II. Group-Theoretic Approach

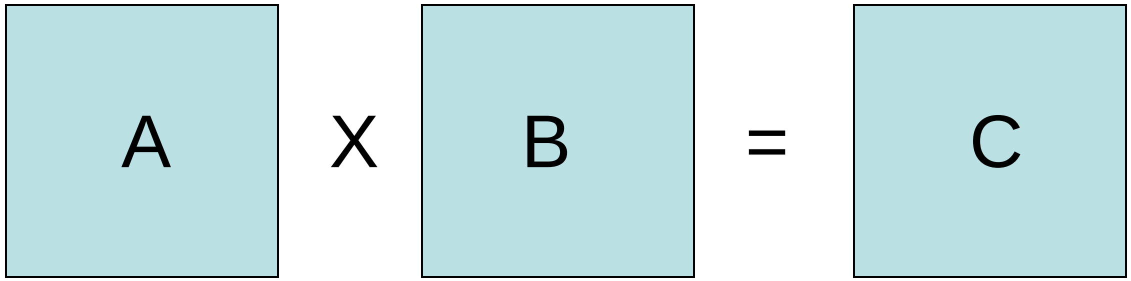
Chris Umans

Caltech

Based on joint work with Noga Alon, Henry Cohn, Bobby Kleinberg, Amir Shpilka, Balazs Szegedy

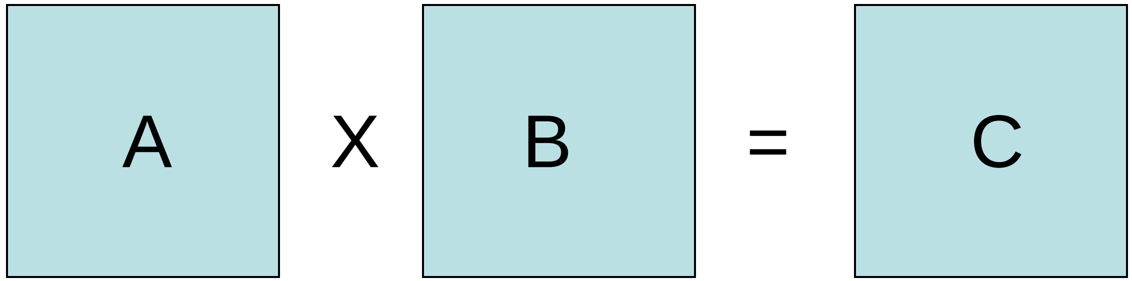
Modern Applications of Representation Theory, IMA, Chicago July 2014

Introduction



- Standard method: $O(n^3)$ operations
- Strassen (1969): $O(n^{2.81})$ operations

Introduction



- Standard method: $O(n^3)$ operations
- Strassen (1969): $O(n^{2.81})$ operations

The exponent of matrix multiplication:

smallest number ω such that for all $\varepsilon>0$

$O(n^{\omega + \varepsilon})$ operations suffice

History

- Standard algorithm $\omega \leq 3$
- Strassen (1969) $\omega < 2.81$
- Pan (1978) $\omega < 2.79$
- Bini; Bini et al. (1979) $\omega < 2.78$
- Schönhage (1981) $\omega < 2.55$
- Pan; Romani; Coppersmith
  + Winograd (1981-1982) $\omega < 2.50$
- Strassen (1987) $\omega < 2.48$
- Coppersmith + Winograd (1987) $\omega < 2.375$
- Stothers (2010) $\omega < 2.3737$
- Williams (2011) $\omega < 2.3729$
- Le Gall (2014) $\omega < 2.37286$

Outline

Lecture I:

- crash course on main ideas from Strassen 1969 through Le Gall 2014
- conjectures implying ! = 2

This lecture and next one:

II. group-theoretic approach

III. extending to coherent configurations

A different approach

- So far...
  - bound border rank of small tensor (by hand)
  - asymptotic bound from high tensor powers
- Disadvantages
  - limited universe of “starting” tensors
  - high tensor powers hard to analyze

- Next: matrix multiplication via groups

The Group Algebra

- Given a group G

Also think of elements as *vectors* **a** with |G| entries

- The group algebra C[G] has elements

$$\Sigma_g\, a_g g$$

with multiplication

$$(\Sigma_g a_g g)(\Sigma_h b_h h) = \Sigma_f\, (\Sigma_{gh\,=\,f}\; a_g b_h) f$$

