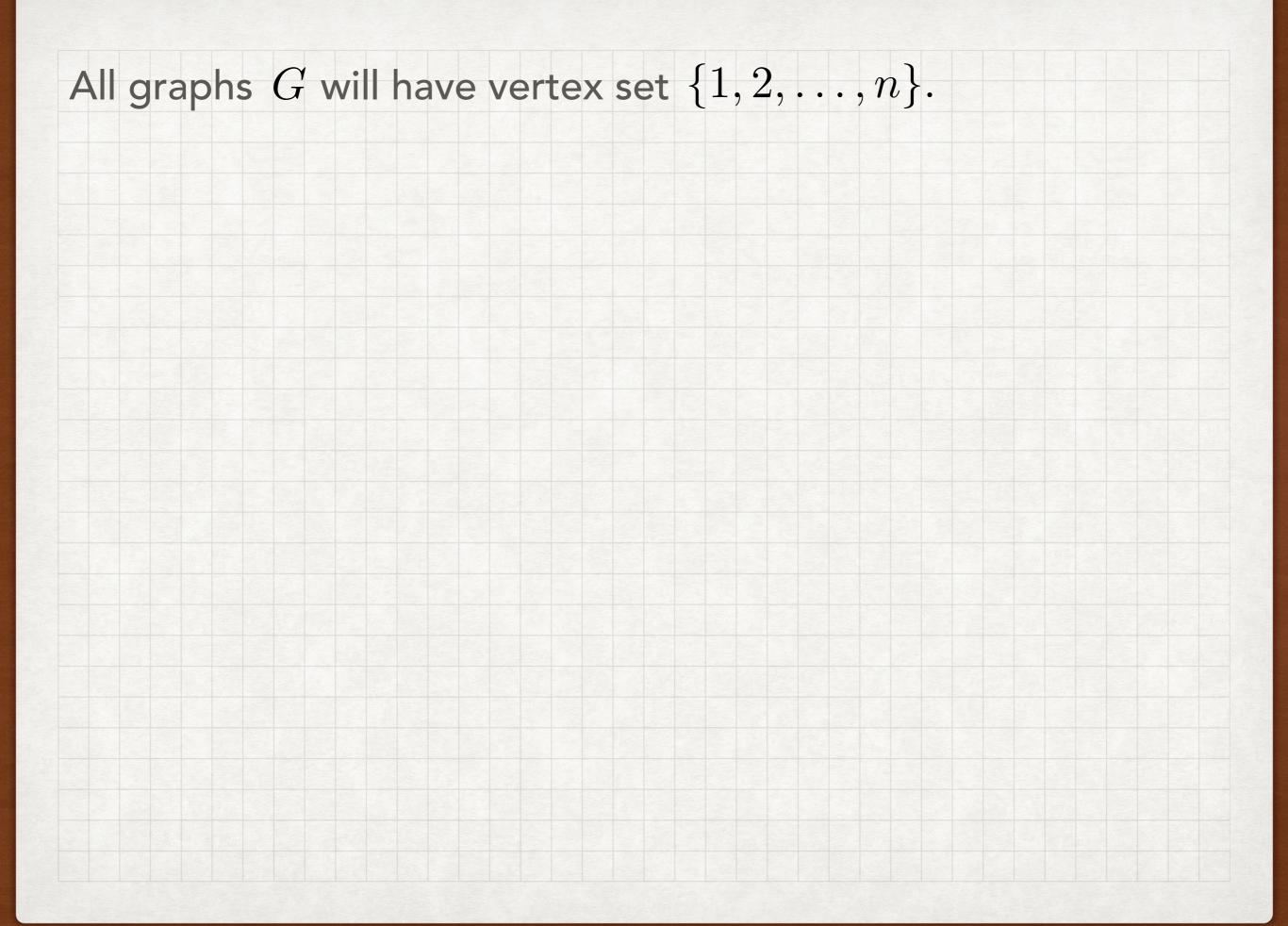
# THE MINRANK OF RANDOM GRAPHS OVER ARBITRARY FIELDS

By: Igor Balla

Joint work with: Noga Alon, Lior Gishboliner, Adva Mond,

Frank Mousset



**Definition:** An orthogonal representation of G in  $\mathbb{R}^k$  is an assignment of a nonzero vector  $v_i \in \mathbb{R}^k$  to each vertex i such that  $\langle v_i, v_j \rangle = 0$  when  $ij \notin E(G)$ .

**Definition:** An orthogonal representation of G in  $\mathbb{R}^k$  is an assignment of a nonzero vector  $v_i \in \mathbb{R}^k$  to each vertex i such that  $\langle v_i, v_j \rangle = 0$  when  $ij \notin E(G)$ .

**Question** (Knuth 94): What is the minimum k such that the random graph G(n, p) for fixed p has an orthogonal representation in  $\mathbb{R}^k$  with high probability?

**Definition:** An orthogonal representation of G in  $\mathbb{R}^k$  is an assignment of a nonzero vector  $v_i \in \mathbb{R}^k$  to each vertex i such that  $\langle v_i, v_j \rangle = 0$  when  $ij \notin E(G)$ .

**Question** (Knuth 94): What is the minimum k such that the random graph G(n, p) for fixed p has an orthogonal representation in  $\mathbb{R}^k$  with high probability?

Note that any graph G has an orthogonal representation in  $\mathbb{R}^{\chi(G)}$ .

**Definition:** An orthogonal representation of G in  $\mathbb{R}^k$  is an assignment of a nonzero vector  $v_i \in \mathbb{R}^k$  to each vertex i such that  $\langle v_i, v_j \rangle = 0$  when  $ij \notin E(G)$ .

**Question** (Knuth 94): What is the minimum k such that the random graph G(n, p) for fixed p has an orthogonal representation in  $\mathbb{R}^k$  with high probability?

Note that any graph G has an orthogonal representation in  $\mathbb{R}^{\chi(G)}$ .

**Theorem (**Grimmett and McDiarmid 75): For p fixed, the random graph  $G \sim G(n, p)$  has chromatic number  $\chi(G) = \Theta(n/\log n)$  with high probability.

**Definition:** The minrank mr(G) of a graph G over a field  $\mathbb{F}$  is the minimum rank rk(M) of a matrix  $M \in \mathbb{F}^{n \times n}$  that represents G.

**Definition:** The minrank mr(G) of a graph G over a field  $\mathbb{F}$  is the minimum rank rk(M) of a matrix  $M \in \mathbb{F}^{n \times n}$  that represents G.

If G has an orthogonal representation in  $\mathbb{R}^k$ , then the Gram matrix of all pairwise inner products represents G and has rank k.

**Definition:** The minrank mr(G) of a graph G over a field  $\mathbb{F}$  is the minimum rank rk(M) of a matrix  $M \in \mathbb{F}^{n \times n}$  that represents G.

If G has an orthogonal representation in  $\mathbb{R}^k$ , then the Gram matrix of all pairwise inner products represents G and has rank k.

**Theorem** (Alon, B., Gishboliner, Mond, Mousset): For any field  $\mathbb{F}$  and any  $1/n \le p \le 1$ , the random graph  $G \sim G(n, p)$  satisfies with high probability that

$$\operatorname{mr}(G) \ge \frac{n \log(1/p)}{80 \log n}.$$

**Definition:** The minrank mr(G) of a graph G over a field  $\mathbb{F}$  is the minimum rank rk(M) of a matrix  $M \in \mathbb{F}^{n \times n}$  that represents G.

If G has an orthogonal representation in  $\mathbb{R}^k$ , then the Gram matrix of all pairwise inner products represents G and has rank k.

**Theorem** (Alon, B., Gishboliner, Mond, Mousset): For any field  $\mathbb{F}$  and any  $1/n \le p \le 1$ , the random graph  $G \sim G(n, p)$  satisfies with high probability that

$$\operatorname{mr}(G) \ge \frac{n \log(1/p)}{80 \log n}.$$

For finite fields  $\mathbb{F}$  with  $|\mathbb{F}| \leq n^{O(1)}$ , this result was already proven recently by Golovnev, Regev, and Weinstein.

A Naive Proof				

For simplicity, lets assume that all matrices are symmetric, p=1/2, and the field  $\mathbb F$  is finite and of constant size.

For simplicity, lets assume that all matrices are symmetric, p=1/2, and the field  $\mathbb F$  is finite and of constant size.

**Definition:** Let  $s(M) = |\{(i, j) : M_{i,j} \neq 0\}|$  denote the sparsity of M.

For simplicity, lets assume that all matrices are symmetric, p=1/2, and the field  $\mathbb F$  is finite and of constant size.

