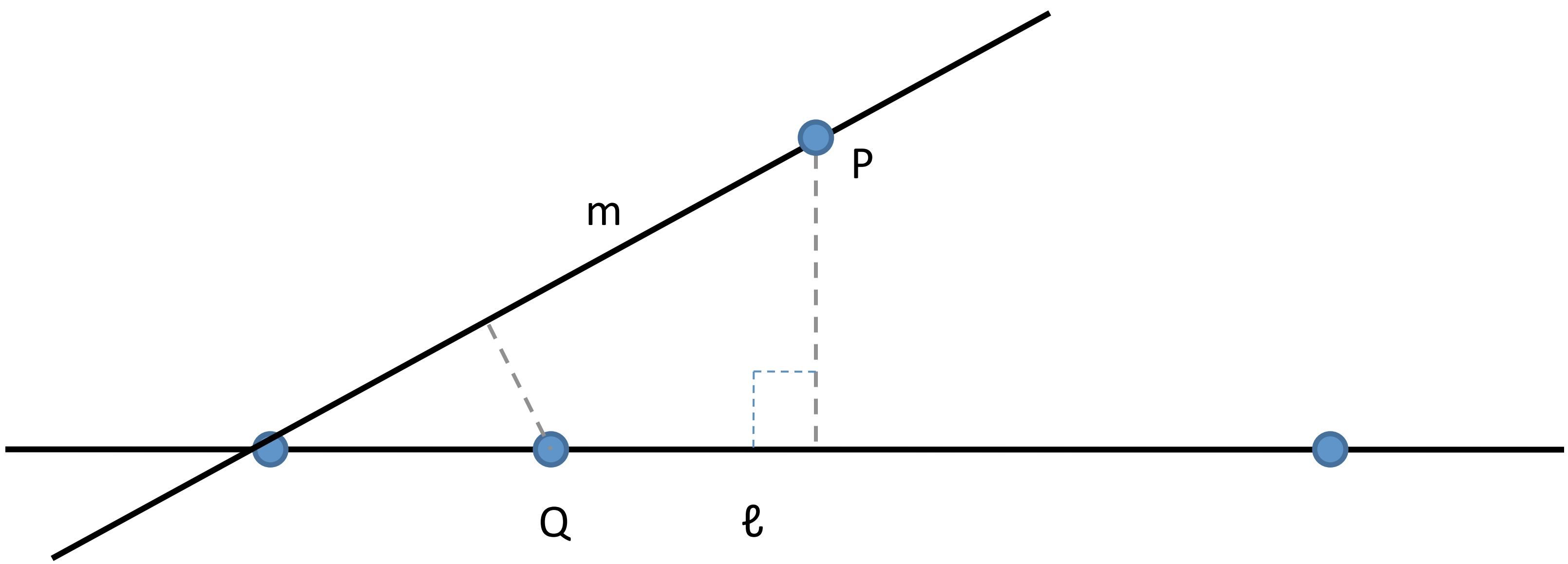
Sylvester-Gallai Theorem (1893)



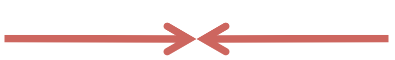
Suppose that every line through *two* points passes through a *third*

Proof of Sylvester-Gallai:

- By contradiction. If possible, for every pair of points, the line through them contains a third.
- Consider the point-line pair with the smallest nonzero distance.



dist(Q, m) < dist(P, ℓ)



Contradiction!

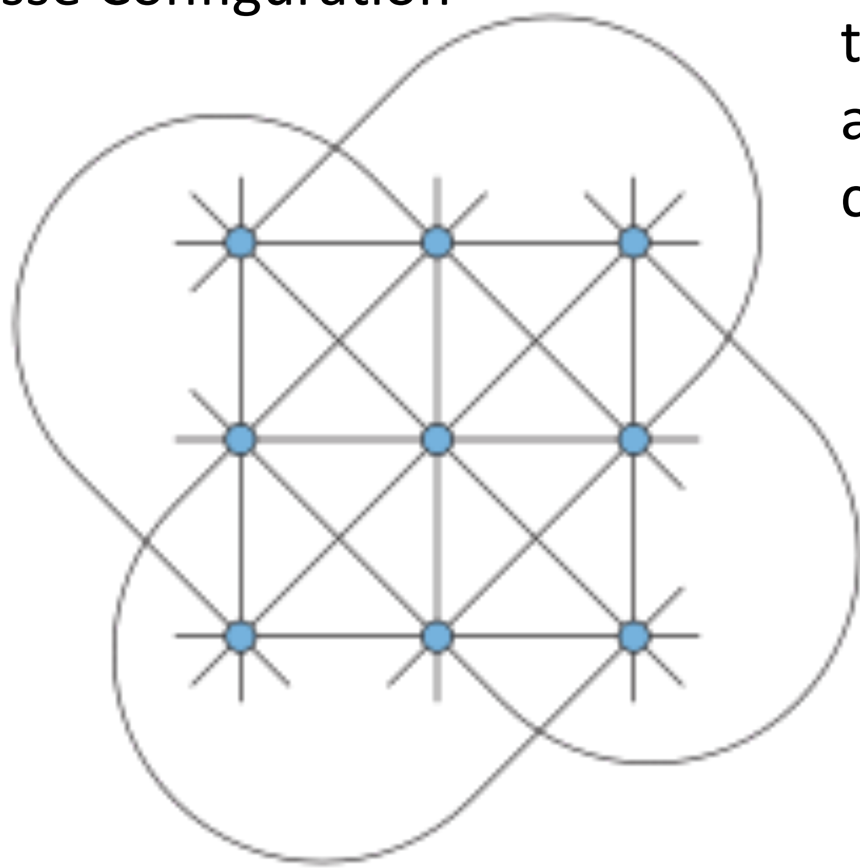
- Several extensions and variations studied
  - Complexes, other fields, colorful, quantitative, approximate, high-dimensional
- Several recent *connections to complexity theory*
  - Structure of arithmetic circuits (DS06, K*S*09, SS11)
  - **Locally Correctable Codes**
- [BDWY11, D*S*W12]: **Quantitative SG thms**
  - Connections of Incidence theorems *to* ***rank bounds for design matrices***
  - Connections to Matrix rigidity
  - 2-query LCCs over the Reals do not exist
- [AD*S*W12]: **Approximate/Stable SG thms**
  - Stable rank of design matrices
  - stable LCCs over R do not exist
- [D*S*W13]: **High dimension/quantitative SG thms**
  - Improved lower bounds for 3 query LCCs over R