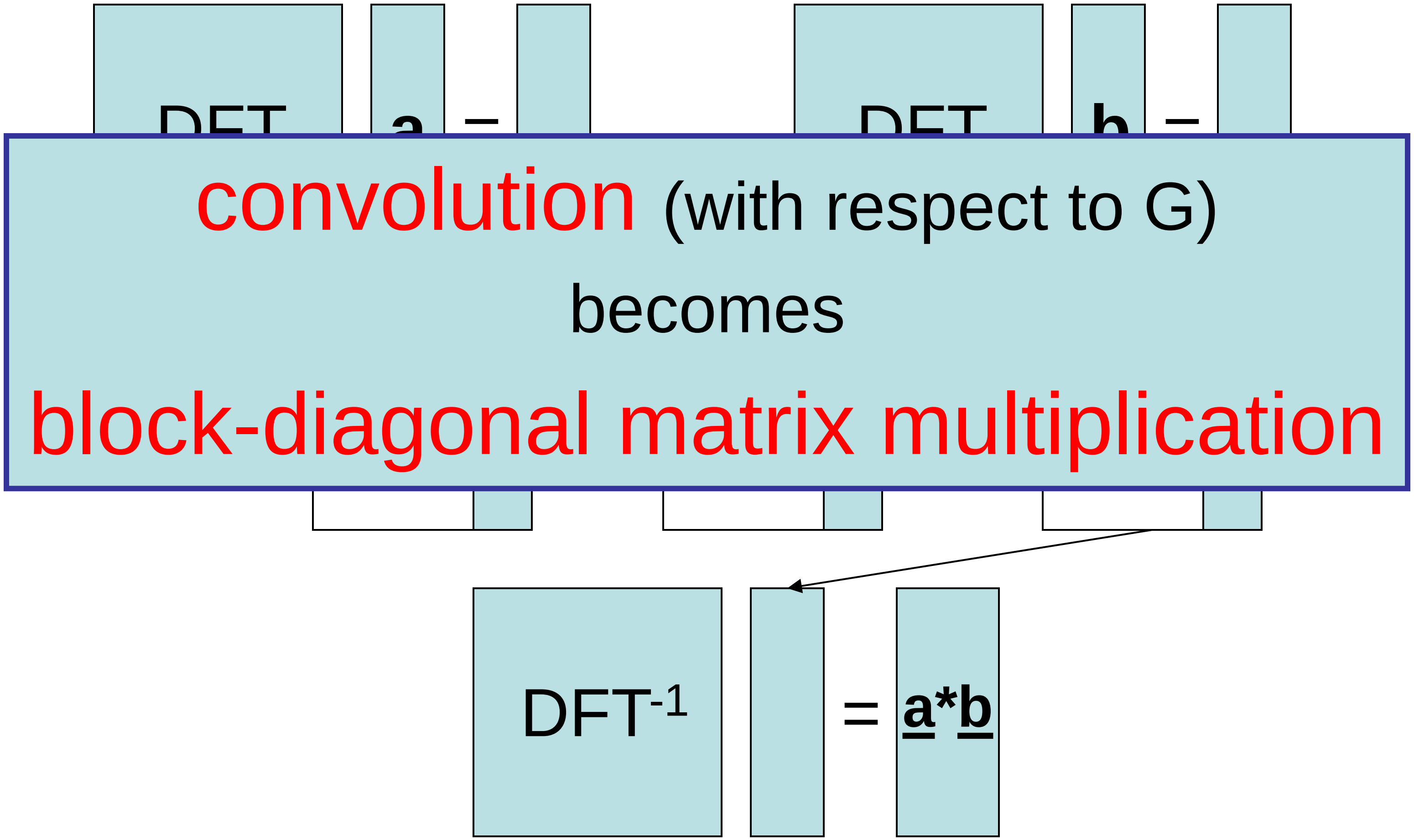
The Group Algebra

$$C[G] \; ' \; (C^{d_1 \times d_1}) \times (C^{d_2 \times d_2}) \times \ldots \times (C^{d_k \times d_k})$$

- $d_1, d_2, \ldots, d_k$ are character degrees of G
- two facts:
  - $\Sigma d_i^2 = |G|$
  - all character degrees are 1 for abelian groups