**Definition:** Let  $s(M) = |\{(i, j) : M_{i,j} \neq 0\}|$  denote the sparsity of M.

By taking a union bound over all matrices of rank at most k, we have

 $P[mr(G) \le k] \le \sum_{M \in \mathbb{F}^{n \times n} : rk(M) \le k} P[M \text{ represents } G]$ 

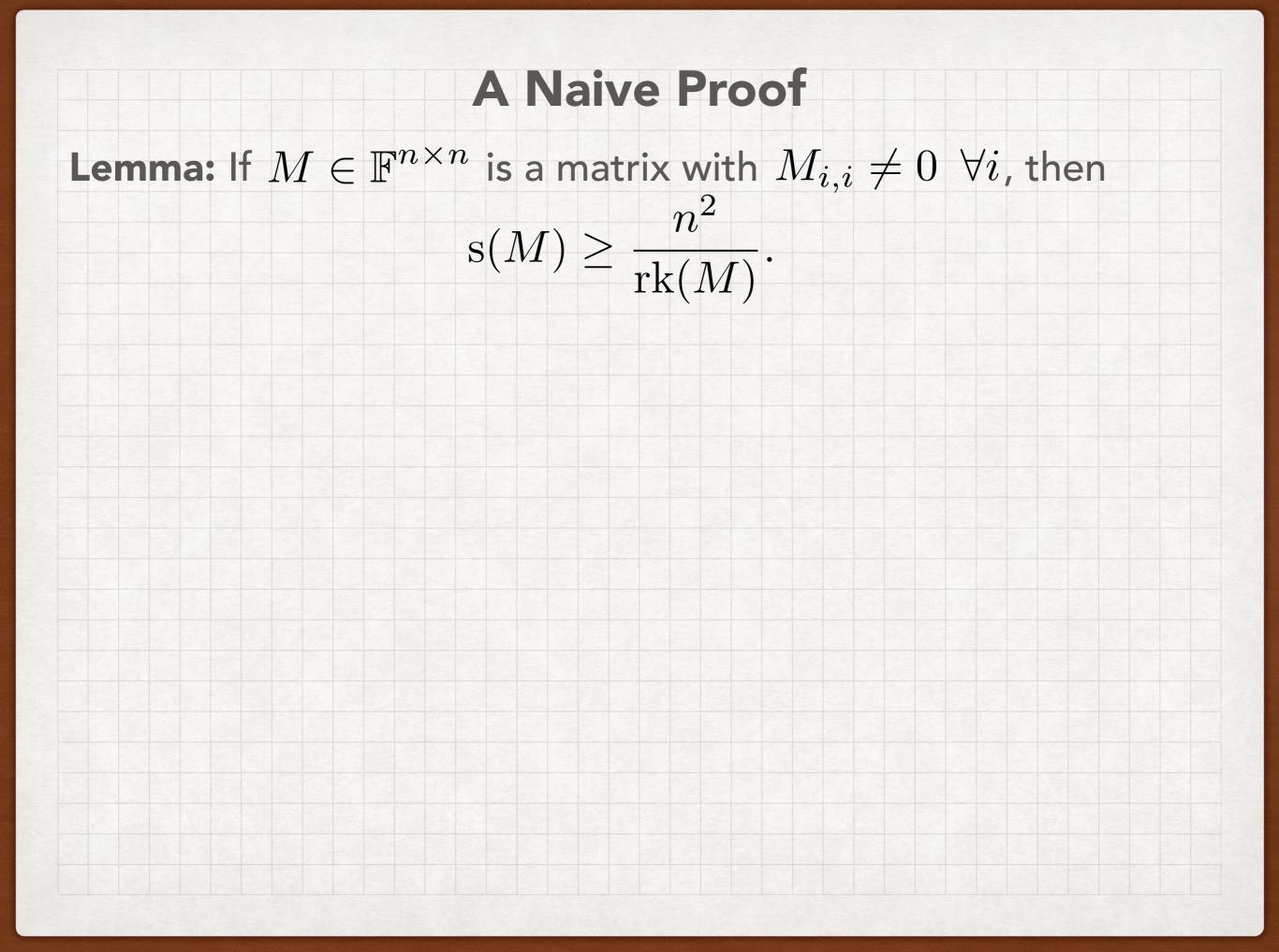
For simplicity, lets assume that all matrices are symmetric, p=1/2, and the field  $\mathbb F$  is finite and of constant size.

**Definition:** Let  $s(M) = |\{(i, j) : M_{i,j} \neq 0\}|$  denote the sparsity of M.

By taking a union bound over all matrices of rank at most k, we have

$$P[\operatorname{mr}(G) \le k] \le \sum_{\substack{M \in \mathbb{F}^{n \times n} : \operatorname{rk}(M) \le k}} P[M \text{ represents } G]$$
$$= \sum_{\substack{M \in \mathbb{F}^{n \times n} : \operatorname{rk}(M) \le k}} (1/2)^{(s(M)-n)/2}$$

A Naive Proof				



**Lemma:** If  $M \in \mathbb{F}^{n \times n}$  is a matrix with  $M_{i,i} \neq 0 \quad \forall i$ , then  $s(M) \geq \frac{n^2}{rk(M)}.$ 

**Proof:** Let G be the graph where ij is an edge if and only if  $M_{i,j} \neq 0$ . Note that it has e(G) = (s(M) - n)/2 edges.

**Lemma:** If  $M \in \mathbb{F}^{n \times n}$  is a matrix with  $M_{i,i} \neq 0 \quad \forall i$ , then  $\mathbf{s}(M) \geq \frac{n^2}{\mathrm{rk}(M)}.$ 

**Proof:** Let G be the graph where ij is an edge if and only if  $M_{i,j} \neq 0$ . Note that it has e(G) = (s(M) - n)/2 edges.

By Turán's theorem, G has an independent set of size

$$\frac{n}{1+2e(G)/n} = \frac{n^2}{\mathrm{s}(M)}.$$

**Lemma:** If  $M \in \mathbb{F}^{n \times n}$  is a matrix with  $M_{i,i} \neq 0 \quad \forall i$ , then  $\mathbf{s}(M) \geq \frac{n^2}{\mathrm{rk}(M)}.$ 

**Proof:** Let G be the graph where ij is an edge if and only if  $M_{i,j} \neq 0$ . Note that it has e(G) = (s(M) - n)/2 edges.

By Turán's theorem, G has an independent set of size  $\frac{n}{1+2e(G)/n} = \frac{n^2}{s(M)}.$ 

On the other hand, any independent set in G corresponds to a full rank submatrix of M, and so must have size at most rk(M).

A Naive Proof				

Note that any matrix  $M \in \mathbb{F}^{n \times n}$  with  $\operatorname{rk}(M) \leq k$  can be written as M = UV, where  $U \in \mathbb{F}^{n \times k}$  and  $V \in \mathbb{F}^{k \times n}$ .

Note that any matrix  $M \in \mathbb{F}^{n \times n}$  with  $\operatorname{rk}(M) \leq k$  can be written as M = UV, where  $U \in \mathbb{F}^{n \times k}$  and  $V \in \mathbb{F}^{k \times n}$ .

Thus the number of matrices with rank  $\leq k$  is at most  $|\mathbb{F}|^{2nk}$ .

Note that any matrix  $M \in \mathbb{F}^{n \times n}$  with  $rk(M) \leq k$  can be written as M = UV, where  $U \in \mathbb{F}^{n \times k}$  and  $V \in \mathbb{F}^{k \times n}$ .

Thus the number of matrices with rank  $\leq k$  is at most  $|\mathbb{F}|^{2nk}$ .

Using the previous lemma, we conclude

 $P[mr(G) \le k] \le \qquad \sum (1/2)^{(s(M)-n)/2}$ 

 $M \in \mathbb{F}^{n \times n} : \operatorname{rk}(M) \leq k$ 

 $\leq |\mathbb{F}|^{2nk} (1/2)^{(n^2/k-n)/2}$ 

Note that any matrix  $M \in \mathbb{F}^{n \times n}$  with  $rk(M) \leq k$  can be written as M = UV, where  $U \in \mathbb{F}^{n \times k}$  and  $V \in \mathbb{F}^{k \times n}$ .

Thus the number of matrices with rank  $\leq k$  is at most  $|\mathbb{F}|^{2nk}$ .

Using the previous lemma, we conclude

$$P[mr(G) \le k] \le \qquad \sum (1/2)^{(s(M)-n)/2}$$

 $M \in \mathbb{F}^{n \times n} : \operatorname{rk}(M) \leq k$ 

$$\leq |\mathbb{F}|^{2nk} (1/2)^{(n^2/k-n)/2}$$

$$= o(1)$$

if we take  $k = \Theta(\sqrt{n})$ .

A Better Proof				

**Observation:** Any rank k matrix is uniquely determined by specifying k linearly independent rows, k linearly independent columns, and the indices of those rows and columns.

