# Lecture 9: SOS Lower Bound for Knapsack

# Lecture Outline

- Part I: Knapsack Eqations and Pseudo-expectation Values
- Part II: Johnson Scheme
- Part III: Proving PSDness
- Part IV: Further Work

# Part I: Knapsack Eqations and Pseudo-expectation Values

# Knapsack Problem

- Knapsack problem: Given weights $w_1, \ldots, w_n$ and a knapsack with total capacity $C$, what is the maximum weight that can be carried?
- In other words, defining $w_I = \sum_{i \in I} w_i$ for each subset $I \subseteq [1, n]$, what is $\max\{w_I : I \subseteq [1, n], w_I \leq C\}$?
- Here we'll consider the simple case where $w_i = 1$ for all $i$ and $C \in [0, n]$ is not an integer.
- Answer is $\lfloor C \rfloor$, but can SOS prove it?

# Knapsack Equations

- Want $x_i = 1$ if $i \in I$ and $x_i = 0$ otherwise.
- Knapsack equations:
  1. $\forall i, x_i^2 = x_i$
  2. $\sum_{i=1}^{n} x_i = k$
- Here we take $k \in [0, n]$ to be a non-integer.
- Equations are infeasible because $\sum_{i=1}^{n} x_i \in \mathbb{Z}$

# SOS Lower Bound for Knapsack

- Theorem[Gri01]: SOS needs degree at least $2\min\{k, n-k\}$ to refute these equations
- We'll follow the presentation of [MPW15] and show a lower bound of $\min\{k, n-k\}$
- Note: This presentation was already in the retracted paper [MW13]

# Review: SOS Lower Bound Strategy

- Recall: To prove an SOS lower bound, we generally do the following:
  1. Come up with pseudo-expectation values $\tilde{E}$ which obey the required linear equations
  2. Show that the moment matrix $M$ is PSD
- Here we'll use symmetry for part 1 and some combinatorics for part 2.

# Pseudo-expectation Values

- Define $x_I = \prod_{i \in I} x_i$
- $\forall I, (\sum_{j=1}^{n} x_j) x_I = \sum_{j \in I} x_j x_I + \sum_{j \notin I} x_j x_I = k x_I$
- If $\tilde{E}[x_I]$ only depends on $|I|$,

$$\forall I, j \notin I, |I| \tilde{E}[x_I] + (n - |I|) \tilde{E}[x_{I \cup \{j\}}] = k \tilde{E}[x_I]$$

$$\forall I, j \notin I, \tilde{E}\left[x_{I \cup \{j\}}\right] = \frac{k - |I|}{n - |I|} \tilde{E}[x_I]$$

- Thus, $\tilde{E}[x_I] = \frac{k(k-1)\ldots(k-|I|+1)}{n(n-1)\ldots(n-|I|+1)} = \frac{\binom{k}{|I|}}{\binom{n}{|I|}}$

# Viewing $\tilde{E}$ as an Expectation

- $\tilde{E}[x_I] = \frac{\binom{k}{|I|}}{\binom{n}{|I|}}$
- Could have predicted this as follows: If we had a set $A$ of 1s of size $k$, then of the $\binom{n}{|I|}$ possible sets of size $|I|$, $\binom{k}{|I|}$ of them will be contained in $A$.
- Bayesian view: $\tilde{E}[x_I]$ is the expected value of $x_I$ given what we can compute (in SOS).
- Here it is a true expectation if $k \in \mathbb{Z}$

# Reduction to Multilinear Indices

- Recall from last lecture: If we have constraints $x_i^2 = x_i$ or $x_i^2 = 1$, it is sufficient to consider $\tilde{E}[g^2]$ for multilinear $g$.
- Reason: For every polynomial $g$, there is a multilinear polynomial $g'$ with $\deg(g') \leq \deg(g)$ such that $\tilde{E}\left[g'^2\right] = \tilde{E}[g^2]$.
- Thus, it is sufficient to consider the restriction of $M$ to multilinear indices.

# Reduction to Degree $\frac{d}{2}$ Indices

- Lemma: If we also have the constraint $\sum_{i=1}^{n} x_i = k$, for every polynomial $g$ of degree at most $\frac{d}{2}$, there is a homogeneous, multilinear polynomial $g'$ of degree exactly $\frac{d}{2}$ such that $\tilde{E}\left[g'^2\right] = \tilde{E}[g^2]$.
- Proof idea: Use the following reductions:
  1. $\forall i, x_i^2 f = x_i f$
  2. $\forall I \subseteq [1,n]: |I| < \frac{d}{2}, x_I = \frac{\sum_{i \notin I}^{n} x_{I \cup \{i\}}}{k-|I|}$. To see this, note that $\left(\sum_{i=1}^{n} x_i\right) x_I = k x_I = |I| x_I + \sum_{i \notin I}^{n} x_{I \cup \{i\}}$

# Reduction to Degree $\frac{d}{2}$ Indices

- Corollary: To prove that $M \succcurlyeq 0$, it is sufficient to prove that the submatrix of $M$ with multilinear entries of degree exactly $\frac{d}{2}$ is PSD.