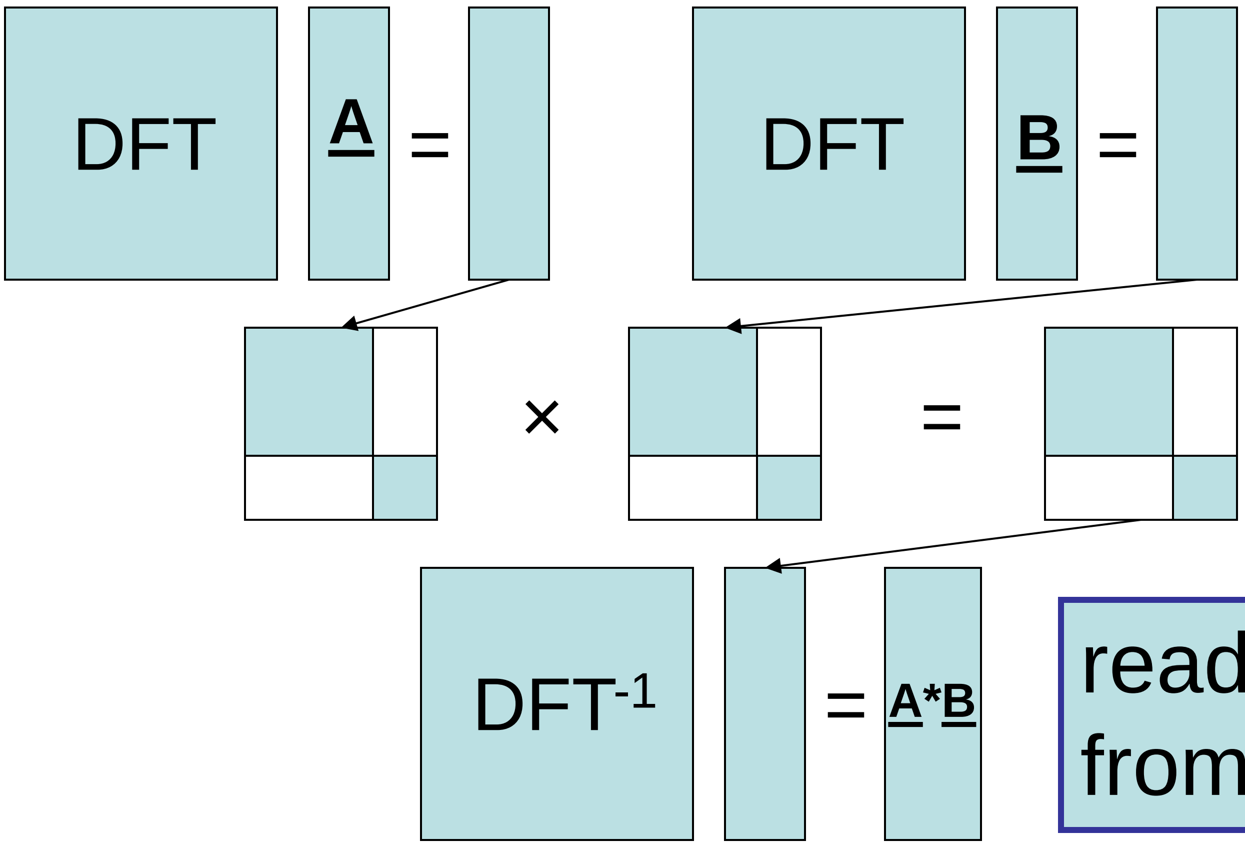
Main idea: multiply in Fourier domain

$$C[G]\ '\ (C^{d_1 \times d_1}) \times (C^{d_2 \times d_2}) \times \ldots \times (C^{d_k \times d_k})$$



Matrix Multiplication

- Two input matrices: $A=(a_{ij})$, $B=(b_{kl})$
- "embed" $A \rightarrow \underline{\mathbf{A}} \in C[G]$, $B \rightarrow \underline{\mathbf{B}} \in C[G]$



read off $C = AB$
from $\underline{\mathbf{A}}*\underline{\mathbf{B}}$

Can this work?

- All depends on choice of group G
- need G to permit an embedding

$$A \to \underline{\mathbf{A}} \in C[G], \quad B \to \underline{\mathbf{B}} \in C[G]$$

so that we can read off entries of AB from

$$\underline{\mathbf{A}}*\underline{\mathbf{B}}.$$

The embedding:

Subgroups X, Y, Z of G satisfy the

triple product property

if for all $x \in X$, $y \in Y$, $z \in Z$:

$xyz = 1$ iff $x = y = z = 1$.

The embedding:

$Q(S) = \{s^{-1}t: s, t \in S\}$

Subsets X, Y, Z of G satisfy the **triple product property** if for all $x \in Q(X)$, $y \in Q(Y)$, $z \in Q(Z)$:

$xyz = 1$ iff $x = y = z = 1$.

A $= \Sigma a_{x\ ,y}\ (x\ y^{-1})$ **B** $= \Sigma b_{y\ ,z}\ (y\ z^{-1})$

Claim: $(AB)_{x\ ,z}$ = coeff. on $(x\ z^{-1})$ in **A*B**.

The embedding:

$Q(S) = \{s^{-1}t : s, t \in S\}$

Subsets X, Y, Z of G satisfy the

triple product property

if for all $x \in Q(X)$, $y \in Q(Y)$, $z \in Q(Z)$:

$xyz = 1$ iff $x = y = z = 1$.

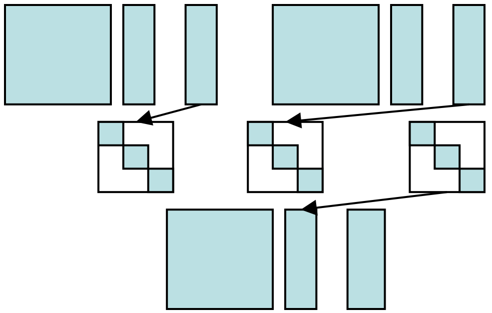
A $= \Sigma a_{x_1,y_1}(x_1 y_1^{-1})$ **B** $= \Sigma b_{y_2,z_2}(y_2 z_2^{-1})$

Claim: $(AB)_{x_3,z_3}$ = coeff. on $(x_3 z_3^{-1})$ in **A*****B**.

$(x_1 y_1^{-1})(y_2 z_2^{-1}) = x_3 z_3^{-1} \Rightarrow x_3^{-1} x_1 y_1^{-1} y_2 z_2^{-1} z_3 = 1$

How many multiplications?