The Plan

- Extensions of the Sylvester-Gallai Theorem
  - Complex numbers
  - Quantitative versions
  - Stable versions
  - High dimensions
- Connection to Locally Correctable Codes
- New lower bounds for 3-query LCCs

Points in Complex space

Hesse Configuration



Kelly's Theorem:

For every pair of points in $\mathbf{C}^d$, the line through them contains a third, then all points contained in a complex plane

[Elkies, Pretorius, Swanpoel 2006]:
First elementary proof

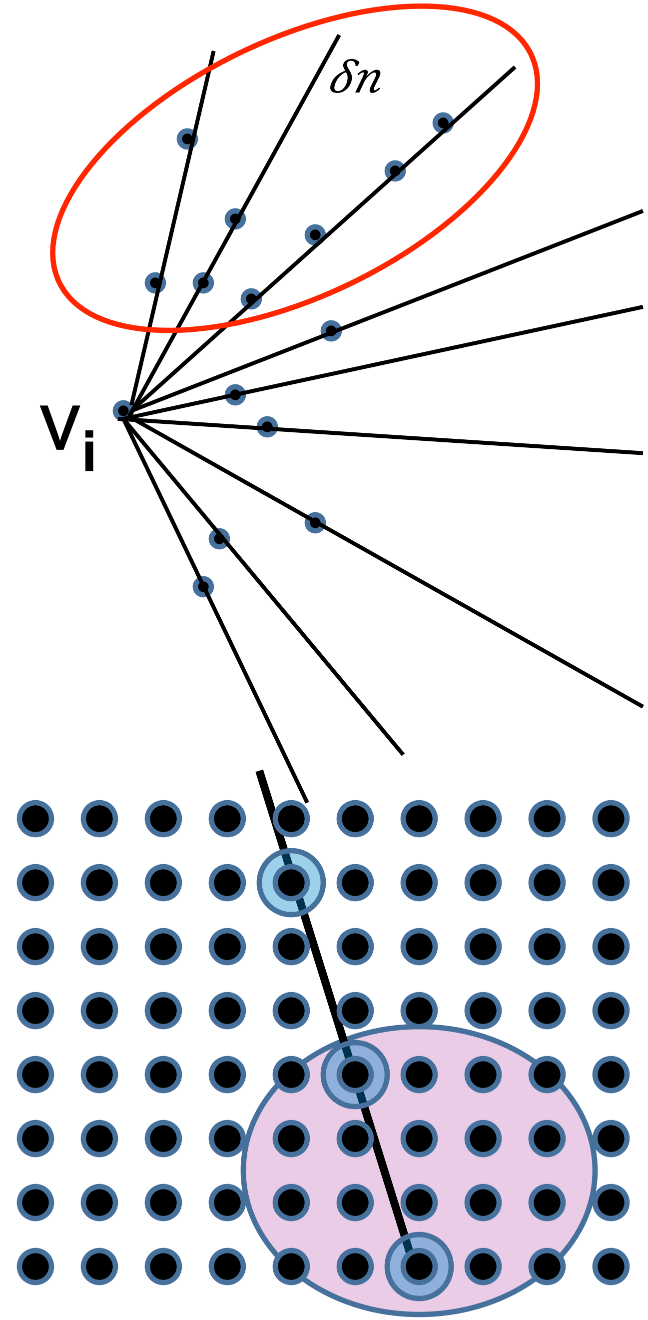
[D*S*W12]:
New proof using basic linear algebra

Quantitative SG

For every point there are at least δn points s.t there is a third point on the line

[BDWY11]: dimension $\leq O(1/\delta^2)$

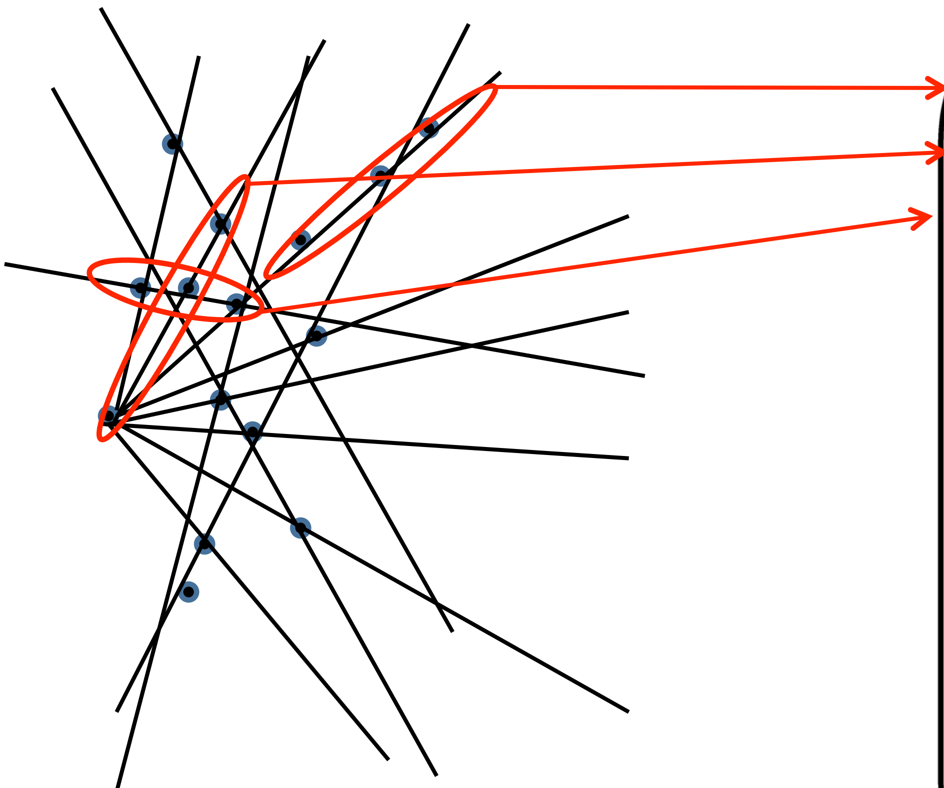
[DSW12]: dimension $\leq O(1/\delta)$



Few words about the proof

Incidence Theorems from Rank Bounds

- Given $v_1, v_2, \ldots, v_n \in \mathbf{C}^d$
- For every *collinear triple* v_i, v_j, v_k, $\exists\ \alpha_i, \alpha_j, \alpha_k$ so that $\alpha_i v_i + \alpha_j v_j + \alpha_k v_k = 0$
- Construct $n \times d$ matrix **V** s.t i^{th} row is v_i
- Construct $m \times n$ matrix $\boldsymbol{A}$ s.t for each collinear triple v_i, v_j, v_k there is a row with $\alpha_i, \alpha_j, \alpha_k$ in positions i,j,k resp.
- $\boldsymbol{A} \cdot \boldsymbol{V} = \mathbf{0}$



$$\begin{pmatrix} 2 & 0 & 2 & 0 & -4 & 0 & 0 \\ 0 & 0 & 7 & -9 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & -4 & 0 & 0 \\ 0 & 0 & 7 & -9 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & -4 & 0 & 0 \\ 0 & 0 & 7 & -9 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & -4 & 0 & 0 \\ 0 & 0 & 7 & -9 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Incidence Theorems from Rank Bounds

- Given set of vectors V, find a matrix A s.t

$A \cdot V = 0$

- Careful pruning of the matrix gives a *design matrix*!