# Part II: Johnson Scheme

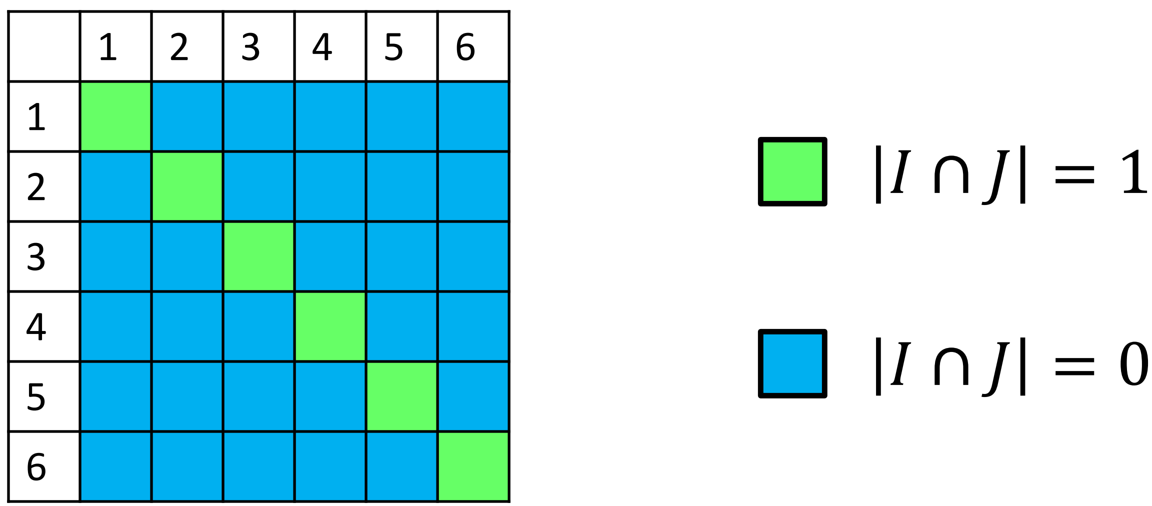
# Johnson Scheme

- Algebra of matrices $M$ such that:
  1. The rows and columns of $M$ are indexed by subsets of $[1, n]$ of size $r$ for some $r$.
  2. $M_{IJ}$ only depends on $|I \cap J|$
- Equivalently, the Johnson Scheme is the algebra of matrices which are invariant under permutations of $[1, n]$.
- Claim: The matrices $M$ in the Johnson scheme are all symmetric and commute with each other

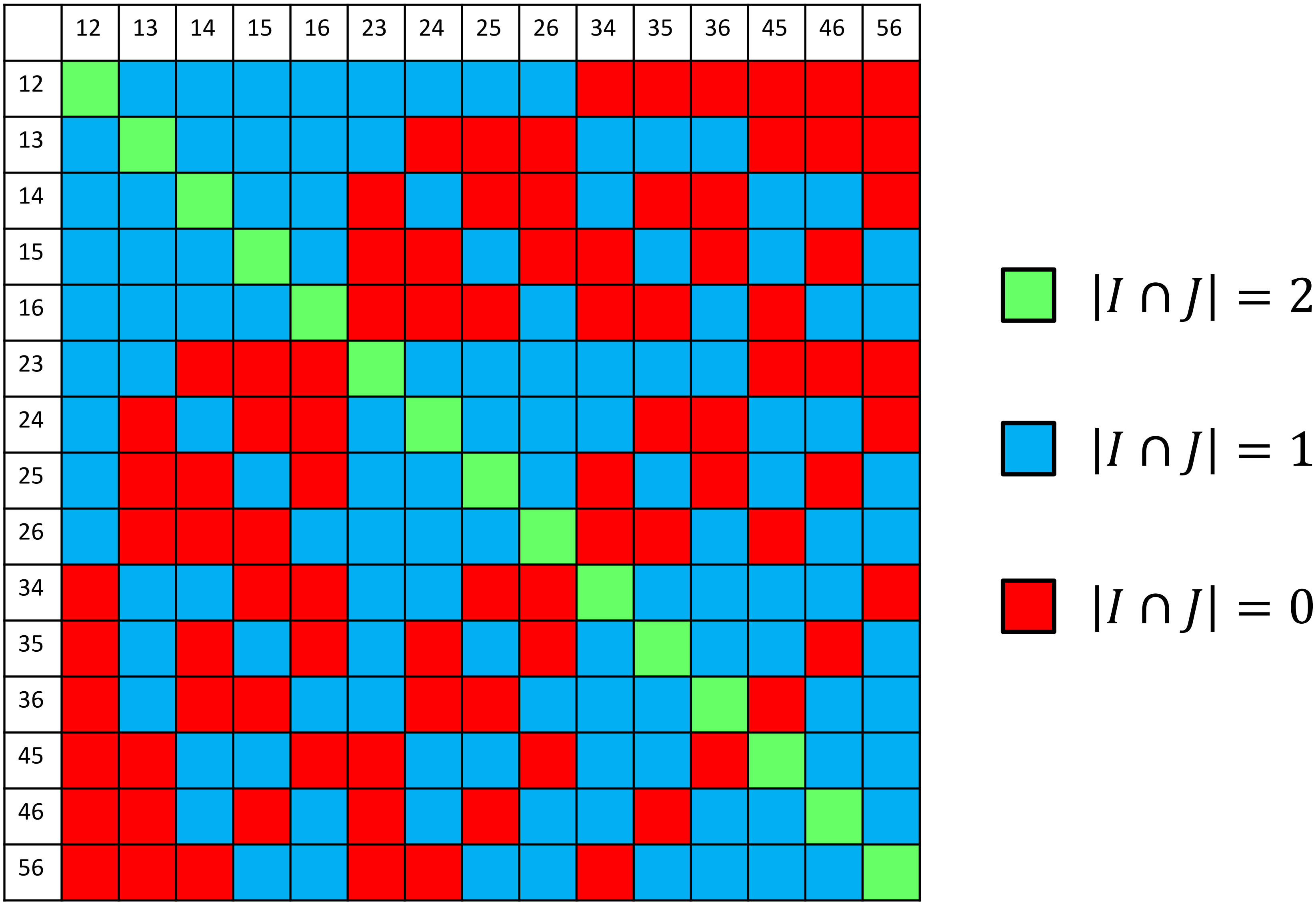
# Johnson Scheme Claim Proof

- Claim: For all $A, B$ in the Johnson scheme, $A^T = A$, $AB$ is in the Johnson scheme as well, and $AB = BA$
- Proof: For the first part, $\forall I, J$, $A_{IJ} = A_{JI}$ because $|I \cap J| = |J \cap I|$. For the second part, $AB_{IK} = \sum_{J \in \binom{n}{r}} A_{IJ} B_{JK}$. Now observe that for any permutation $\sigma$ of $[1, n]$, $AB_{IK} = \sum_{J \in \binom{n}{r}} A_{IJ} B_{JK} = \sum_{J \in \binom{n}{r}} A_{\sigma(I)J} B_{J\sigma(K)} = AB_{\sigma(I)\sigma(K)}$
- For the third part, $AB = (AB)^T = B^T A^T = BA$