Fact: method to multiply $k \times k$ matrices using m multiplications proves $! \leq \log_k m$



- we use $m \leq \Sigma d_i^3$ mults
- really $m = \Sigma d_i^{!}$ mults
- *at least* $m \geq \Sigma d_i^2 = |G|$ mults

First Challenge: embed $k \times k$ matrix multiplication in group of size $\frac{1}{4} k^2$

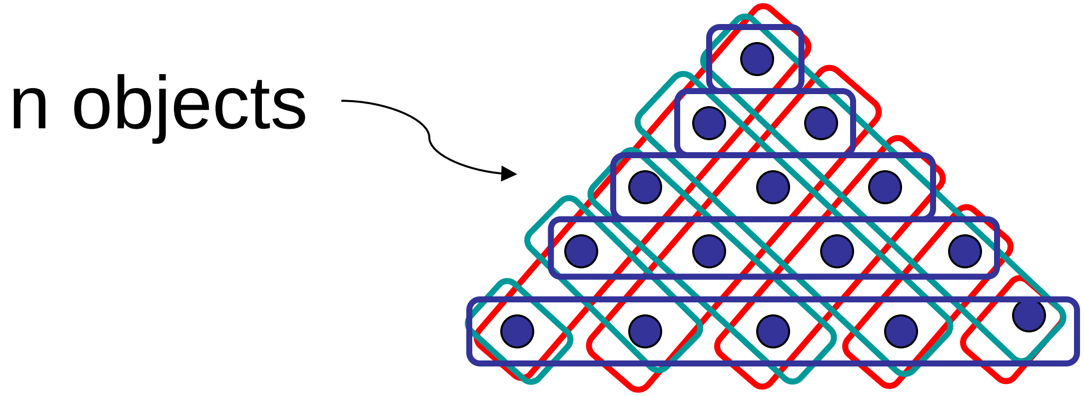
The embedding

First Challenge: embed $k \times k$ matrix multiplication in group of size $\frac{1}{4} k^2$

- simple pigeonhole argument:
  - embedding in an abelian group requires group to have size k^3

The triangle construction

Theorem: can embed k × k matrix multiplication in symmetric group of size $k^{2 + o(1)}$

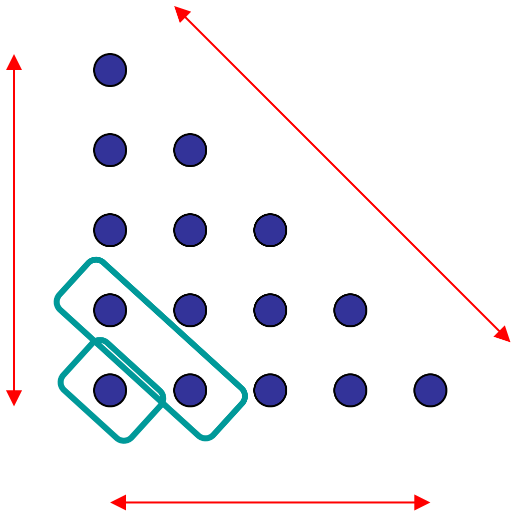


- subgroup X
- subgroup Y
- subgroup Z

need X, Y, Z in S_n all with size ≈ $|S_n|^{1/2}$

The triangle construction

– X moves points within rows
– Y moves points within columns
– Z moves points within diagonals
– want: $xyz = 1 \quad \Rightarrow \quad x = y = z = 1$



Character degrees

- We have described a *reduction* from $k \times k$ mat. mult. to block-diagonal mat. mult.

Theorem: in group G with character degrees $d_1, d_2, d_3, \ldots$, we obtain:

$$k^{\omega} \cdot \sum_i d_i^{\omega}$$

Potential barrier

Can use this framework to prove $\omega < 3$

if and only if

can find X, Y, Z subsets of G satisfying the triple product property, and

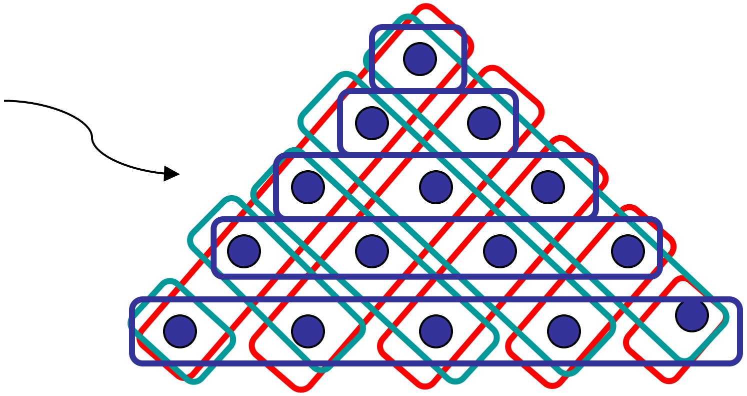
$|X||Y||Z| > \sum d_i^3$.

"beating the sum of the cubes"

Recall: the triangle construction

Theorem: can embed k × k matrix multiplication in symmetric group of size $k^{2 + o(1)}$

n objects



- subgroup X
- subgroup Y
- subgroup Z

unfortunately, $d_{max} > |X| \ (= |Y| = |Z|)$

What should we be aiming for?