**Observation:** Any rank k matrix is uniquely determined by specifying k linearly independent rows, k linearly independent columns, and the indices of those rows and columns.

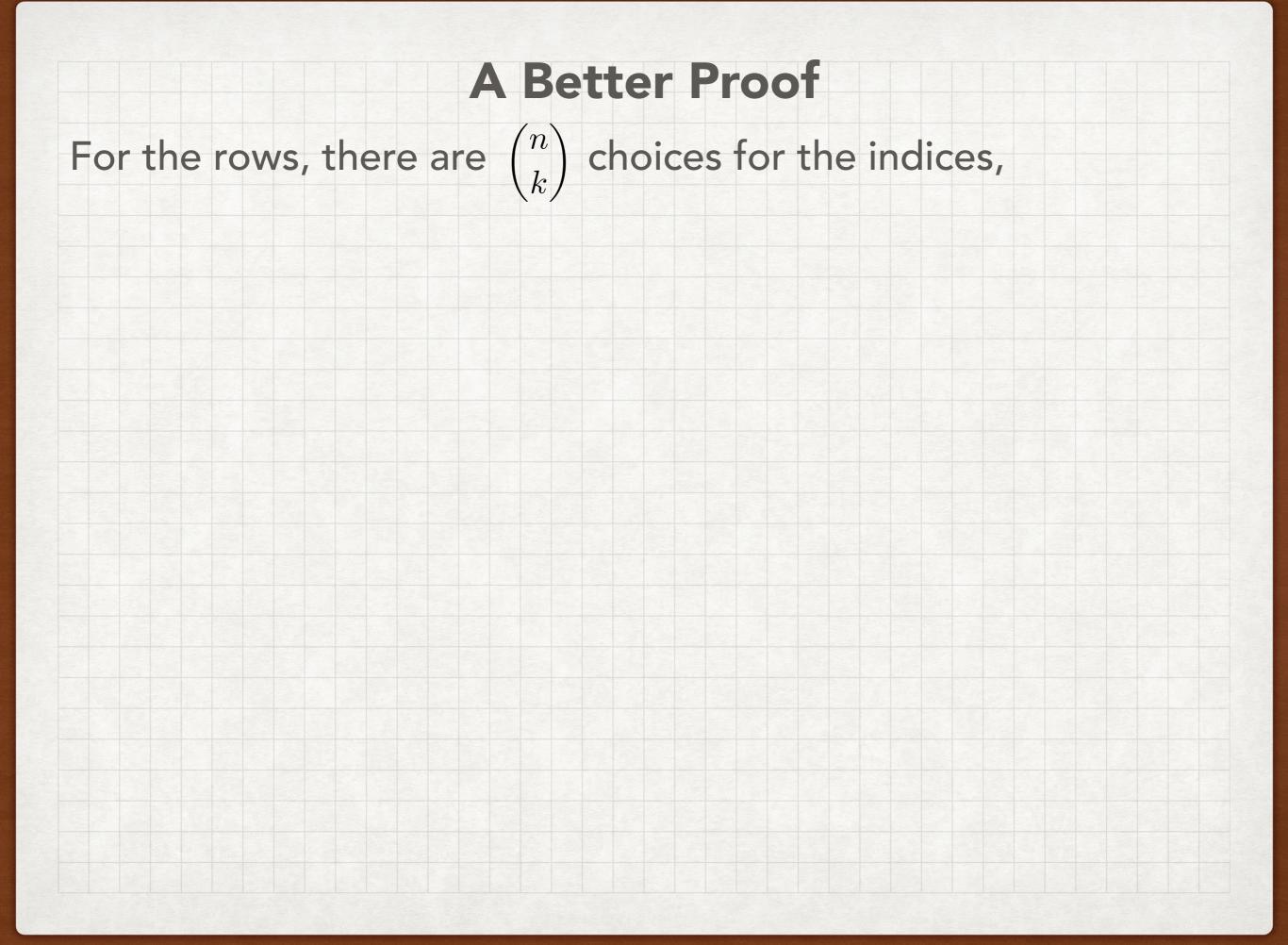
If  $M \in \mathbb{F}^{n \times n}$  is a matrix of rank k with sparsity s which is sufficiently "nice", then each row and column will have  $\approx s/n$ nonzero entries and so the k linearly independent rows/ columns determining M will each have  $\approx ks/n$  nonzero entries.

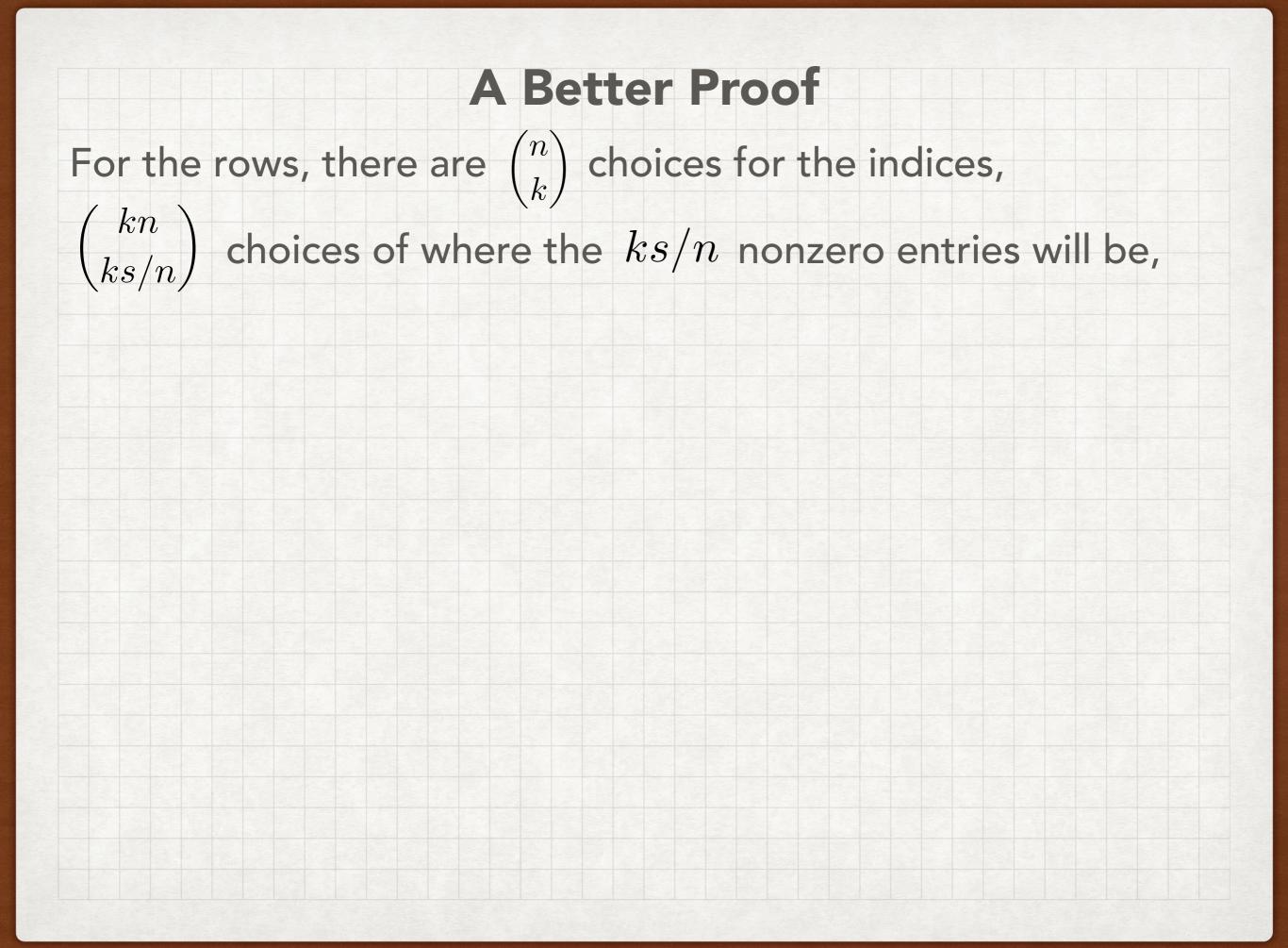
**Observation:** Any rank k matrix is uniquely determined by specifying k linearly independent rows, k linearly independent columns, and the indices of those rows and columns.

If  $M \in \mathbb{F}^{n \times n}$  is a matrix of rank k with sparsity s which is sufficiently "nice", then each row and column will have  $\approx s/n$ nonzero entries and so the k linearly independent rows/ columns determining M will each have  $\approx ks/n$  nonzero entries.

Lets bound the number of "nice" matrices with rank at most k and sparsity s.