# Johnson Scheme Picture for $r = 1$



# Johnson Scheme Picture for $r = 2$



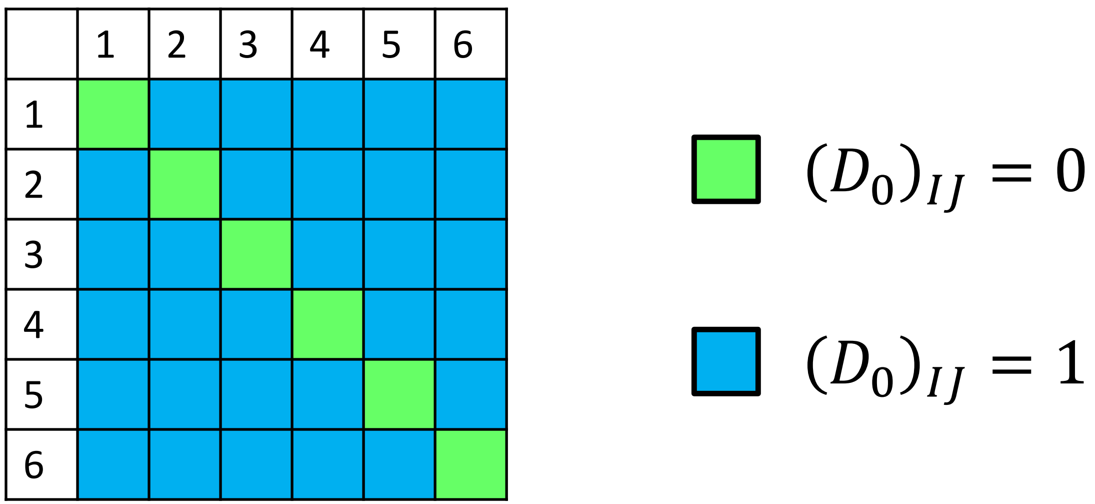
# Basis for Johnson Scheme

- Natural basis for Johnson Scheme: Define $D_a \in \mathbb{R}^{\binom{n}{r} \times \binom{n}{r}}$ to have entries $(D_a)_{IJ} = 1$ if $|I \cap J| = a$ and $(D_i)_{IJ} = 0$ if $|I \cap J| \neq a$.
- Easy to express matrices in this basis, but not so easy to show PSDness

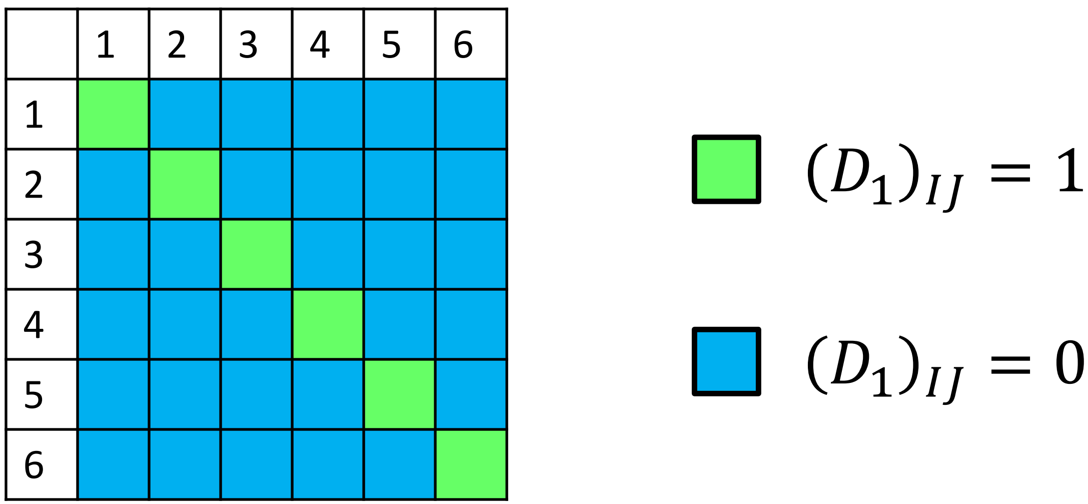
# PSD Basis for Johnson Scheme

- Want a convenient basis of PSD matrices.
- Building block: Define $v_A$ so that $(v_A)_I = 1$ if $\mathrm{A} \subseteq I$ and $0$ otherwise
- PSD basis for Johnson Scheme: Define $P_a \in \mathbb{R}^{\binom{n}{r} \times \binom{n}{r}}$ to be $P_a = \sum_{A \subseteq [1,n]:|A|=a} v_A v_A^T$
- $P_a$ has entries $(P_a)_{IJ} = \binom{|I \cap J|}{a}$ because $v_A v_A^T = 1$ if and only if $A \subseteq I \cap J$ and there are $\binom{|I \cap J|}{a}$ such $A \subseteq [1, n]$ of size $a$.