Theorem: in group G supporting k x k matrix multiplication with character degrees d_1, d_2, d_3,…, we obtain:

$$k^{\omega} \cdot \sum_i d_i^{\omega}$$

- If X, Y, Z µ G satisfy T.P.P. and
  - $(|X| \text{¢} |Y| \text{¢} |Z|)^{1/3} = k$ ¸ $|G|^{1/2 – o(1)}$
  - $d_{max} \cdot |G|^{1/2 – ²}$

  then ! = 2

$\sum_i d_i^{!} \cdot$
$d_{max}^{! – 2}|G|$

Constructions in linear groups

- Good candidate family:

  SL(n, q) for fixed dimension n

  because $d_{max} \cdot |G|^{1/2 - {}^{2}n}$

- a non-trivial construction (i.e., $k^3 > |G|$):

$$X = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \right\} \quad Y = \left\{ \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \right\} \quad Z = \left\{ \begin{pmatrix} 1+z & z \\ -z & 1-z \end{pmatrix} \right\}$$

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} = \begin{pmatrix} 1+xy & x \\ y & 1 \end{pmatrix}$$

Constructions in linear groups

- Good candidate family:

  SL(n, q) for fixed dimension n

  because $d_{max} \cdot |G|^{1/2 - {}^{2}n}$

- best we know, in SL(2, q) for $q = p^2$:

$$X = \left\{ \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \right\} \quad Y = \left\{ \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} \right\} \quad Z = \left\{ \begin{bmatrix} z & w \\ \overline{w} & \overline{z} \end{bmatrix} \right\}$$

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} = \begin{bmatrix} 1+xy & x \\ y & 1 \end{bmatrix}$$

Constructions in linear groups

- Good candidate family:

  SL(n, q) for fixed dimension n

  because d_{max} · $|G|^{1/2\ -\ ²\mathbf{n}}$

- best we know, in SL(2, q) for $q = p^2$:

$$X = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \right\} \quad Y = \left\{ \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \right\} \quad Z = \left\{ \begin{pmatrix} z & w \\ \overline{w} & \overline{z} \end{pmatrix} \right\}$$

  – $(|X|¢|Y|¢|Z|)^{1/3} = |G|^{18/7 – o(1)}$

Constructions in linear groups

- Good candidate family:

  SL(n, q) for fixed dimension n

- In SL(n, R) these three subgroups satisfy the triple product property:
  - upper-triangular with ones on the diagonal
  - lower-triangular with ones on the diagonal
  - the special orthogonal group SO(n, R)

and dim. of each is ½ dim. of G as n ! 1

an example

yielding $\omega < 3$

Wreath product groups

- A abelian group
- G semidirect product of

$(A^w)^N$ and S_N (symmetric group)

Ã w !

N rows

6	9	8
0	7	4
8	3	0
0	6	2

+

3	7	3
8	4	5
1	3	2
0	9	5

=

9	6	1
8	1	9
9	6	2
0	5	7

Wreath product groups

- A abelian group
- G semidirect product of

$(A^W)^N$ and S_N (symmetric group)

$$\pi \begin{bmatrix} 6 & 9 & 8 \\ 0 & 7 & 4 \\ 8 & 3 & 0 \\ 0 & 6 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 6 & 2 \\ 6 & 9 & 8 \\ 0 & 7 & 4 \\ 8 & 3 & 0 \end{bmatrix} \pi$$

Beating the sum of the cubes

Three subsets of $(A^3)^2$ semidirect S_2:

$$X = \left\{ \begin{bmatrix} X & 0 & 0 \\ 0 & X' & 0 \end{bmatrix} \pi \right\}$$

$$Y = \left\{ \begin{bmatrix} 0 & Y & 0 \\ 0 & 0 & Y' \end{bmatrix} \pi \right\}$$

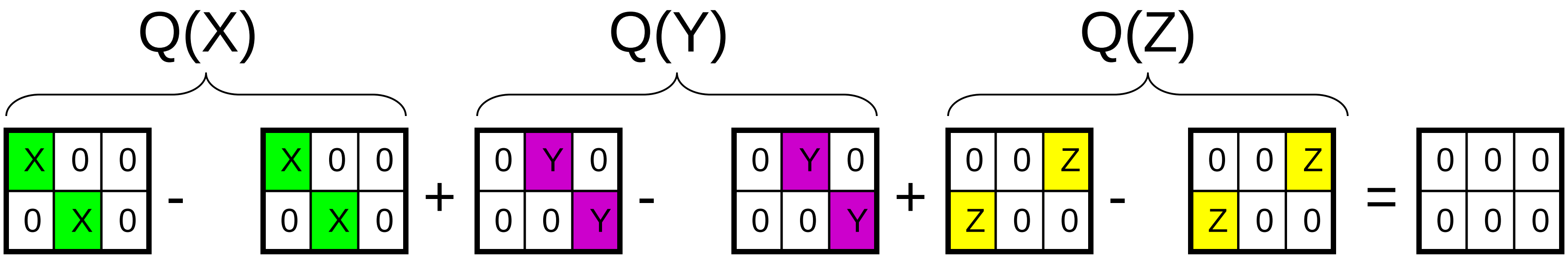
$$Z = \left\{ \begin{bmatrix} 0 & 0 & Z \\ Z' & 0 & 0 \end{bmatrix} \pi \right\}$$