A Better Proof				





For the rows, there are  $\binom{n}{k}$  choices for the indices,  $\binom{kn}{ks/n}$  choices of where the ks/n nonzero entries will be, and  $|\mathbb{F}|^{ks/n}$  choices for what those nonzero entries will be.

For the rows, there are  $\binom{n}{k}$  choices for the indices,  $\binom{kn}{ks/n}$  choices of where the ks/n nonzero entries will be, and  $|\mathbb{F}|^{ks/n}$  choices for what those nonzero entries will be.

The same holds for the columns, and so we conclude that the number of "nice" matrices with rank at most k and sparsity sis at most

For the rows, there are  $\binom{n}{k}$  choices for the indices,  $\binom{kn}{ks/n}$  choices of where the ks/n nonzero entries will be,

and  $|\mathbb{F}|^{ks/n}$  choices for what those nonzero entries will be.

The same holds for the columns, and so we conclude that the number of "nice" matrices with rank at most k and sparsity sis at most

 $\left(\binom{n}{k}\binom{kn}{ks/n}|\mathbb{F}|^{ks/n}\right)^2 \leq \left(n^k(kn)^{ks/n}|\mathbb{F}|^{ks/n}\right)^2$ 

## **A Better Proof**

For the rows, there are  $\binom{n}{k}$  choices for the indices,  $\binom{kn}{ks/n}$  choices of where the ks/n nonzero entries will be,

and  $|\mathbb{F}|^{ks/n}$  choices for what those nonzero entries will be.

The same holds for the columns, and so we conclude that the number of "nice" matrices with rank at most k and sparsity sis at most

 $\left( \binom{n}{k} \binom{kn}{ks/n} |\mathbb{F}|^{ks/n} \right)^2 \leq \left( n^k (kn)^{ks/n} |\mathbb{F}|^{ks/n} \right)^2 \\ \leq \left( n^{ks/n} (n^2)^{ks/n} |\mathbb{F}|^{ks/n} \right)^2$ 

## **A Better Proof**

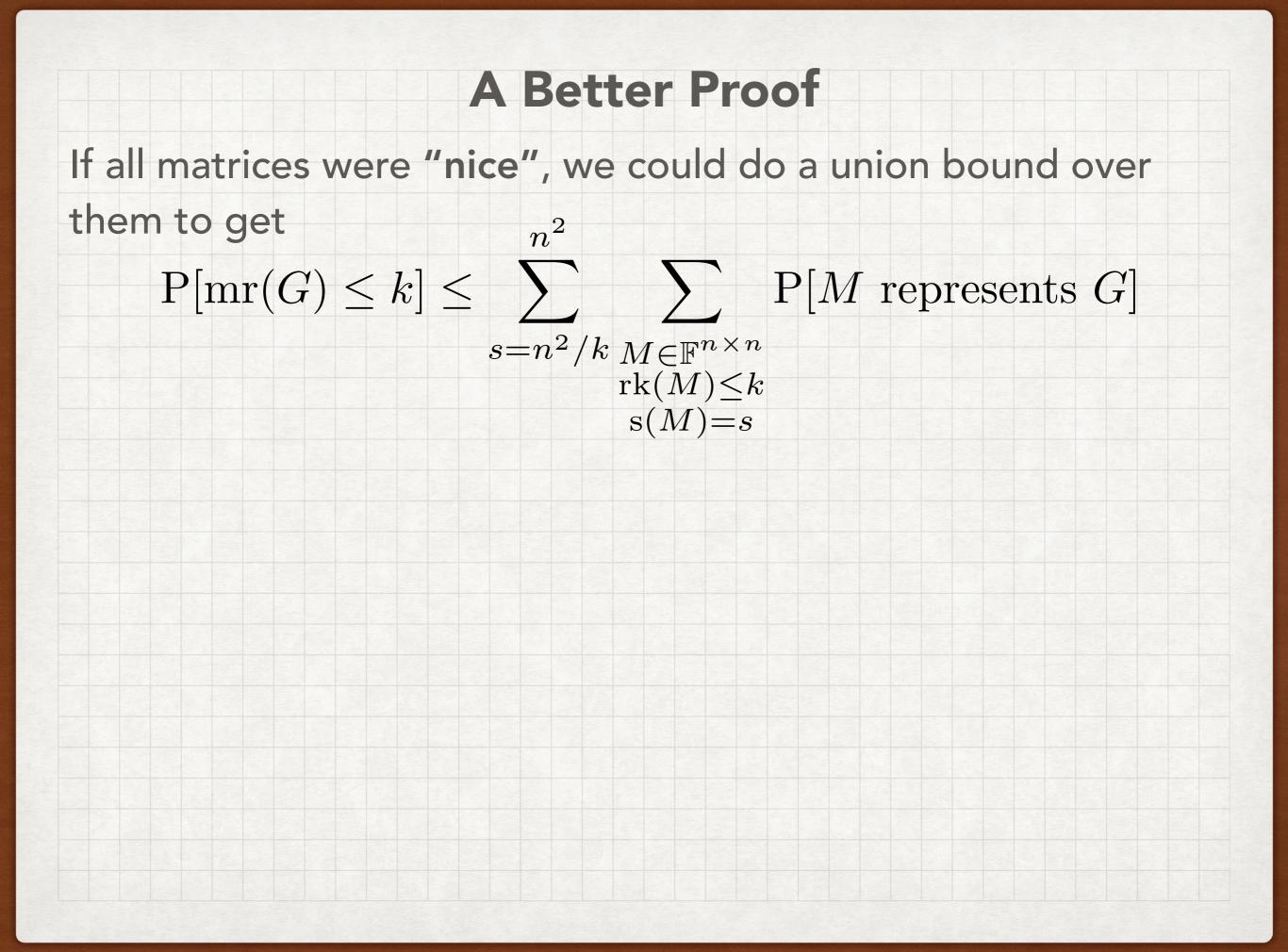
For the rows, there are  $\binom{n}{k}$  choices for the indices,  $\binom{kn}{ks/n}$  choices of where the ks/n nonzero entries will be,

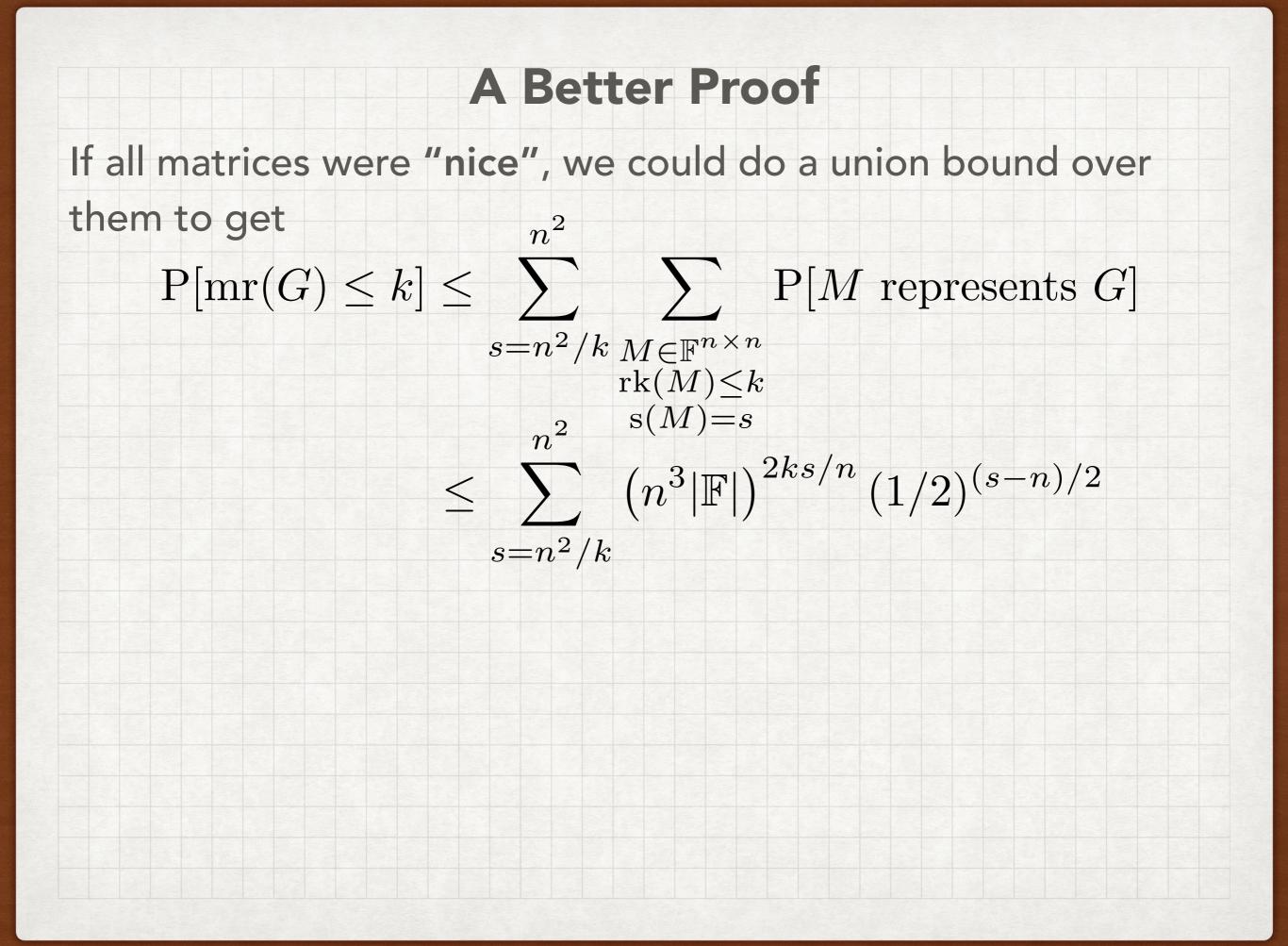
and  $|\mathbb{F}|^{ks/n}$  choices for what those nonzero entries will be.