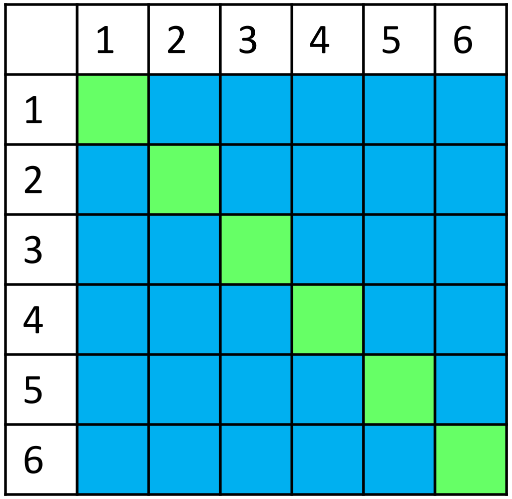
# Basis for $r = 1$



# Basis for $r = 1$



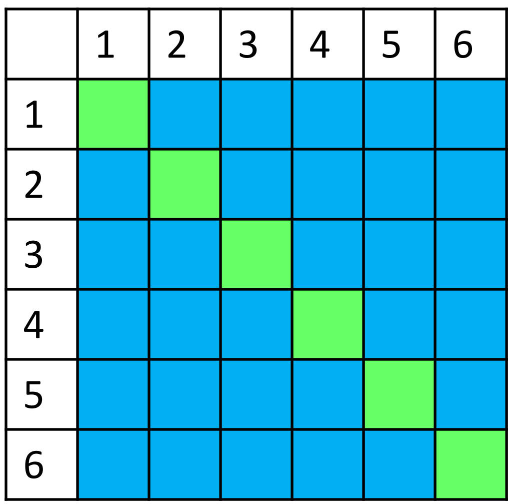
# PSD Basis for $r = 1$




$$(P_0)_{IJ} = \binom{1}{0} = 1$$

$$(P_0)_{IJ} = \binom{0}{0} = 1$$

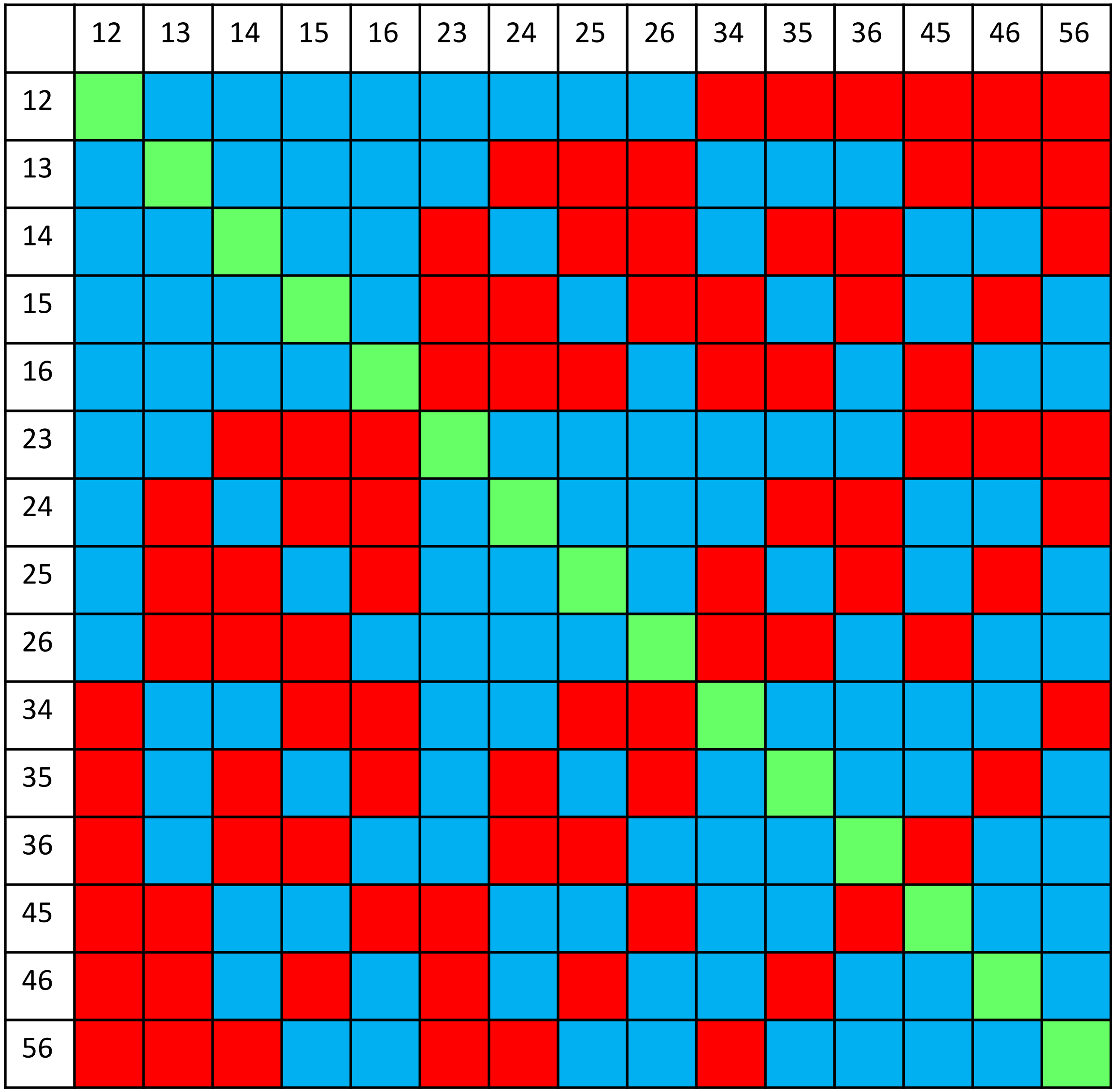
# PSD Basis for $r = 1$




$$(P_1)_{IJ} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1$$

$$(P_1)_{IJ} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

# PSD Basis for $r = 2$

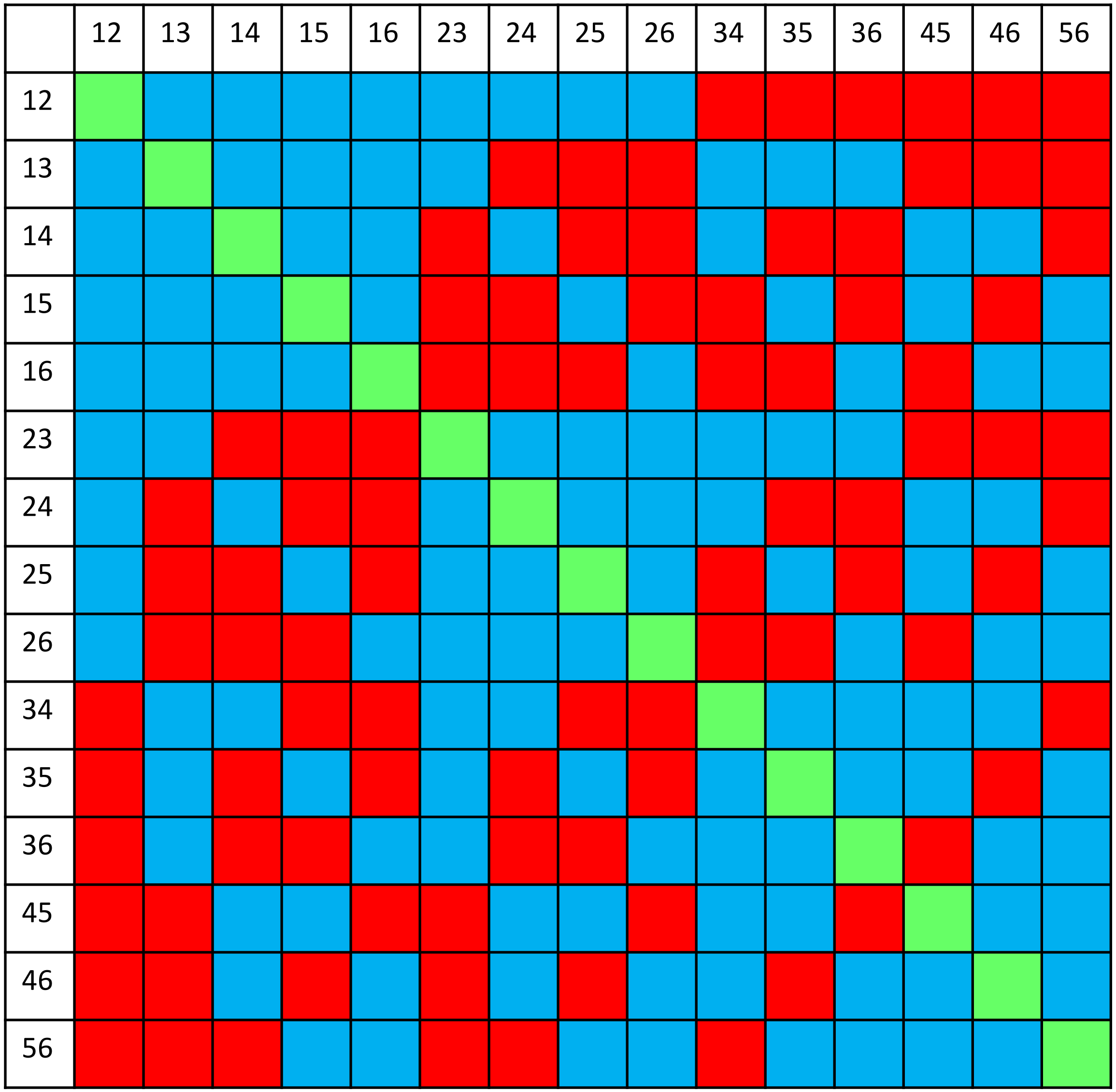



$$(P_0)_{IJ} = \binom{2}{0} = 1$$

$$(P_0)_{IJ} = \binom{1}{0} = 1$$

$$(P_0)_{IJ} = \binom{0}{0} = 1$$

# PSD Basis for $r = 2$

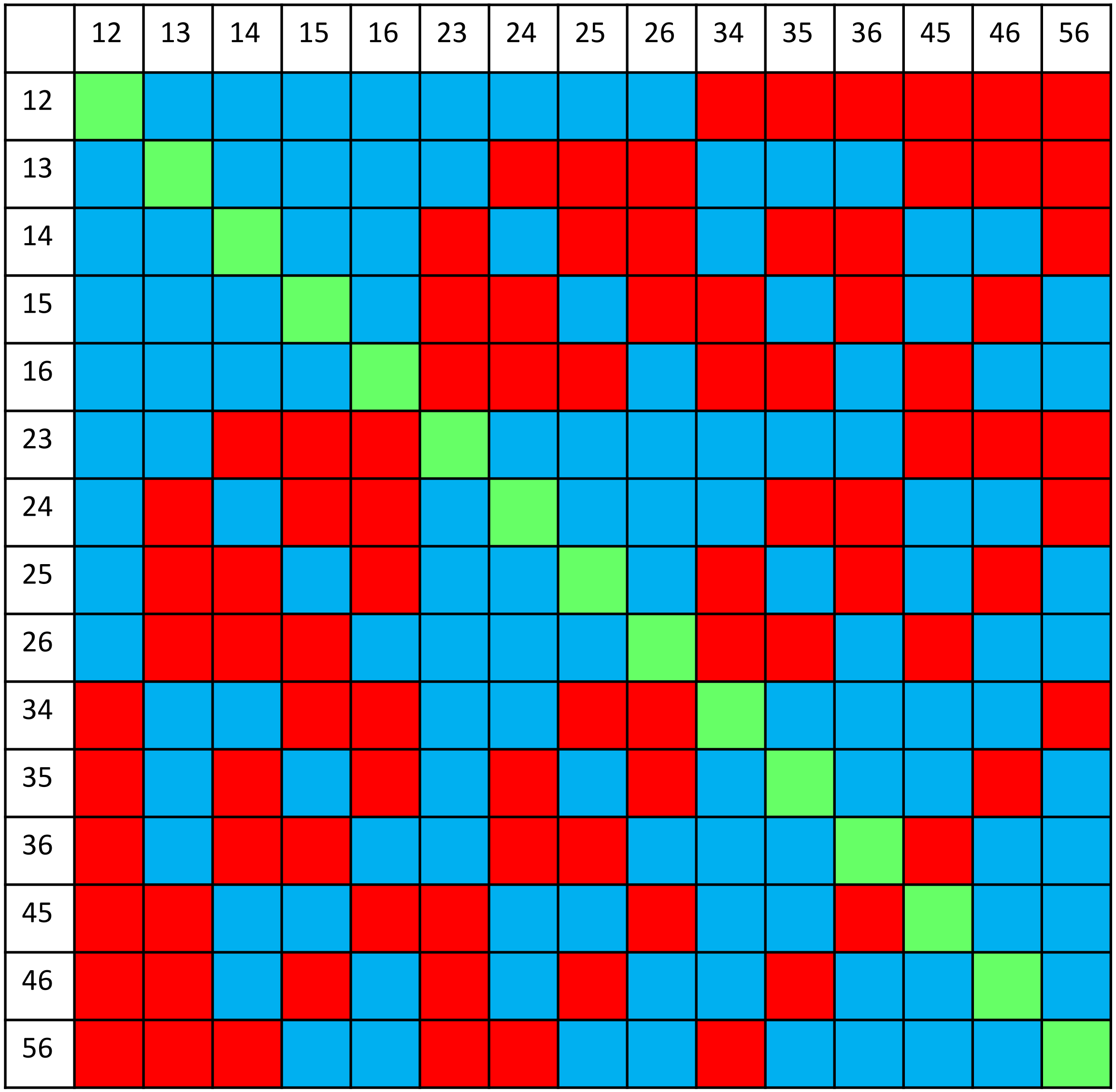



$(P_1)_{IJ} = \binom{2}{1} = 2$

$(P_1)_{IJ} = \binom{1}{1} = 1$

$(P_1)_{IJ} = \binom{0}{1} = 0$

# PSD Basis for $r = 2$




$(P_2)_{IJ} = \binom{2}{2} = 1$

$(P_2)_{IJ} = \binom{1}{2} = 0$

$(P_2)_{IJ} = \binom{0}{2} = 0$

# Shifting Between Bases

- Basis for Johnson Scheme: $(D_a)_{IJ} = \delta_{a|I \cap J|}$
- PSD Basis for Johnson Scheme : $(P_a)_{IJ} = \binom{|I \cap J|}{a}$
- Want to shift between bases.
- Lemma:
  1. $P_a = \sum_{b=a}^{r} \binom{b}{a} D_b$
  2. $D_a = \sum_{b=a}^{r} (-1)^{b-a} \binom{b}{a} P_b$