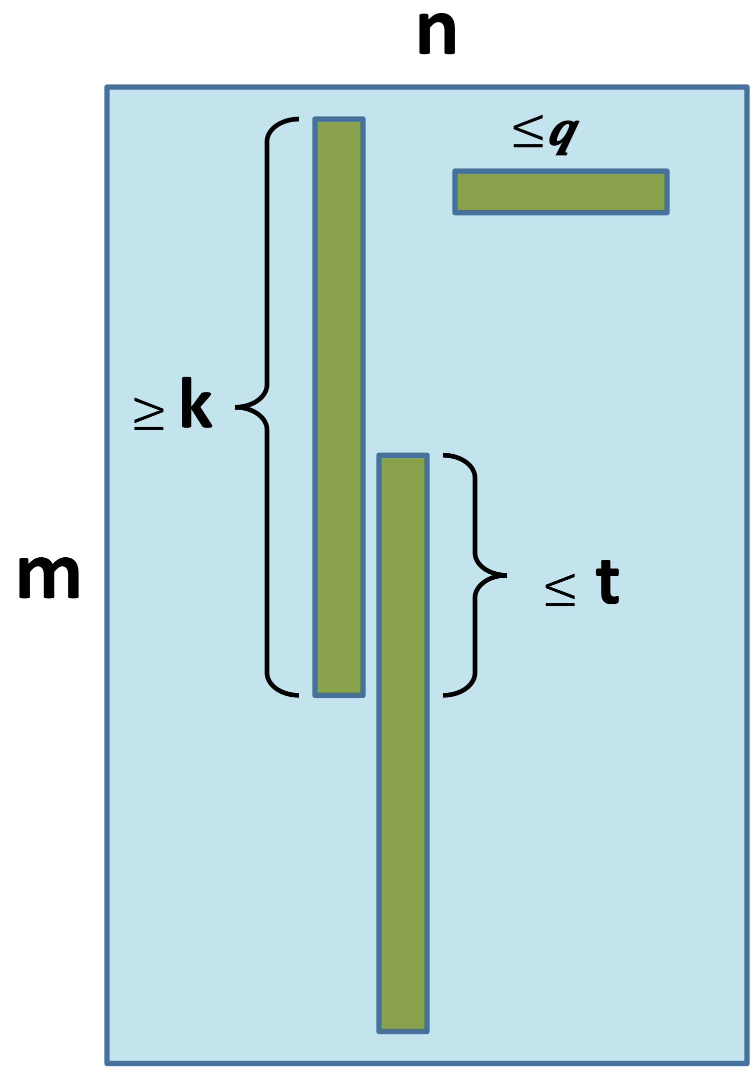
- Want: Upper bound on rank of V

- How?: Lower bound on rank of A

Design Matrices

An m x n matrix is a (q,k,t)-design matrix if:

1. Each row has at most q non-zeros
2. Each column has at least k non-zeros
3. The supports of every two columns intersect in at most t rows



Main Theorem: Rank Bound

Thm [BDWY11, DSW12]: Let A be an m x n complex (q,k,t)-design matrix then:

$$rank \geq n - ntq^2/k$$

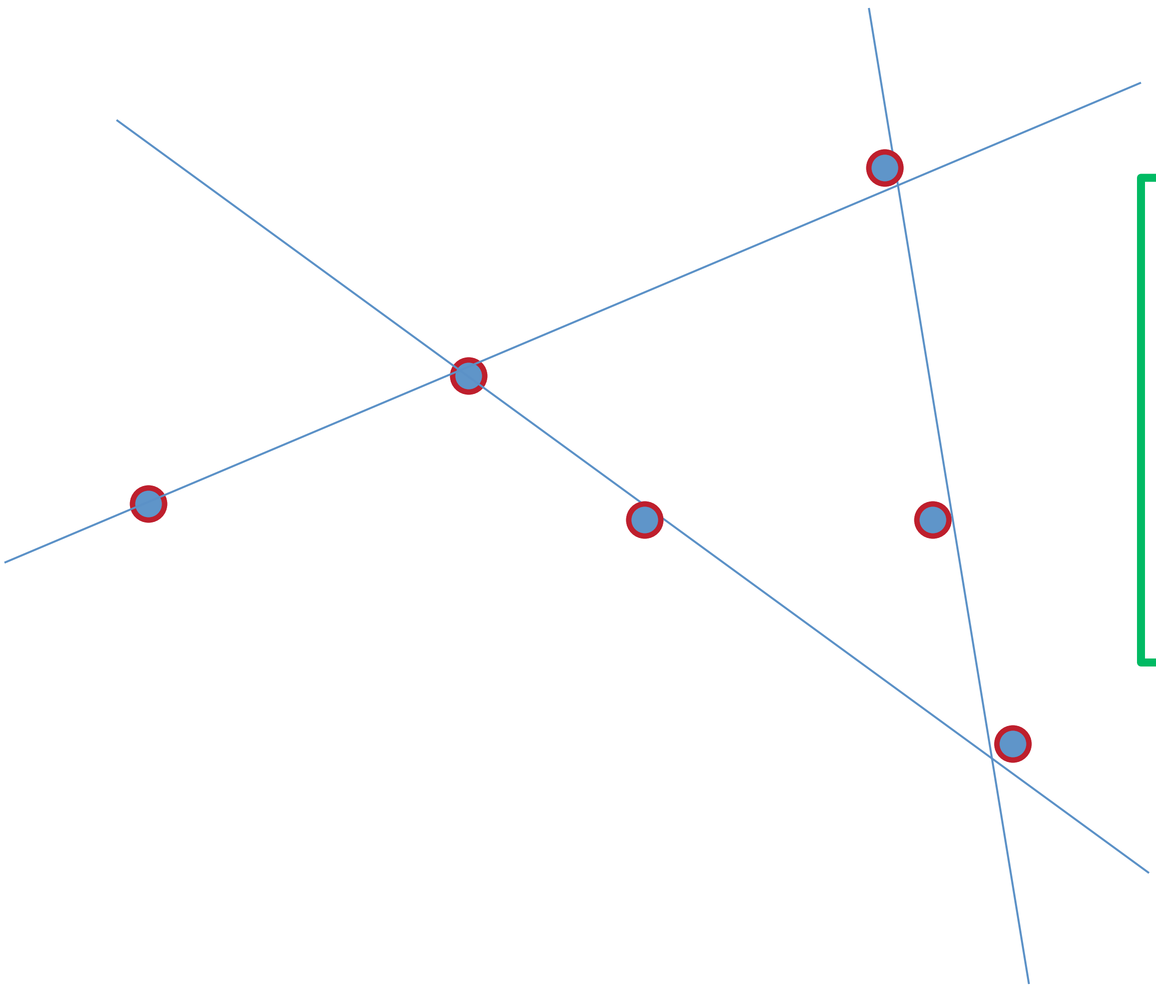
Main idea: Matrix scaling

Holds for any field of char=0 (or very large positive char)

Not true over fields of small characteristic!

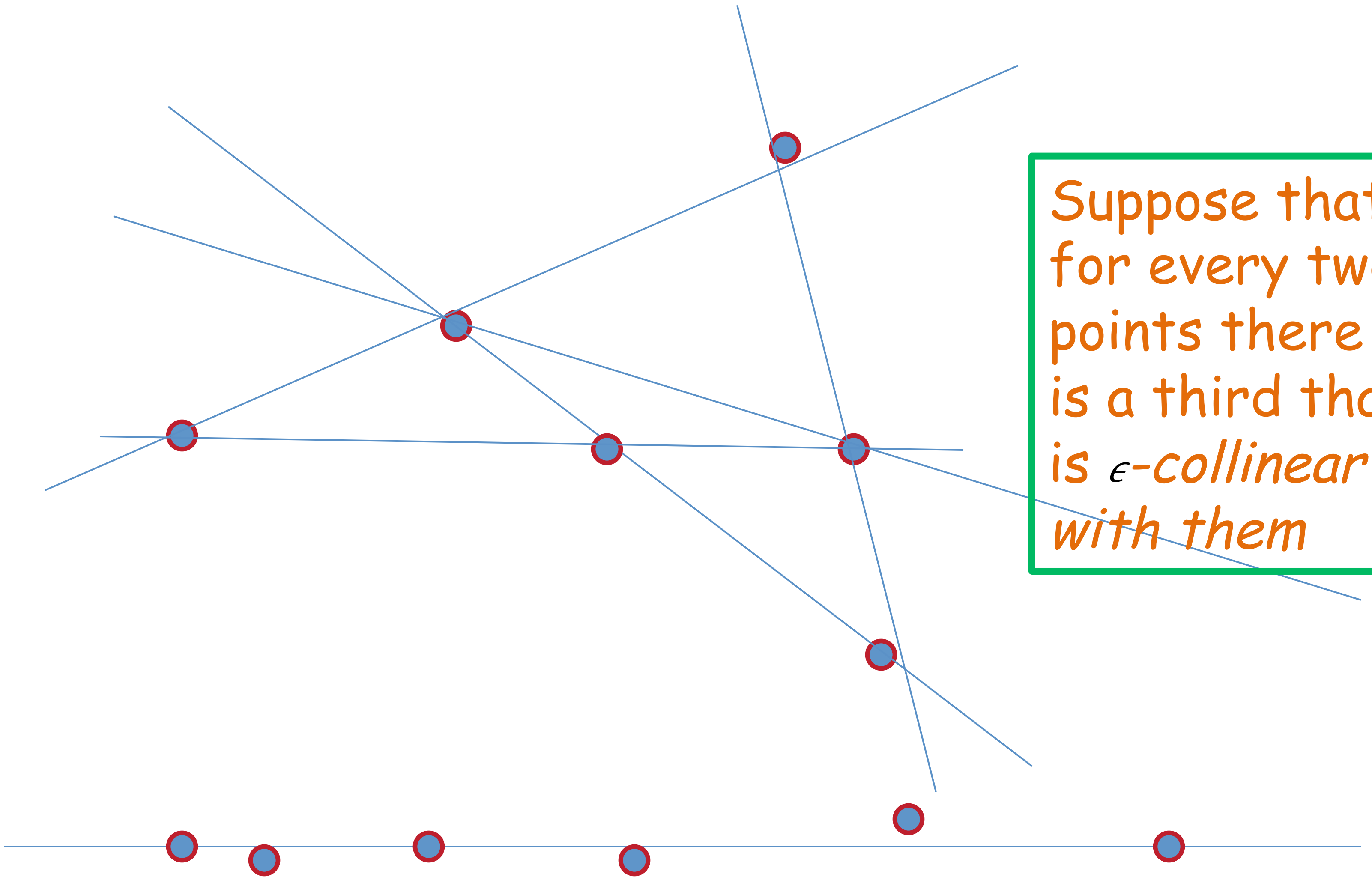
Implies Kelly's theorem (SG over complex numbers)

Stable Sylvester-Gallai Theorem



Suppose that for every two points there is a third that is *ϵ-collinear with them*