Beating the sum of the cubes

$$\overbrace{\begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \end{bmatrix} - \pi \begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \end{bmatrix}}^{Q(X)} + \overbrace{\begin{bmatrix} 0 & Y & 0 \\ 0 & 0 & Y \end{bmatrix} - \rho \begin{bmatrix} 0 & Y & 0 \\ 0 & 0 & Y \end{bmatrix}}^{Q(Y)} + \overbrace{\begin{bmatrix} 0 & 0 & Z \\ Z & 0 & 0 \end{bmatrix} - \tau \begin{bmatrix} 0 & 0 & Z \\ Z & 0 & 0 \end{bmatrix}}^{Q(Z)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

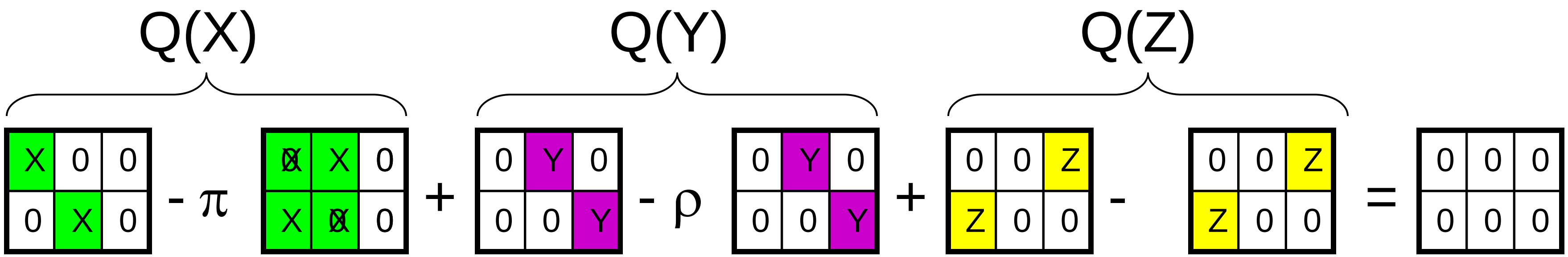
- Group is $(A^3)^2$ semidirect S_2
- π, ρ, τ 2 S_2 either “flip” or “no flip”
- must be even number of flips

Beating the sum of the cubes



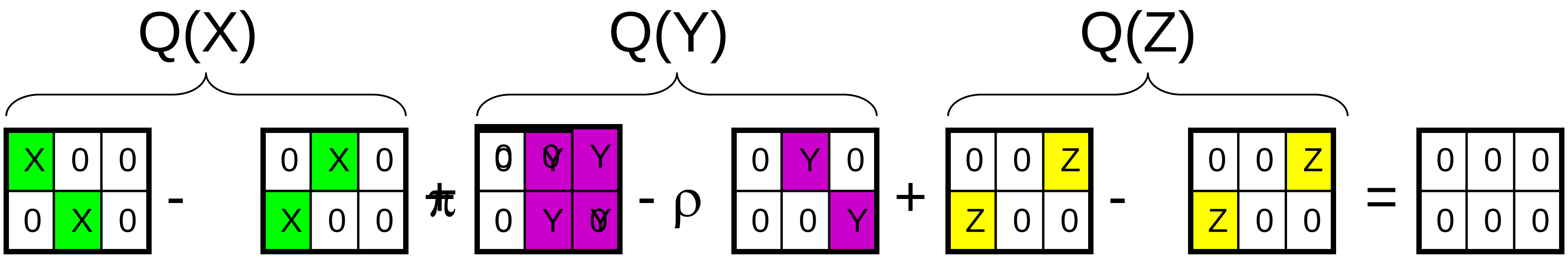
- Group is $(A^3)^2$ semidirect S_2
- π, ρ, τ 2 S_2 either "flip" or "no flip"
- must be even number of flips
  - CASE 1: $\pi = \rho = \tau =$ "no flip"

Beating the sum of the cubes



- Group is $(A^3)^2$ semidirect S_2
- π, ρ, τ 2 S_2 either “flip” or “no flip”
- must be even number of flips
  - CASE 2: π = ρ = “flip”; τ = “no flip”

Beating the sum of the cubes



- Group is $(A^3)^2$ semidirect S_2
- π, ρ, τ 2 S_2 either “flip” or “no flip”
- must be even number of flips
  - – CASE 2: $\pi = \rho =$ “flip”; $\tau =$ “no flip”
  - – contradiction.