- First part is trivial, second part follows from a bit of combinatorics.

# Shifting Between Bases Proof

- Lemma:
  1. $P_a = \sum_{b=a}^{r} \binom{b}{a} D_b$
  2. $D_a = \sum_{b=a}^{r} (-1)^{b-a} \binom{b}{a} P_b$
- Proof of the second part: Observe that $\sum_{b=a}^{r} (-1)^{b-a} \binom{b}{a} P_b = \sum_{a'=a}^{r} \sum_{b=a}^{a'} (-1)^{b-a} \binom{b}{a} D_b$
- Must show that for all $a' \geq a$,

$$\sum_{b=a}^{a'} (-1)^{b-a} \binom{a'}{b} \binom{b}{a} = \delta_{a'a}$$

- In-class exercise: Prove this

# Shifting Between Bases Proof

- Need to show: $\sum_{b=a}^{a'}(-1)^{b-a}\binom{a'}{b}\binom{b}{a} = \delta_{a'a}$
- Answer: Observe that

$$\binom{a'}{b}\binom{b}{a} = \frac{a'!b!}{b!(a'-b)!a!(b-a)!} = \frac{a'!}{a!(a'-a)!}\frac{(a'-a)!}{(a'-b)!(b-a)!}$$

- Our expression is equal to

$$\frac{a'!}{a!(a'-a)!}\sum_{j=0}^{m}(-1)^{j}\binom{m}{j} \text{ where } m = a' - a$$

- Now note that $\sum_{j=0}^{m}(-1)^{j}\binom{m}{j} = \left(1+(-1)\right)^{m}$, which equals $1$ if $m = 0$ and 0 if $m > 0$.

# Part III: Proving PSDness

# Decomposition of $M$

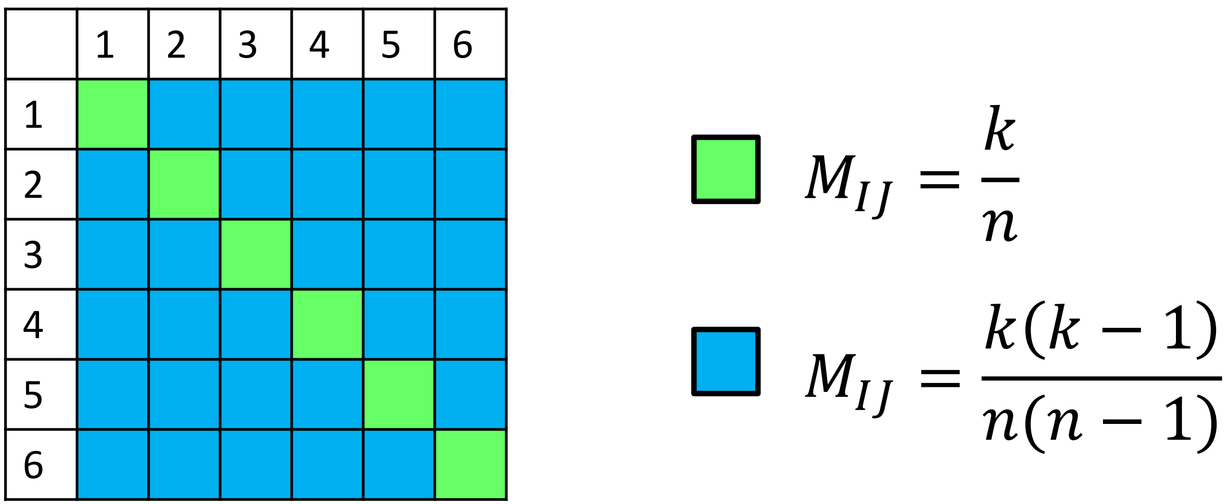
- Recall that $\tilde{E}[x_I] = \frac{\binom{k}{|I|}}{\binom{n}{|I|}}$