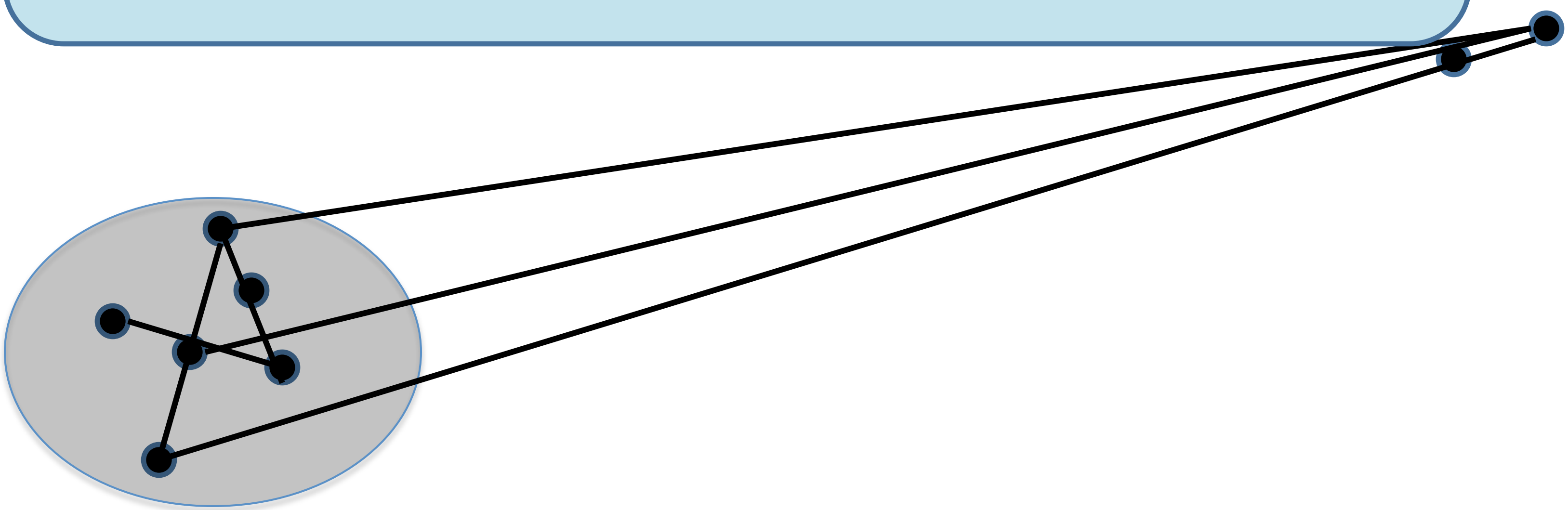
Stable Sylvester Gallai Theorem



Suppose that for every two points there is a third that is ϵ*-collinear with them*

Not true in general ..

$\exists n$ points in $\approx n$ dimensional space s.t for every two points there exists a third point that is ϵ-collinear with them



Stable-SG Theorem [ADSW12]

All distances between 1 and B

Let $v_1, v_2, \ldots, v_n$ be a set of *B-balanced* points in C^d so that for each v_i, v_j there is a point v_k such that the triple is ϵ- *collinear*.

Then

$$Dim_{\epsilon'}(v_1, v_2, \ldots, v_n) \leq O(B^6)$$

The High Dimensional Sylvester-Gallai Theorem

[Hansen 65], [Bonnice-Edelstein 67]

Given a finite set of points in R spanning at least 2k dimensions, there exists a k dimensional hyperplane which is spanned by and contains exactly k+1 points.

[BDWY11] [DSW12]

Extension to the complex numbers

Quantitative versions …

(Bounds far from optimal)

Colorful Sylvester-Gallai [Edelestein-Kelly `66, Kayal-S `10]

Let S be a finite set of points, each colored one of k colors. If every hyperplane containing points of k-1 colors also contains the k^{th} color, then $\dim(S) < k^k$

(Connections to structure of arithmetic circuits)

Dirac-Motzkin conjecture

Green-Tao 2012:

In any set of n noncollinear points in R^2, there must be many ($\geq n/2$) ``ordinary” lines

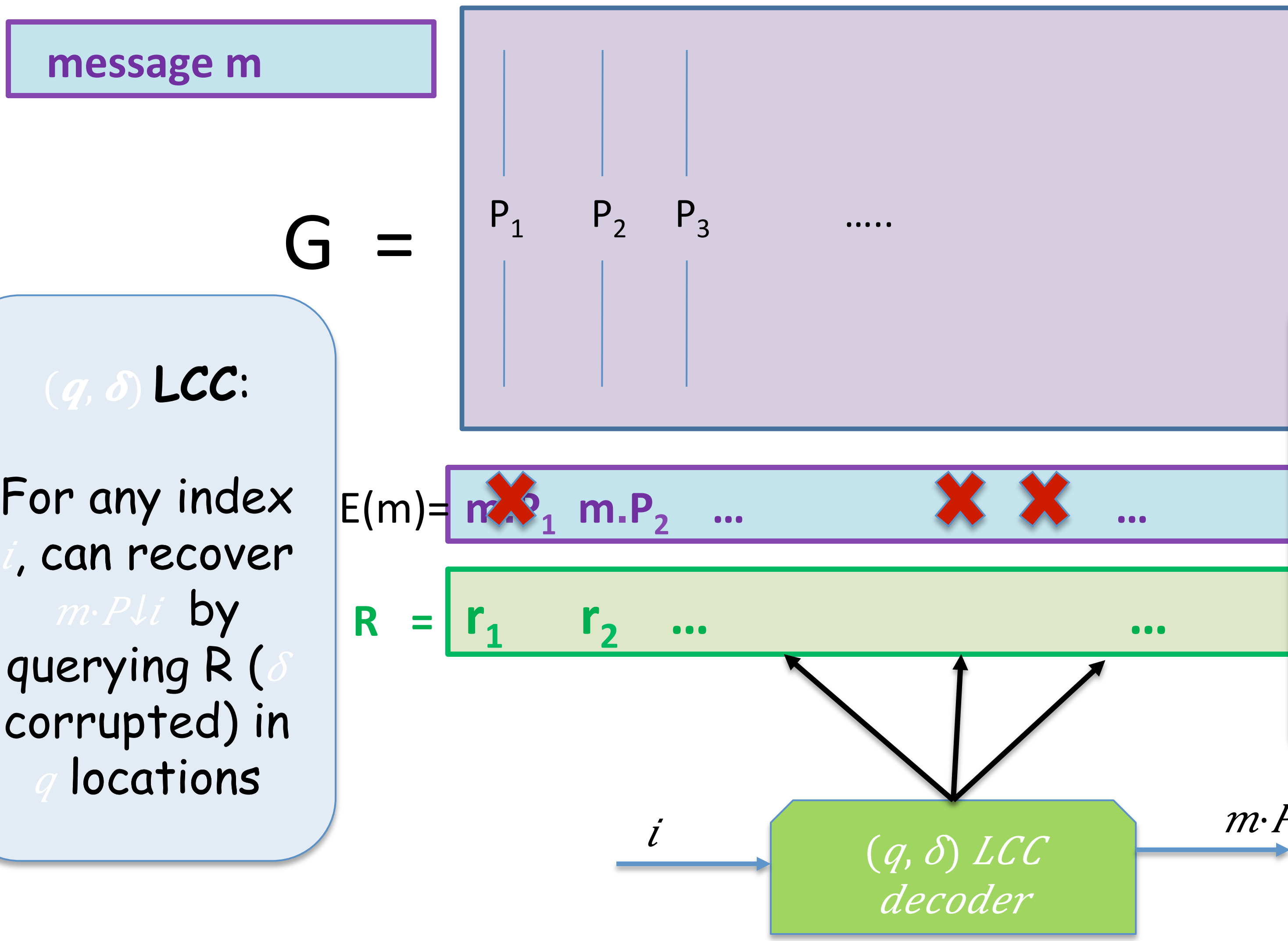
The Plan

- Extensions of the SG Theorem
  - Complexes
  - Quantitative versions
  - Stable versions
  - High dimensions