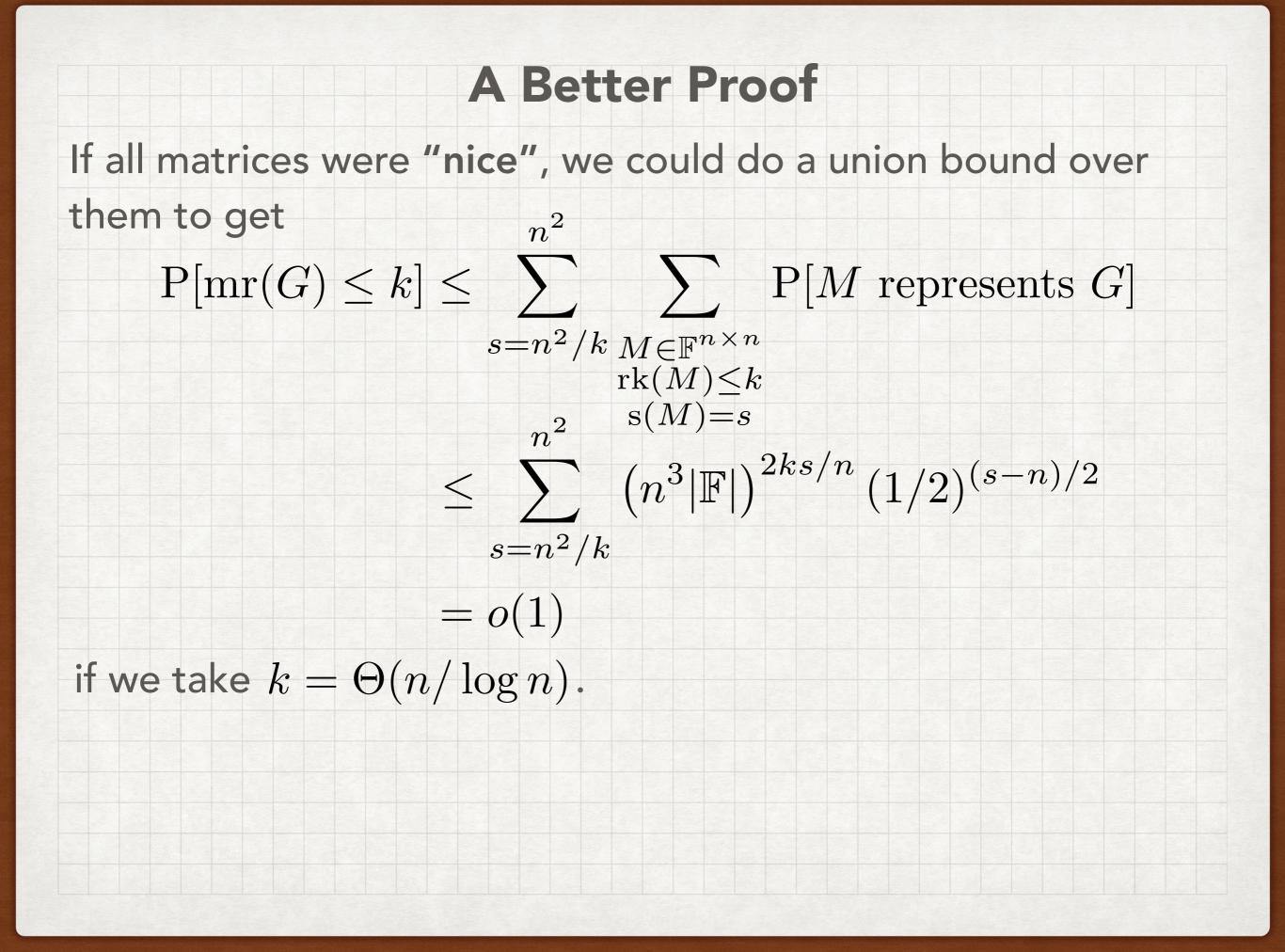
The same holds for the columns, and so we conclude that the number of "nice" matrices with rank at most k and sparsity sis at most

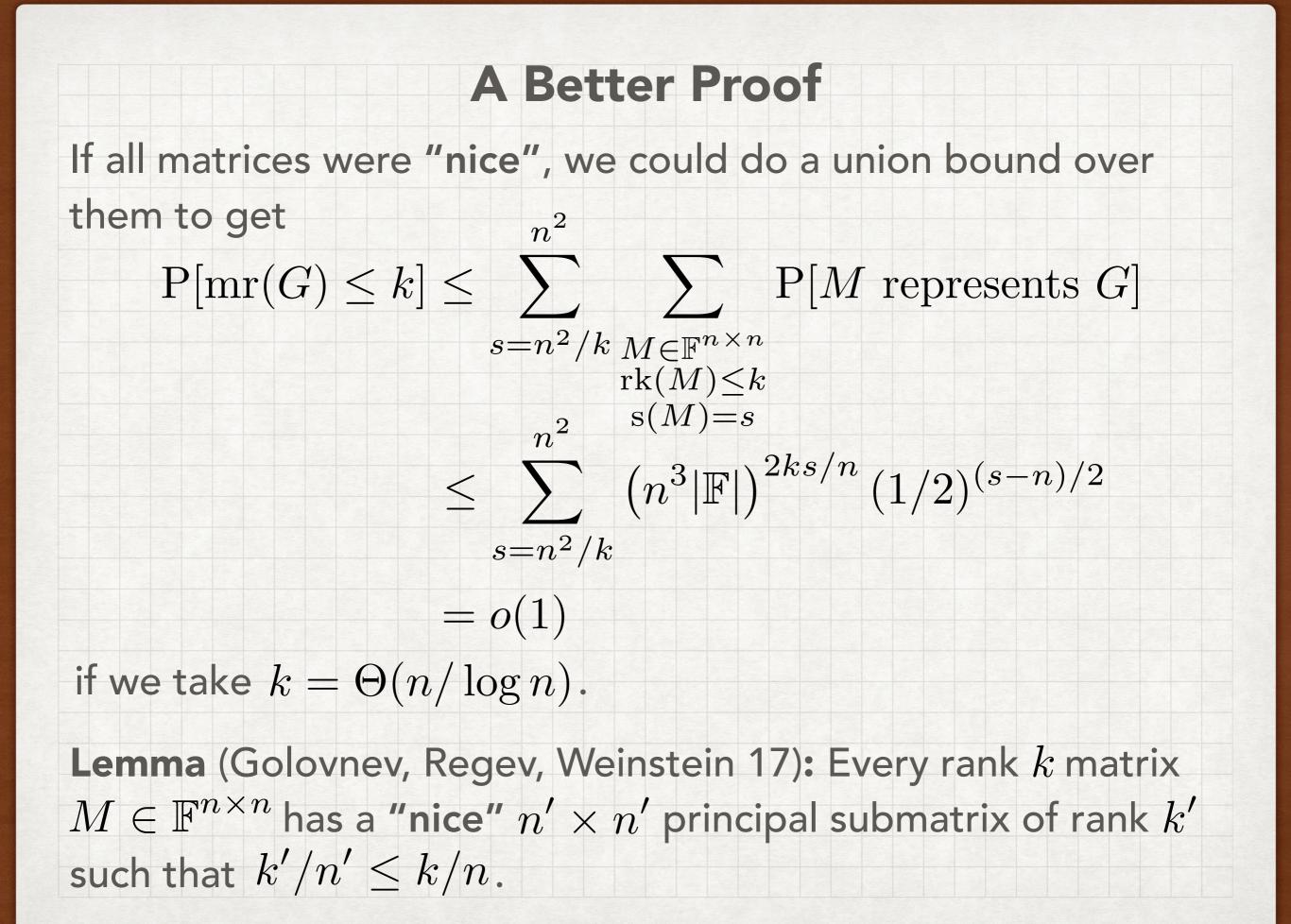
 $\left( \binom{n}{k} \binom{kn}{ks/n} |\mathbb{F}|^{ks/n} \right)^2 \leq \left( n^k (kn)^{ks/n} |\mathbb{F}|^{ks/n} \right)^2 \\ \leq \left( n^{ks/n} (n^2)^{ks/n} |\mathbb{F}|^{ks/n} \right)^2$  $= \left(n^3 |\mathbb{F}|\right)^{2ks/n}.$ 