- $M_{IJ} = \frac{\binom{k}{|I \cup J|}}{\binom{n}{|I \cup J|}}$
- Thus, $M = \sum_{a=0}^{r} \frac{\binom{k}{2r-a}}{\binom{n}{2r-a}} D_a$

# PSD Decomposition

- To prove $M \succcurlyeq 0$, it is sufficient to express $M$ as a non-negative linear combination of the matrices $P_a$.

# Example: Decomposition for $r = 1$

- $M = \frac{k}{n} D_1 + \frac{k(k-1)}{n(n-1)} D_0 = \frac{k}{n} P_1 + \frac{k(k-1)}{n(n-1)} (P_0 - P_1)$
- $M = \left(\frac{k}{n} - \frac{k(k-1)}{n(n-1)}\right) P_1 + \frac{k(k-1)}{n(n-1)} P_0 = \frac{k(n-k)}{n(n-1)} P_1 + \frac{k(k-1)}{n(n-1)} P_0$



# PSD Decomposition

- Claim: $M = \sum_{a=0}^{r} \frac{\binom{k}{2r}}{\binom{n}{2r}} \cdot \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} P_a$
- For the proof, see the appendix
- Corollary: $M \succcurlyeq 0$ if $k \geq 2r$ and $n - k \geq r$ (where $d = 2r$)

# Improving Degree Lower Bound

- $\{P_a\}$ is a nice basis to work with because it is relatively easy to go between $\{D_a\}$ and $\{P_a\}$.
- However, in some sense, it's not the right basis to use.
- Want a basis $\{P'_a\}$ such that all symmetric PSD matrices are a non-negative linear combination of the $\{P'_a\}$.
- With the right basis, can get a higher degree lower bound.