- Connection to Locally Correctable Codes

- New lower bounds for 3-query LCCs

(Linear) Locally Correctable Code

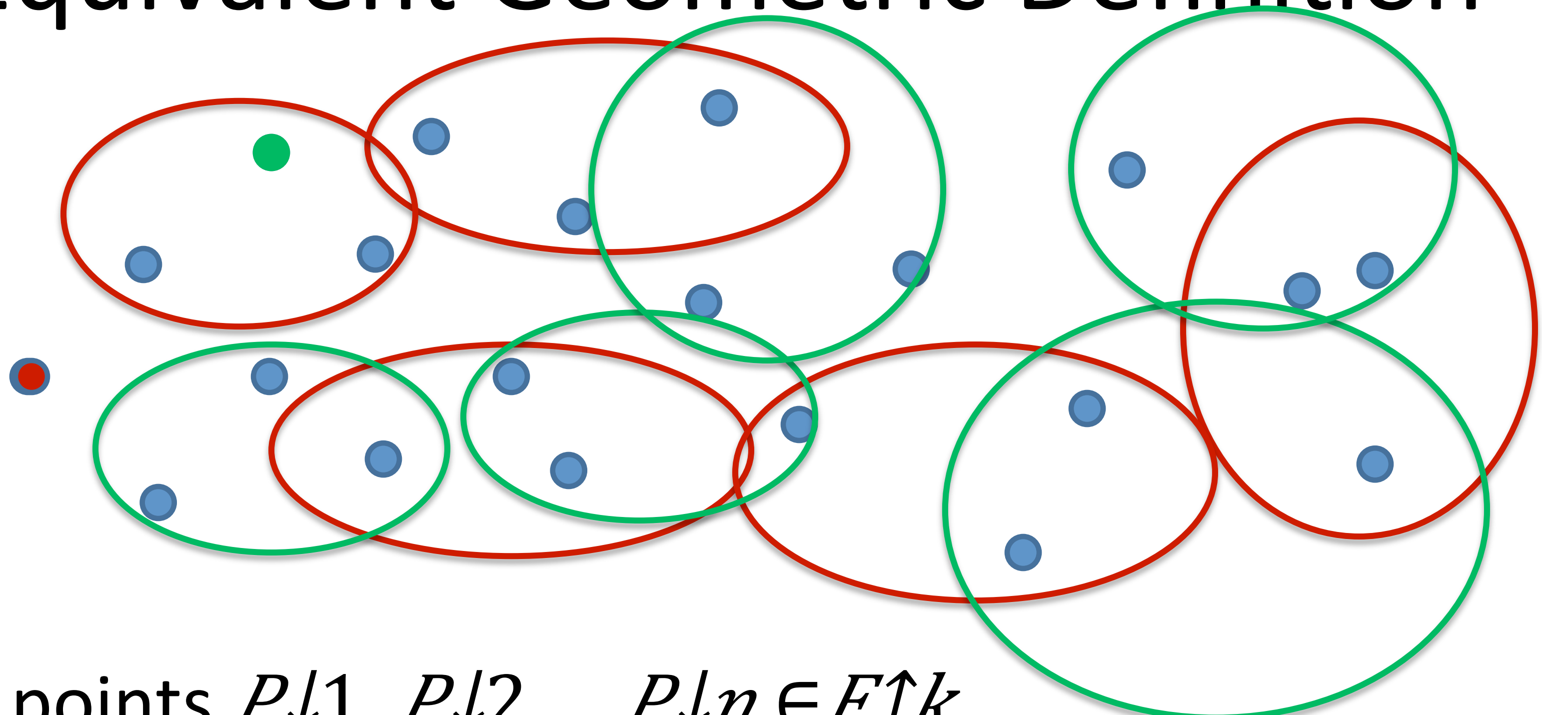


(q, δ) LCC:

For any index i, can recover $m \cdot P \downarrow i$ by querying R (δ corrupted) in q locations

If (j,k,l) locations used to decode i, $P \downarrow i$ is in linear span of $P \downarrow j$, $P \downarrow k$, $P \downarrow l$

Equivalent Geometric Definition



- n points $P_1, P_2, \ldots P_n \in F^k$
- For each P_i, there are δn disjoint q-tuples spanning P_i
- Thus n different *matchings* M_i of q-tuples

Locally correctable codes

- Central role in program testing, PCPs, IP = PSPACE…
- Only examples we know: Hadamard code, Reed Muller code
- Very weak lower bounds known
- [Dvir]: (even mild) lower bounds for polylog-query LCCs implies new lower bounds for *matrix rigidity*

2 Query LCCs

- Only example known: Hadamard Code

- Lower Bounds:
  - [GKST02]: $n=2^{\Omega(k)}$ (over F_2)
  - [BDSS11]: $n=p^{\Omega(k)}$ (Over F_p)

- [BDWY11]: over R they do not exist!

3 Query LCCs

- Best Upper bounds: Reed-Muller codes
  - F_2 : $n=2^{\sqrt{k}}$
  - Over $\mathbb{R}$: no examples

- Best Lower bounds [GKST02, Woo07, Woo10]:
  - Over any field: $n=\Omega(k^2)$

- New result [Dvir-*S*-Wigderson 13]:

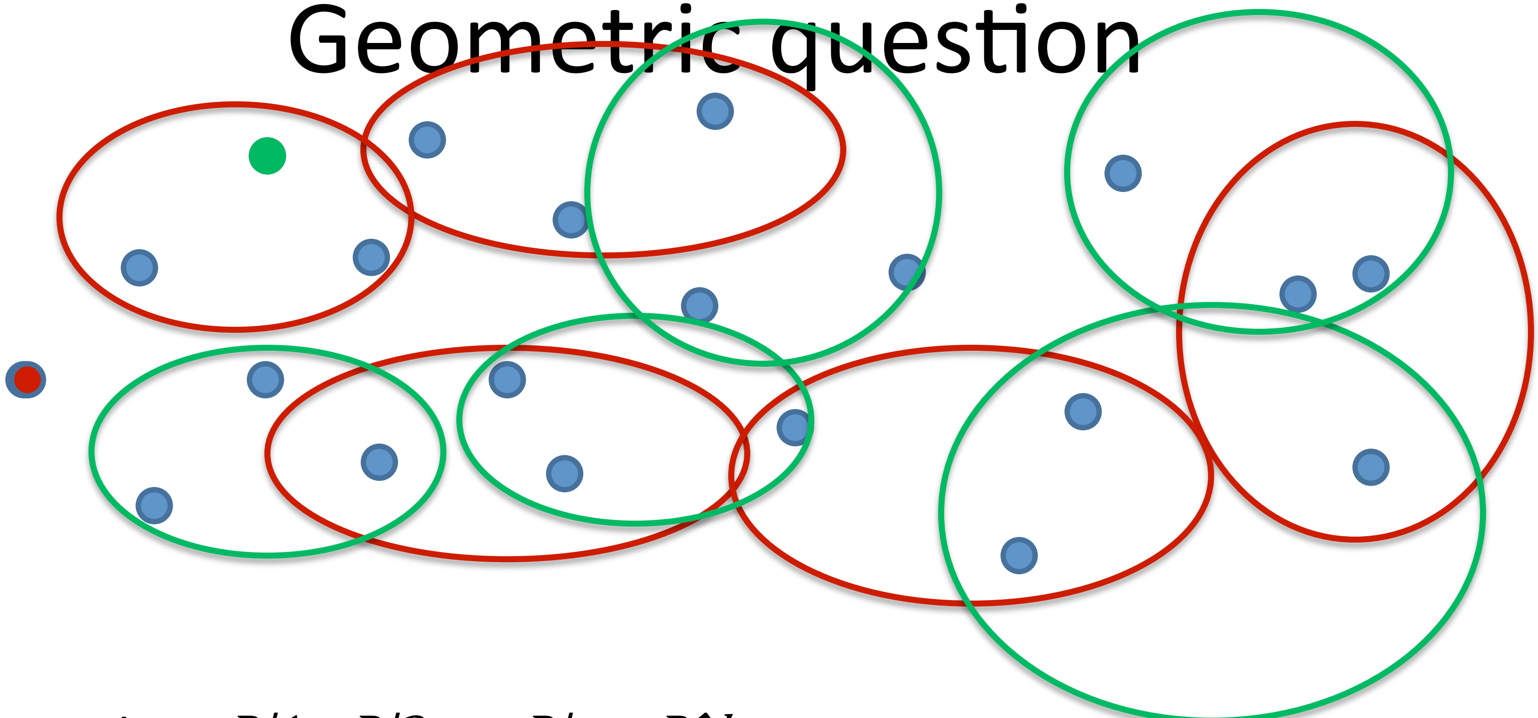
  over $\mathbb{R}$: $n=\Omega(k^{2+\epsilon})$

Rest of the talk

For 3 query LCCs over the real numbers,

$$n > k^{(2+\epsilon)}$$

3 query LCCs over R: Geometric question

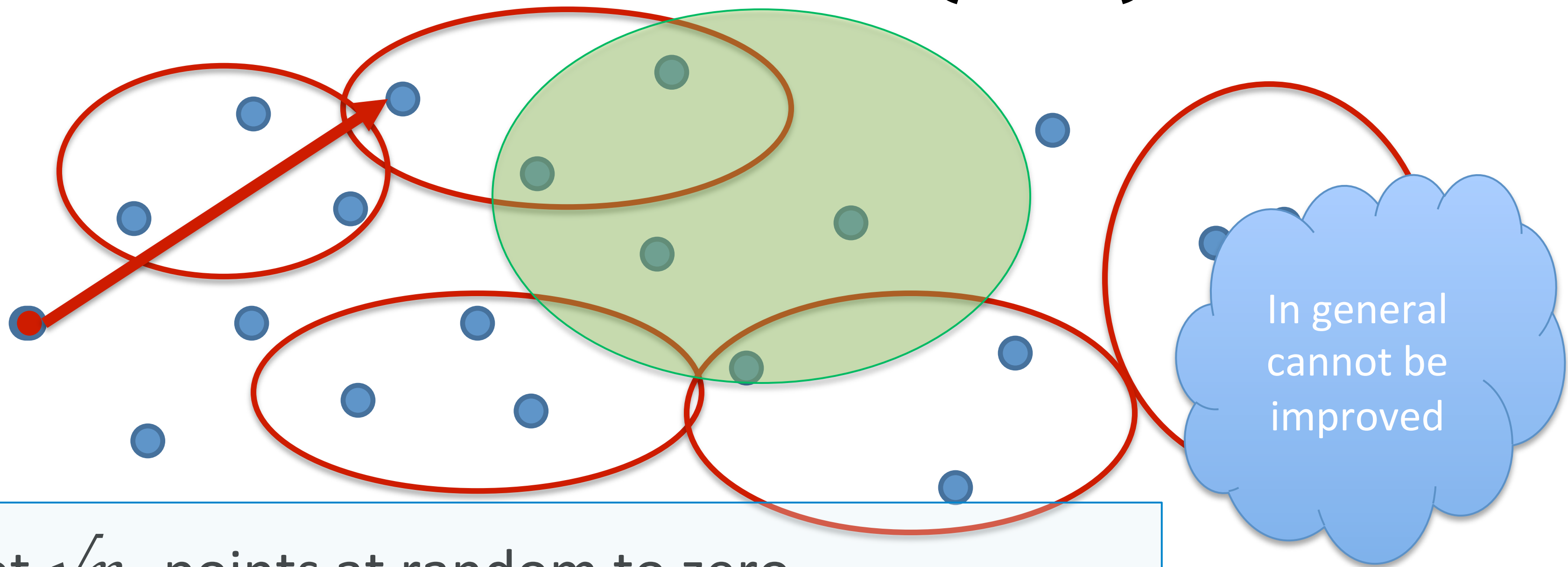


Given n points $P_1, P_2, \ldots, P_n \in R^k$

For each P_i, there is a ``matching'' M_i of triples of size δn spanning P_i

- Woodruff 10: $k < \sqrt{n}$
- DSW13: $k < n^{0.499}$
- Possible: $k < poly(1/\delta)$