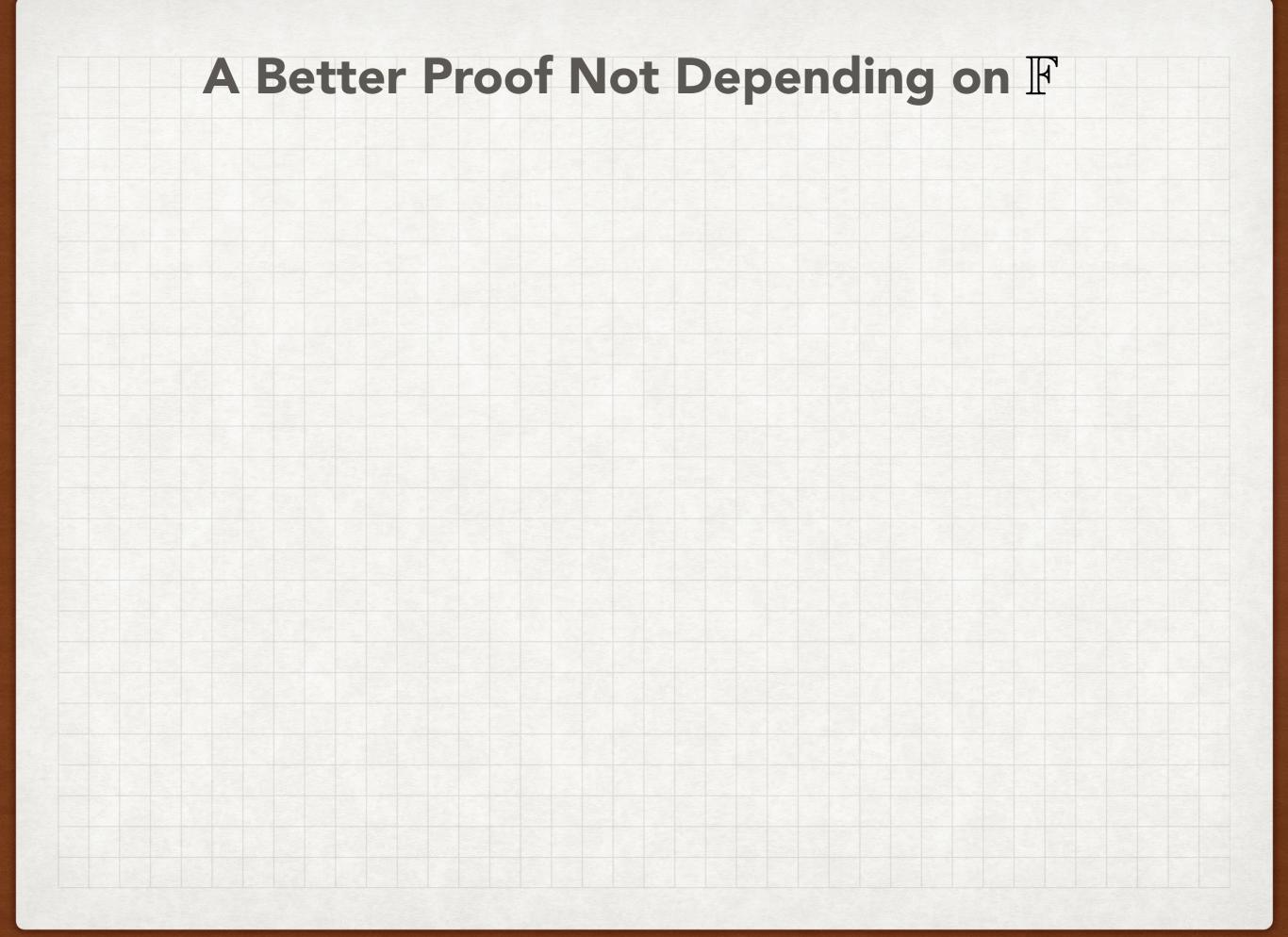
A Better Proof	











**Definition:** The zero-pattern of a vector or a matrix is obtained by replacing all nonzero entries with a  $\star$ .

**Definition:** The zero-pattern of a vector or a matrix is obtained by replacing all nonzero entries with a  $\star$ .

**Observation:** The probability that M represents G only depends on the zero-pattern of M.

**Definition:** The zero-pattern of a vector or a matrix is obtained by replacing all nonzero entries with a  $\star$ .

**Observation:** The probability that M represents G only depends on the zero-pattern of M.

Thus if we let  $Q_s$  denote the set of all zero-patterns of matrices  $M \in \mathbb{F}^{n \times n}$  with  $rk(M) \leq k$  and s(M) = s, then

**Definition:** The zero-pattern of a vector or a matrix is obtained by replacing all nonzero entries with a  $\star$ .

**Observation:** The probability that M represents G only depends on the zero-pattern of M.

 $n^2$ 

Thus if we let  $Q_s$  denote the set of all zero-patterns of matrices  $M \in \mathbb{F}^{n \times n}$  with  $rk(M) \leq k$  and s(M) = s, then

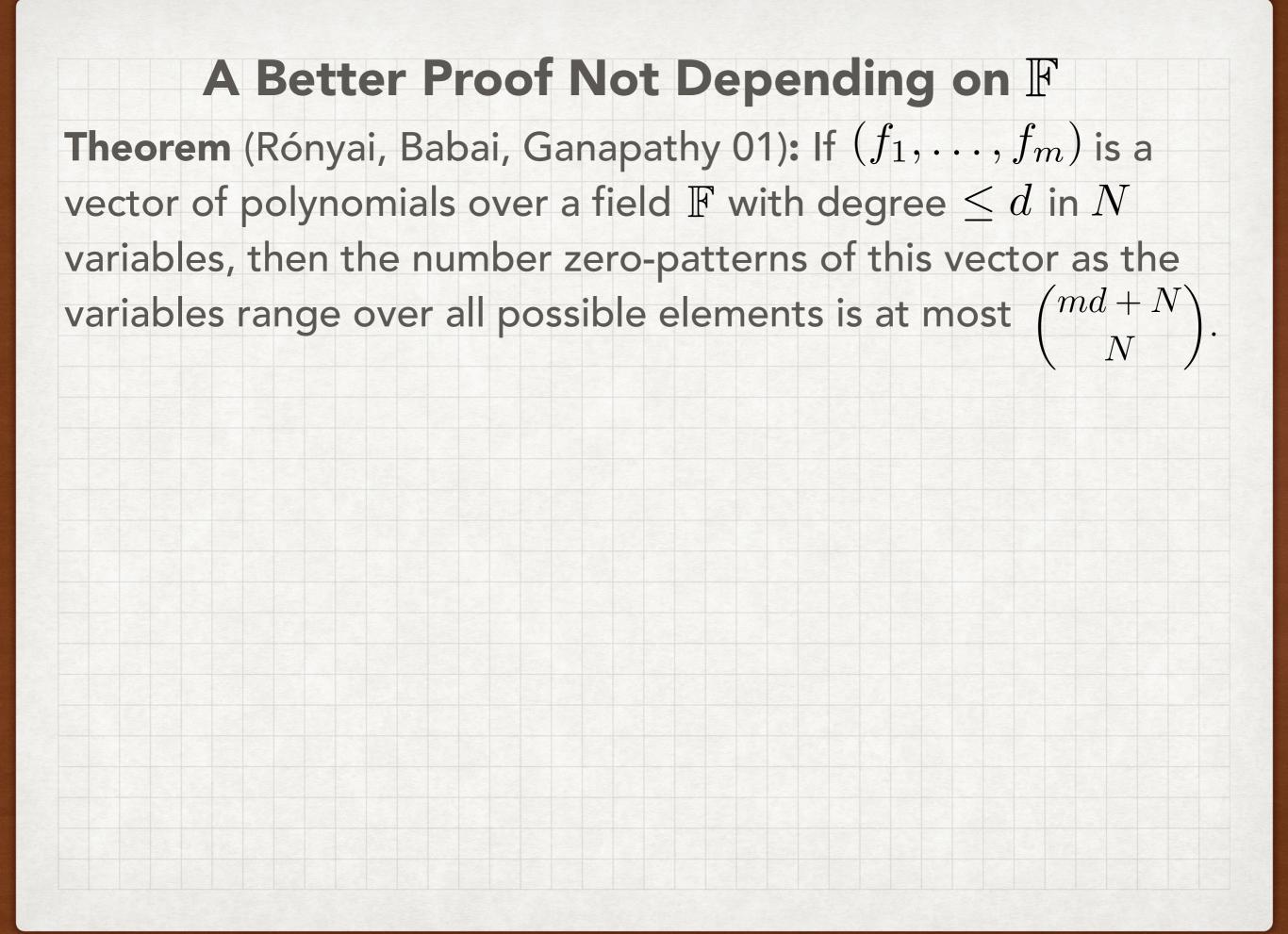
 $\Pr[\operatorname{mr}(G) \le k] \le \sum_{s=n^2/k} \sum_{Z \in Q_s} \Pr\left[\exists M : \overset{Z \text{ is the zero-pattern of } M}{M \text{ represents } G}\right]$ 