# Example

- Let $J$ be the all ones matrix.
- For the case $d = 2, r = 1, P_0 = J$ and $P_1 = Id$
- Better basis: $P_0' = J, P_1' = \frac{n-1}{n}Id - \frac{1}{n}J$

# Part IV: Further Work

# Using Symmetry

- Can we take advantage of symmetry in the problem more generally?
- Yes!

# Using Symmetry

- Proposition: Whenever there are valid pseudo-expectation values, there are valid pseudo-expectation values which are symmetric.
- Proof: Let $S$ be the group of symmetries of the problem. If we have pseudo-expectation values $\tilde{E}$, then for any $\sigma \in S$, $\widetilde{E'}[\mathrm{f}] = \tilde{\mathrm{E}}[\sigma(f)]$ is also valid. Since the conditions for pseudo-expectation values are convex, $\widetilde{E_{avg}}[f] = \frac{\tilde{E}\left[\sum_{\sigma \in S} \sigma(f)\right]}{|S|}$ is valid as well and is symmetric.

# Using Symmetry

- Gatermann and Parrilo [GP04] show how symmetry can be used to drastically reduce the search space for finding pseudo-expectation values.
- Recently, Raymond, Saunderson, Singh, and Thomas [RSST16] showed that if the problem is symmetric, it can be solved with a semidefinite program whose size is independent of $n$.

# Obtaining Lower Bounds Directly

- One way to give intuition for the lower bound: SOS "thinks" that we are choosing $k$ elements out of $n$ and takes the corresponding pseudo-expectation values.
- SOS is very bad at determining functions must be integers and needs degree $\geq k$ to detect a problem.

# Obtaining Lower Bounds Directly

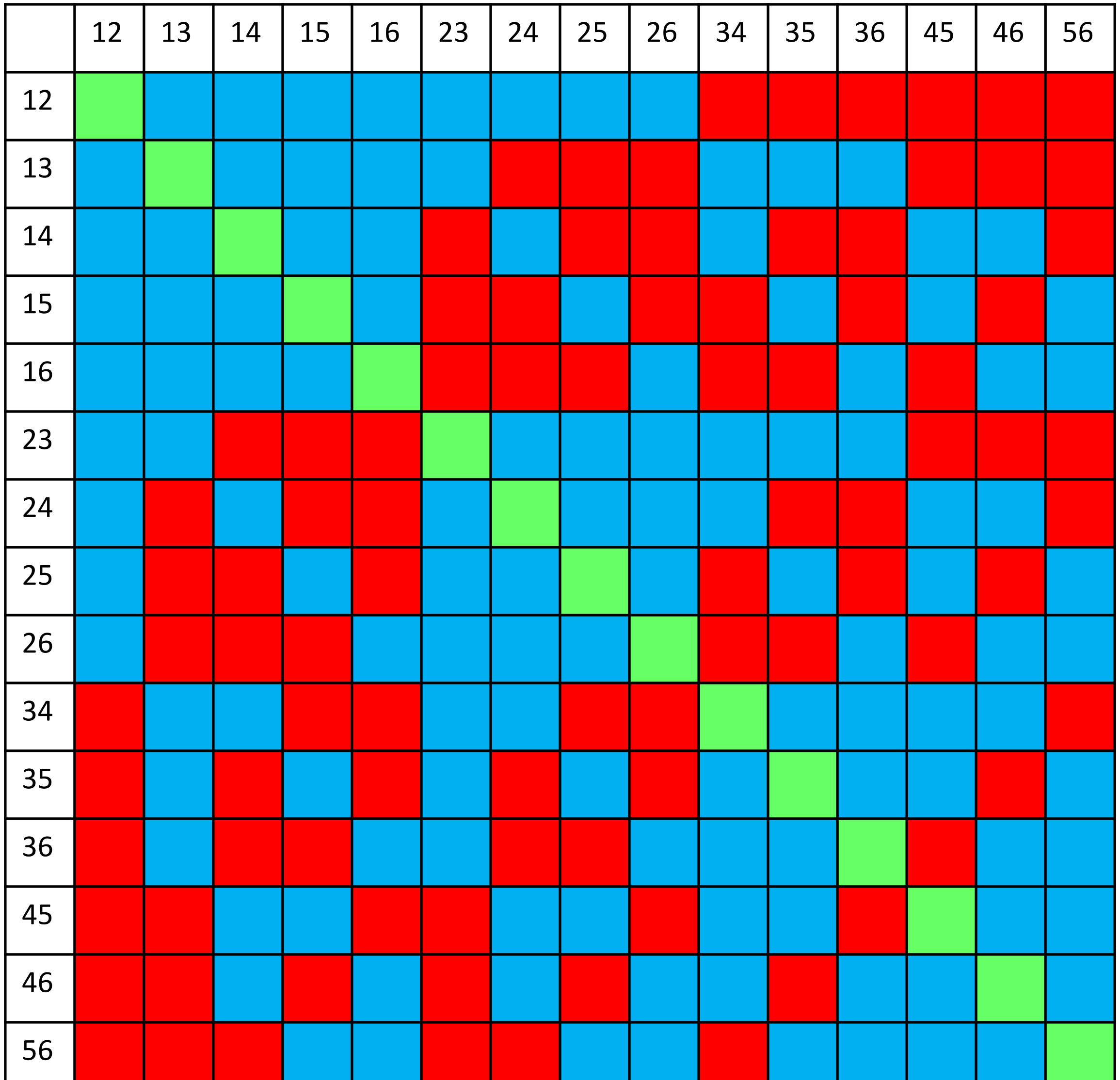
- Is there a way to say that this intuition is good enough to obtain a lower bound without going through the combinatorics?
- Unless I’m mistaken, yes (this is work in progress).