Beating the sum of the cubes

G semidirect product of $(A^3)^2$ and S_2

X = { [matrix below] π : π 2 S_2}

X	0	0
0	X'	0

Y = { [matrix below] π : π 2 S_2}

0	Y	0
0	0	Y'

Z = { [matrix below] π : π 2 S_2}

0	0	Z
Z'	0	0

|A| = 17 yields
$\omega < 2.908\ldots$

- $|X||Y||Z| = 8(|A|-1)^6$ **>**
- $\sum_i d_i^3 \cdot d_{max} \sum_i d_i^2 = 2|G| = 4|A|^6$

generalizing
the construction
via
Uniquely Solvable Puzzles

Three subsets of $(A^w)^N$ semidirect S_N:

$$X = \left\{ \begin{bmatrix} X & X & 0 & 0 & 0 & 0 \\ X & 0 & X & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 \\ 0 & X & X & 0 & 0 & 0 \\ 0 & X & 0 & 0 & 0 & 0 \\ 0 & 0 & X & 0 & 0 & 0 \end{bmatrix} \pi \right\}$$

$$Y = \left\{ \begin{bmatrix} 0 & 0 & Y & Y & Y & 0 \\ 0 & Y & 0 & Y & 0 & Y \\ 0 & Y & Y & Y & 0 & 0 \\ Y & 0 & 0 & 0 & Y & Y \\ Y & 0 & Y & 0 & Y & 0 \\ Y & Y & 0 & 0 & 0 & Y \end{bmatrix} \pi \right\}$$

$$Z = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & Z \\ 0 & 0 & 0 & 0 & Z & 0 \\ 0 & 0 & 0 & 0 & Z & Z \\ 0 & 0 & 0 & Z & 0 & 0 \\ 0 & 0 & 0 & Z & 0 & Z \\ 0 & 0 & 0 & Z & Z & 0 \end{bmatrix} \pi \right\}$$

Want: partition of table that ensures Triple Product Property holds.

X = { table π }

X	X	0	0	0	0
X	0	X	0	0	0
X	0	0	0	0	0
0	X	X	0	0	0
0	X	0	0	0	0
0	0	X	0	0	0

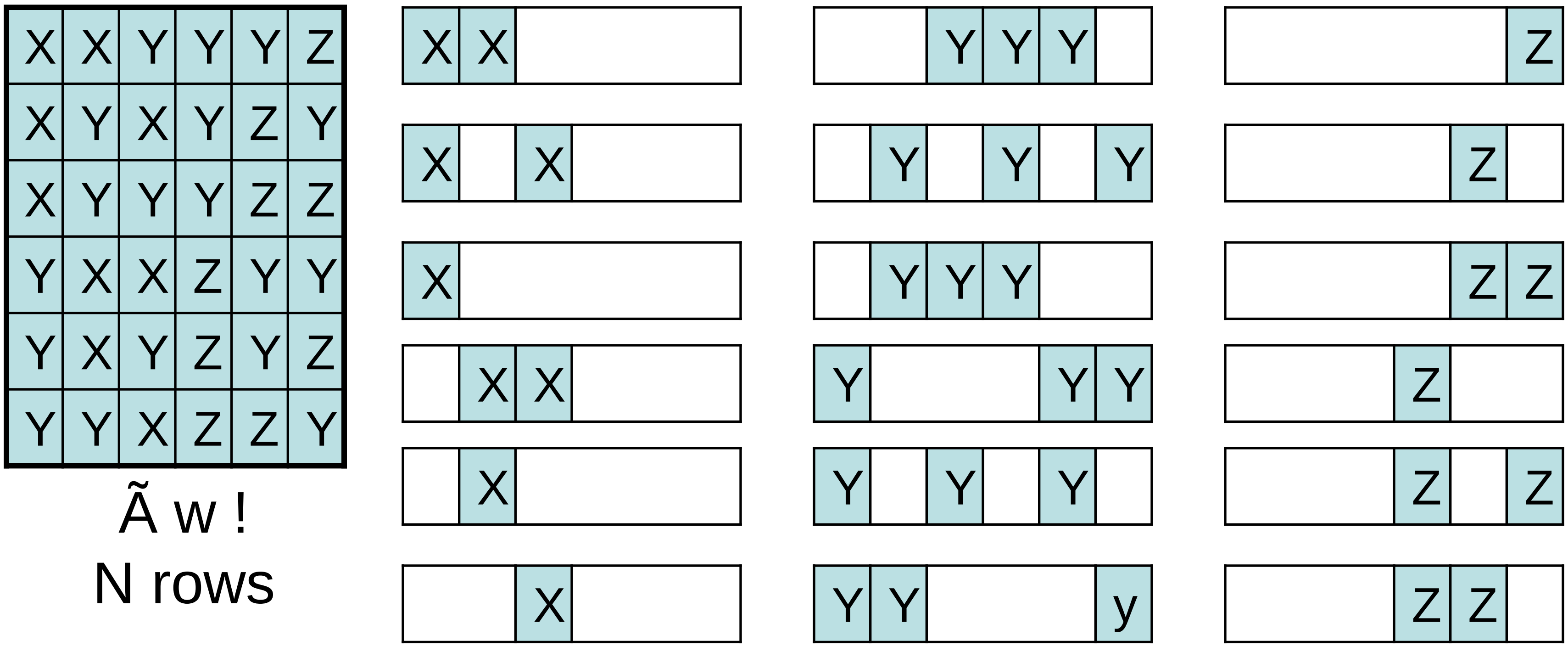
Y = { table π }

0	0	Y	Y	Y	0
0	Y	0	Y	0	Y
0	Y	Y	Y	0	0
Y	0	0	0	Y	Y
Y	0	Y	0	Y	0
Y	Y	0	0	0	Y

Z = { table π }

0	0	0	0	0	Z
0	0	0	0	Z	0
0	0	0	0	Z	Z
0	0	0	Z	0	0
0	0	0	Z	0	Z
0	0	0	Z	Z	0

Uniquely Solvable Puzzle (USP)