**Definition:** The zero-pattern of a vector or a matrix is obtained by replacing all nonzero entries with a  $\star$ .

**Observation:** The probability that M represents G only depends on the zero-pattern of M.

Thus if we let  $Q_s$  denote the set of all zero-patterns of matrices  $M \in \mathbb{F}^{n \times n}$  with  $rk(M) \leq k$  and s(M) = s, then

 $P[\operatorname{mr}(G) \le k] \le \sum_{s=n^2/k}^{n^2} \sum_{Z \in Q_s} P\left[\exists M : Z \text{ is the zero-pattern of } M\right]$  $\le \sum_{s=n^2/k}^{n^2} \sum_{Z \in Q_s} (1/2)^{(s-n)/2}$ 



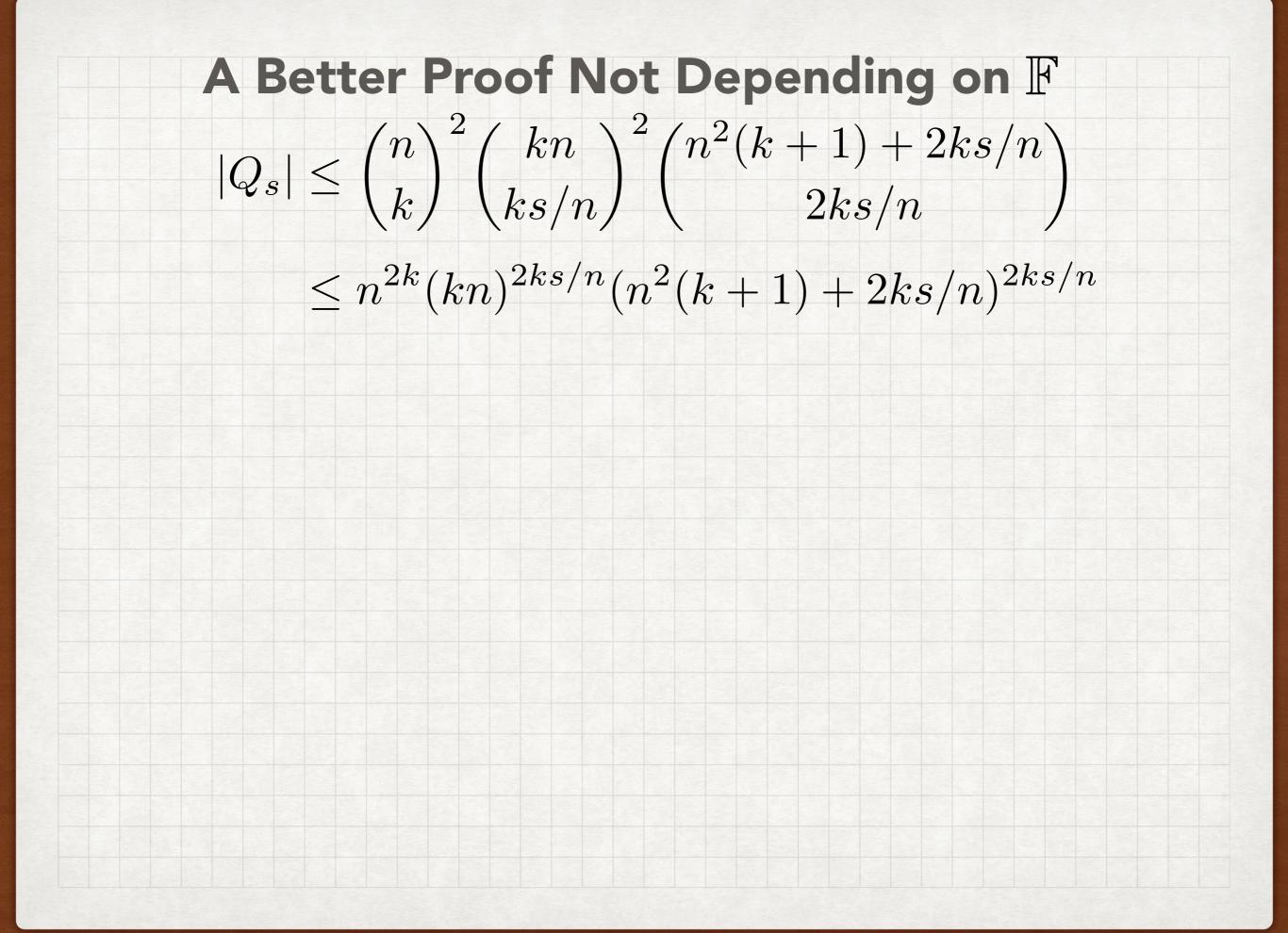
A Better Proof Not Depending on  $\mathbb F$ Theorem (Rónyai, Babai, Ganapathy 01): If  $(f_1, \ldots, f_m)$  is a vector of polynomials over a field  $\mathbb F$  with degree  $\leq d$  in Nvariables, then the number zero-patterns of this vector as the variables range over all possible elements is at most  $\binom{md+N}{N}$ . Let M be a "nice"  $n \times n$  matrix of rank k. Then there are k linearly independent rows and columns that uniquely determine it. These rows and columns have pprox 2ks/n nonzero entries. Call these entries E.

A Better Proof Not Depending on  $\mathbb F$ Theorem (Rónyai, Babai, Ganapathy 01): If  $(f_1, \ldots, f_m)$  is a vector of polynomials over a field  $\mathbb F$  with degree  $\leq d$  in Nvariables, then the number zero-patterns of this vector as the variables range over all possible elements is at most  $\binom{md+N}{N}$ . Let M be a "nice"  $n \times n$  matrix of rank k. Then there are k linearly independent rows and columns that uniquely determine it. These rows and columns have pprox 2ks/n nonzero entries. Call these entries E.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

A Better Proof Not Depending on  $\mathbb F$ Theorem (Rónyai, Babai, Ganapathy 01): If  $(f_1, \ldots, f_m)$  is a vector of polynomials over a field  $\mathbb{F}$  with degree  $\leq d$  in N variables, then the number zero-patterns of this vector as the variables range over all possible elements is at most  $\binom{md+N}{N}$ . Let M be a "nice"  $n \times n$  matrix of rank k. Then there are klinearly independent rows and columns that uniquely determine it. These rows and columns have pprox 2ks/n nonzero entries. Call these entries E. **Claim:** Each entry of M is a polynomial of degree  $\leq k+1$  in the variables of E. Thus we have  $n^2$  polynomials of degree  $\leq k+1$  in 2ks/n $|Q_s| \le \binom{n}{k}^2 \binom{kn}{ks/n}^2 \binom{n^2(k+1) + 2ks/n}{2ks/n}.$ variables, so

# **A Better Proof Not Depending on** $\mathbb{F}$ $|Q_s| \leq {\binom{n}{k}}^2 {\binom{kn}{ks/n}}^2 {\binom{n^2(k+1)+2ks/n}{2ks/n}}$