# References


- [GP04] K. Gatermann and P. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. J. Pure Appl. Algebra, 192(1-3):95–128, 2004.
- [Gri01] D. Grigoriev. Complexity of Positivstellensatz proofs for the knapsack. Computational Complexity 10(2):139–154, 2001
- [MW13] R. Meka and A. Wigderson Association schemes, non-commutative polynomial concentration, and sum-of-squares lower bounds for planted clique. https://arxiv.org/abs/1307.7615v1
- [MPW15] R. Meka, A Potechin, A. Wigderson, Sum-of-squares lower bounds for planted clique. STOC p.87–96, 2015
- [RSST16] A. Raymond, J. Saunderson, M. Singh, R. Thomas. Symmetric sums of squares over k-subset hypercubes. https://arxiv.org/abs/1606.05639, 2016

# Appendix: PSD Decomposition Calculations

# Picture for $r = 2$




$$M_{IJ} = \frac{\binom{k}{2}}{\binom{n}{2}}$$

$$M_{IJ} = \frac{\binom{k}{3}}{\binom{n}{3}}$$

$$M_{IJ} = \frac{\binom{k}{4}}{\binom{n}{4}}$$

# Decomposition for $r = 2$

- $M = \frac{\binom{k}{2}}{\binom{n}{2}} D_2 + \frac{\binom{k}{3}}{\binom{n}{3}} D_1 + \frac{\binom{k}{4}}{\binom{n}{4}} D_0$
- $M = \frac{\binom{k}{2}}{\binom{n}{2}} P_2 + \frac{\binom{k}{3}}{\binom{n}{3}} (P_1 - 2P_2) + \frac{\binom{k}{4}}{\binom{n}{4}} (P_0 - P_1 + P_2)$
- $M = \left(\frac{\binom{k}{2}}{\binom{n}{2}} - 2\frac{\binom{k}{3}}{\binom{n}{3}} + \frac{\binom{k}{4}}{\binom{n}{4}}\right) P_2 + \left(\frac{\binom{k}{3}}{\binom{n}{3}} - 2\frac{\binom{k}{4}}{\binom{n}{4}}\right) P_1 + \frac{\binom{k}{4}}{\binom{n}{4}} P_0$
- $\frac{\binom{k}{4}}{\binom{n}{4}} = \frac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)}$
- $\left(\frac{\binom{k}{3}}{\binom{n}{3}} - \frac{\binom{k}{4}}{\binom{n}{4}}\right) = \frac{k(k-1)(k-2)((n-3)-(k-3))}{n(n-1)(n-2)(n-3)} = \frac{k(k-1)(k-2)(n-k)}{n(n-1)(n-2)(n-3)}$

# Decomposition for $r = 2$

- Claim: $\left(\frac{\binom{k}{2}}{\binom{n}{2}} - 2\frac{\binom{k}{3}}{\binom{n}{3}} + \frac{\binom{k}{4}}{\binom{n}{4}}\right) = \frac{k(k-1)(n-k)(n-k-1)}{n(n-1)(n-2)(n-3)}$
- Proof: Consider $\frac{n(n-1)(n-2)(n-3)}{k(k-1)}\left(\frac{\binom{k}{2}}{\binom{n}{2}} - 2\frac{\binom{k}{3}}{\binom{n}{3}} + \frac{\binom{k}{4}}{\binom{n}{4}}\right)$. This equals $(n-2)(n-3) - 2(k-2)(n-3) + (k-2)(k-3)$ which equals

$$\big(n - 2 - (k-2)\big)(n-3) - (k-2)(n-3\ - (k-3))$$
$$= (n-k)\big((n-3) - (k-2)\big) = (n-k)(n-k-1)$$

# General Pattern

- $M = \frac{k(k-1)(n-k)(n-k-1)}{n(n-1)(n-2)(n-3)} P_2 + \frac{k(k-1)(k-2)(n-k)}{n(n-1)(n-2)(n-3)} P_1 + \frac{k(k-1)(k-2)(k-3)}{n(n-1)(n-2)(n-3)} P_0$
- Can you see the pattern?
- General Pattern: $M = \frac{\binom{k}{2r}}{\binom{n}{2r}} \left( \sum_{a=0}^{r} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} P_a \right)$

# General Pattern Proof

- Claim: $M = \frac{\binom{k}{2r}}{\binom{n}{2r}}\left(\sum_{a=0}^{r} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} P_a\right)$
- This gives $M = \frac{\binom{k}{2r}}{\binom{n}{2r}}\left(\sum_{a=0}^{r} \sum_{b=a}^{r} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} \binom{b}{a} D_b\right)$
- $M = \frac{\binom{k}{2r}}{\binom{n}{2r}}\left(\sum_{b=0}^{r} \sum_{a=0}^{b} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} \binom{b}{a} D_b\right)$
- Need to show: $\sum_{a=0}^{b} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} \binom{b}{a} = \frac{\binom{n-2r+b}{b}}{\binom{k-2r+b}{b}}$

# General Pattern Proof

- Claim: $\sum_{a=0}^{b} \frac{\binom{n-k}{a}}{\binom{k-2r+a}{a}} \binom{b}{a} = \frac{\binom{n-2r+b}{b}}{\binom{k-2r+b}{b}}$

- Proof: Note that $\frac{\binom{k-2r+b}{b}}{\binom{k-2r+a}{a}} = \frac{\binom{k-2r+b}{b-a}}{\binom{b}{a}}$, so this is equivalent to the following:

$$\sum_{a=0}^{b} \binom{n-k}{a} \binom{k-2r+b}{b-a} = \binom{n-2r+b}{b}$$

# General Pattern Proof

- Claim: $\sum_{a=0}^{b}\binom{n-k}{a}\binom{k-2r+b}{b-a} = \binom{n-2r+b}{b}$
- Proof: One way to choose $b$ elements out of $[1, n-2r+b]$ elements is to first choose the number $a$ of elements which will be in $[1, n-k]$. We then choose $a$ elements from $[1, n-k]$ and choose the remaining $b-a$ elements from $[n-k+1, n-2r+b]$, which gives $\binom{n-k}{a}\binom{k-2r+b}{b-a}$ choices for each $a \in [0, b]$.