Warmup : $k < O(\sqrt{n})$



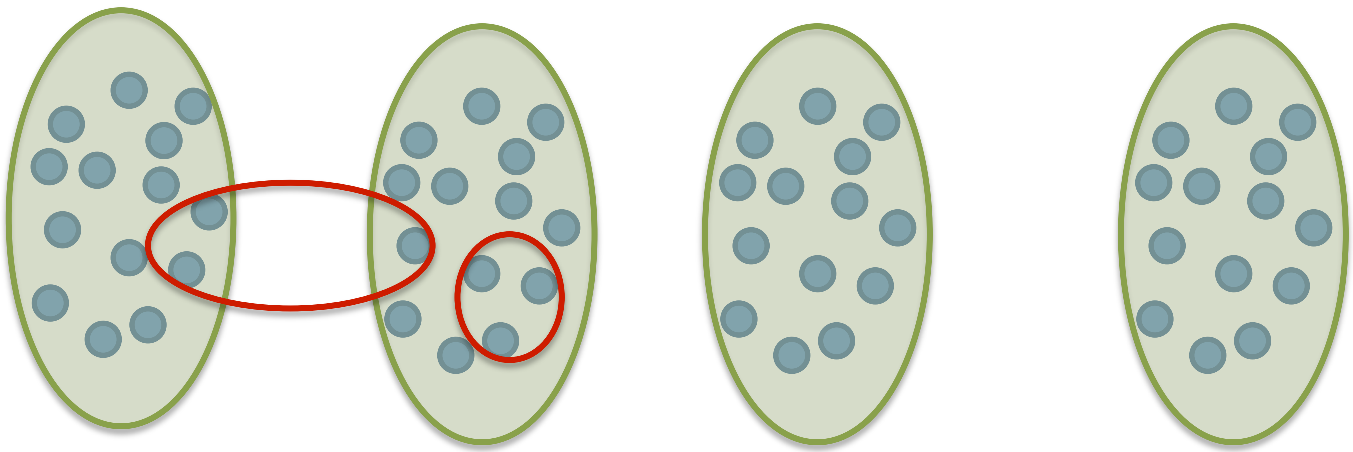
- Set $\sqrt{n}$ points at random to zero.
  - Dimension reduces by at most $\sqrt{n}$
  - For each P_i, some triple in M_i becomes a *singleton*
    - P_i gets identified with it
  - Set shrinks by a constant amount
- Repeat $log(n)$ times …

[Dvir-S-Wigderson 13]: $k < n^{0.499}$

If possible $k > n^{0.499}$

1) The triples in the LCC must be *structured*
2) Exploit structure to get *improved random restriction*

Structure theorem



If possible $k > n^{0.499}$

- There is a *clustering* of points:
  $\approx\sqrt{n}$ points in $\approx\sqrt{n}$ clusters
- Every triple intersects some cluster in 2 points

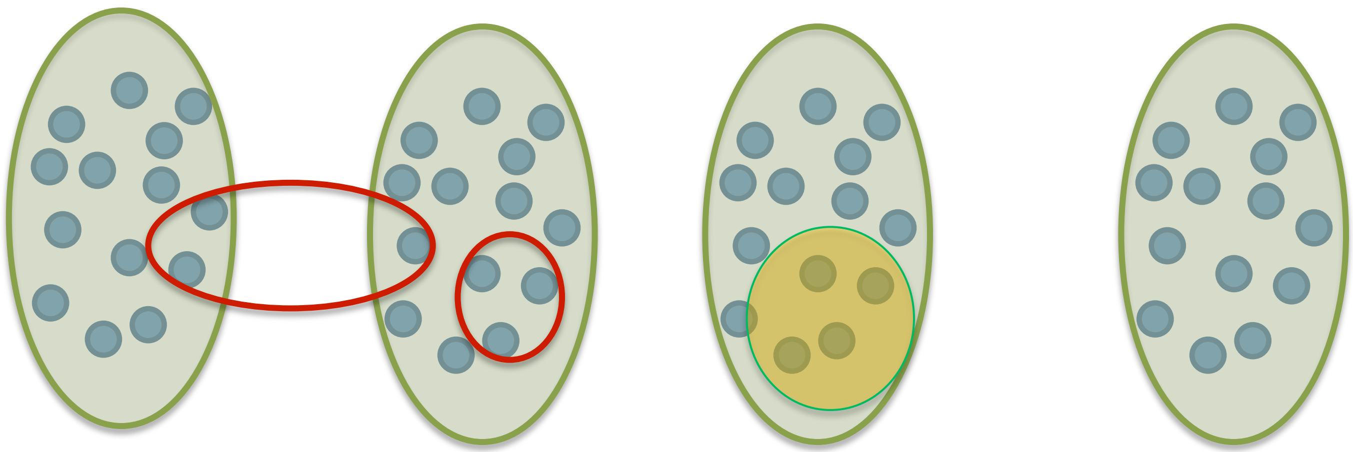
Structure theorem: Main idea

- Barthe`98
  - Given n points in R^k s.t. *no large subset in a low dimension*, then there exists an invertible linear transformation M st for $\boldsymbol{P_i' = MP_i/|MP_i|}$ the new points are ``well spread”

For every unit $\boldsymbol{w \in R^k}$ we have
$(*)\quad \sum_{i=1}^{n} |\boldsymbol{w \cdot P'_i}|^2 \le \boldsymbol{O(n/k)} < \boldsymbol{n^{0.501}}$

  - No point correlates with too many other points
  - But in a 3-LCC, many correlated pairs
    - Many dependent 4-tuples
    - For every dependent 4-tuple, there is some pair of points that has nontrivial correlation

Improved Random Restriction



- Totally $\approx n^2$ triples.
- $\approx \sqrt{n}$ clusters - thus typical cluster has $\approx \sqrt{n}$ edges *per matching*

- Pick random cluster and set random $n^{1/4}$ points in it to zero
  - Dimension reduces by at most $n^{1/4}$
  - For each P_i, some triple in M_i becomes a *singleton*
    - P_i gets identified with it
  - Size of set shrinks by a constant factor
- Repeat $log\ n$ times

Summary

- Several variations of the SG thm
  - Many local linear dependencies => global dimension bound

- Similar to Freiman-Ruzsa thm in additive combinatorics:
  $|A + A| < k\,|A|$ => *structure*
  - (lots of additive triples implies structure/low dims)
  - [BD*S*S11] optimal lower bounds for 2-query LCCS over F_p (BSG +Ruzsa)

- Very little understood about high dimensional versions
  - Extremely interesting for lower bounds for LCCs

Future Directions

- Lower bounds for 3 query LCCs over F_2 ?
  - Random restriction part still works
  - Clustering?
- Show that there are no 3 query LCCs over Reals
- Improved lower bounds for more queries?
  - Barthe, correlations etc still work
  - Strong enough lower bounds imply new lower bounds for matrix rigidity
- Improved bounds for stable SG?

Thanks!