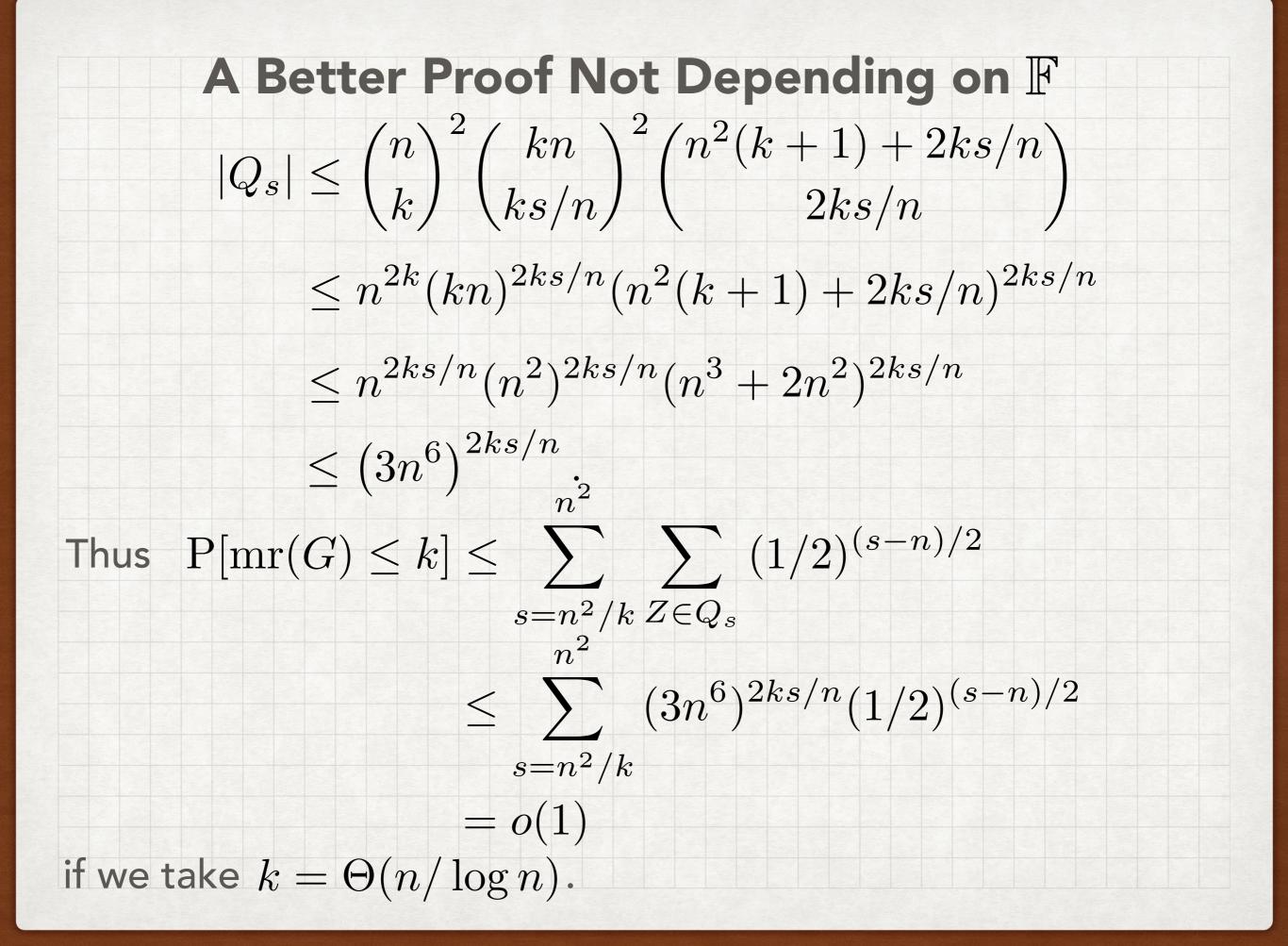


A Better Proof Not Depending on  $\mathbb F$  $|Q_s| \le \binom{n}{k}^2 \binom{kn}{ks/n}^2 \binom{n^2(k+1) + 2ks/n}{2ks/n}$  $\leq n^{2k} (kn)^{2ks/n} (n^2(k+1) + 2ks/n)^{2ks/n}$  $\leq n^{2ks/n} (n^2)^{2ks/n} (n^3 + 2n^2)^{2ks/n}$ 

A Better Proof Not Depending on  $\mathbb F$  $|Q_s| \le \binom{n}{k}^2 \binom{kn}{ks/n}^2 \binom{n^2(k+1) + 2ks/n}{2ks/n}$  $\leq n^{2k} (kn)^{2ks/n} (n^2(k+1) + 2ks/n)^{2ks/n}$  $\leq n^{2ks/n} (n^2)^{2ks/n} (n^3 + 2n^2)^{2ks/n}$  $\leq (3n^6)^{2ks/n}$ .

A Better Proof Not Depending on  $\mathbb F$  $|Q_s| \le \binom{n}{k}^2 \binom{kn}{ks/n}^2 \binom{n^2(k+1) + 2ks/n}{2ks/n}$  $\leq n^{2k} (kn)^{2ks/n} (n^2(k+1) + 2ks/n)^{2ks/n}$  $\leq n^{2ks/n} (n^2)^{2ks/n} (n^3 + 2n^2)^{2ks/n}$  $\leq \left(3n^6\right)^{2ks/n}$ Thus  $P[mr(G) \le k] \le \sum (1/2)^{(s-n)/2}$  $s=n^2/k Z \in Q_s$ 

A Better Proof Not Depending on  $\mathbb F$  $|Q_s| \le \binom{n}{k}^2 \binom{kn}{ks/n}^2 \binom{n^2(k+1) + 2ks/n}{2ks/n}$  $\leq n^{2k} (kn)^{2ks/n} (n^2(k+1) + 2ks/n)^{2ks/n}$  $\leq n^{2ks/n} (n^2)^{2ks/n} (n^3 + 2n^2)^{2ks/n}$  $\leq \left(3n^6\right)^{2ks/n}$ Thus  $P[mr(G) \le k] \le \sum (1/2)^{(s-n)/2}$  $s=n^2/k \ Z \in Q_s$  $n^2$  $\leq \sum (3n^6)^{2ks/n} (1/2)^{(s-n)/2}$  $s = n^2 / k$ 



**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these

columns.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these columns.

Thus for any column v of M, there exists a vector x of length k such that Cx = v.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these columns.

Thus for any column v of M, there exists a vector x of length k such that Cx = v.

Now let C', v' be obtained by keeping only the first k rows of C, v respectively, so that C'x = v'.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these columns.

Thus for any column v of M, there exists a vector x of length k such that Cx = v.

Now let C', v' be obtained by keeping only the first k rows of C, v respectively, so that C'x = v'.

Note that all entries of C', v' lie in E.

**Claim:** Each entry of M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these columns.

Thus for any column v of M, there exists a vector x of length k such that Cx = v.

Now let C', v' be obtained by keeping only the first k rows of C, v respectively, so that C'x = v'.

Note that all entries of C', v' lie in E.

**Cramer's rule:** If  $C'_i$  is the matrix obtained by replacing the *i*-th column of C' by v', then  $x_i = \det(C'_i) / \det(C')$ .

**Correct Claim:** Each entry of det(C')M is a polynomial of degree  $\leq k + 1$  in the variables of E.

**Proof:** Suppose the first k rows and columns of M are linearly independent, and let C be the matrix consisting of these columns.

Thus for any column v of M, there exists a vector x of length k such that Cx = v.

Now let C', v' be obtained by keeping only the first k rows of C, v respectively, so that C'x = v'.

Note that all entries of C', v' lie in E.

**Cramer's rule:** If  $C'_i$  is the matrix obtained by replacing the *i*-th column of C' by v', then  $x_i = \det(C'_i) / \det(C')$ .

Concluding	Remarks	

• Note that our result applies to other problems in which one represents graphs via polynomial relations, such as unit distance graphs, or graphs of touching spheres.

• Note that our result applies to other problems in which one represents graphs via polynomial relations, such as unit distance graphs, or graphs of touching spheres.

• Nelson recently showed that the zero-pattern bound of Rónyai, Babai, and Ganapathy holds for all fields simultaneously and this can be used to show that for the random graph, the minrank over all fields is  $\Theta(n/\log n)$  with high probability.

• Note that our result applies to other problems in which one represents graphs via polynomial relations, such as unit distance graphs, or graphs of touching spheres.

• Nelson recently showed that the zero-pattern bound of Rónyai, Babai, and Ganapathy holds for all fields simultaneously and this can be used to show that for the random graph, the minrank over all fields is  $\Theta(n/\log n)$  with high probability.

• Haviv used the methods of Golovnev, Regev, and Weinstein to randomly construct graphs with large minrank and such that their complement does not contain a copy of some *H*.

• Note that our result applies to other problems in which one represents graphs via polynomial relations, such as unit distance graphs, or graphs of touching spheres.

• Nelson recently showed that the zero-pattern bound of Rónyai, Babai, and Ganapathy holds for all fields simultaneously and this can be used to show that for the random graph, the minrank over all fields is  $\Theta(n/\log n)$  with high probability.

• Haviv used the methods of Golovnev, Regev, and Weinstein to randomly construct graphs with large minrank and such that their complement does not contain a copy of some *H*.

Done!