- <u>**USP**</u>: unique way to assemble "puzzle pieces" without overlap (implies no duplicate pieces)

Uniquely Solvable Puzzle (USP)

- **Goal**: maximize N as function of w

- No duplicate pieces implies:

  N · (w choose w/3)

- Coppersmith/Winograd: USPs with N = (w choose w/3)$^{1 - o(1)}$ exist

X	X	Y	Y	Y	Z
X	Y	X	Y	Z	Y
X	Y	Y	Y	Z	Z
Y	X	X	Z	Y	Y
Y	X	Y	Z	Y	Z
Y	Y	X	Z	Z	Y

Ã w !

N rows

Uniquely Solvable Puzzle (USP)

USP: every *unintended* way of assembling pieces produces

~~overlap of 2 or 3 symbols in~~ some cell

overlap of exactly 2 symbols in some cell

X	X	Y	Y	Y	Z
X	Y	X	Y	Z	Y
X	Y	Y	Y	Z	Z
Y	X	X	Z	Y	Y
Y	X	Y	Z	Y	Z
Y	Y	X	Z	Z	Y

Ã w !

N rows

Conjecture: Strong USPs with

$N = (w \text{ choose } w/3)^{1 - o(1)}$ rows exist.

a different generalization via the Two Families Conjecture

Three subsets of
$(H^3)^N$ semidirect S_N:

$X = \{$

A_1	B_1	0
A_2	B_2	0
A_3	B_3	0
A_4	B_4	0
□	□	□
A_n	B_n	0

$\pi \}$

$Y = \{$

0	A_1	B_1
0	A_2	B_2
0	A_3	B_3
0	A_4	B_4
□	□	□
0	A_n	B_n

$\pi \}$

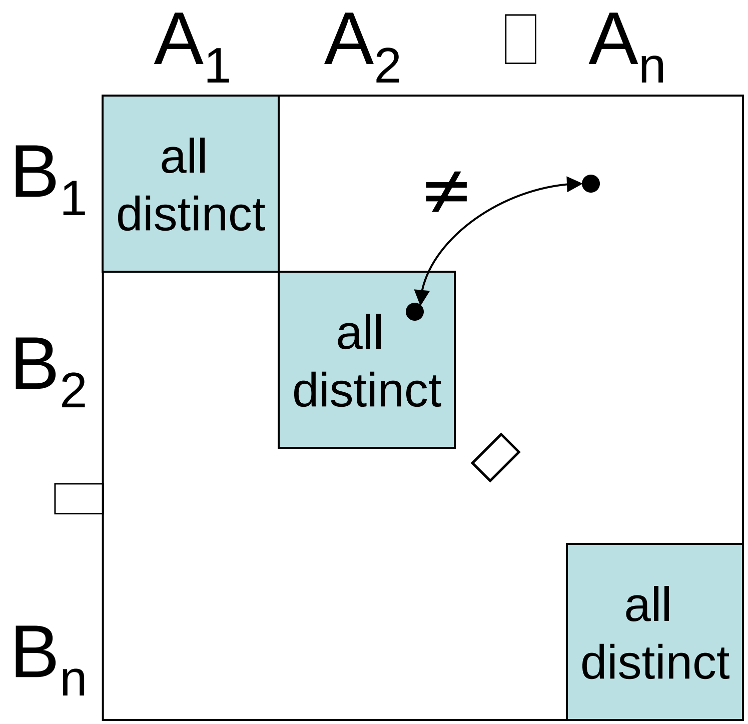
$Z = \{$

B_1	0	A_1
B_2	0	A_2
B_3	0	A_3
B_4	0	A_4
□	□	□
B_n	0	A_n

$\pi \}$

Sufficient condition for Triple Product Property

- subsets $A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_n$ of Abelian group H



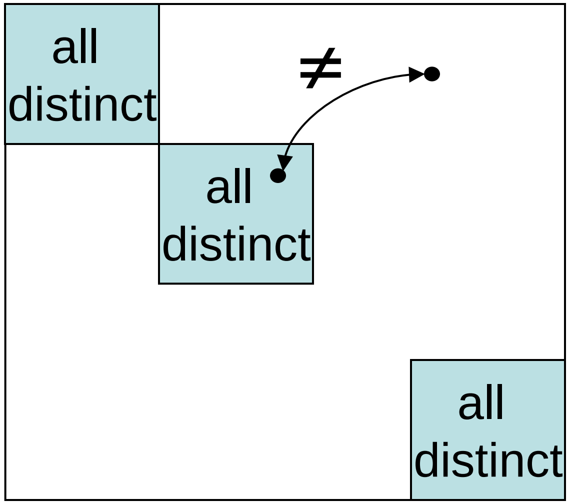
1. $|A_i + B_i| = |A_i|$¢$|B_i|$

2. $(A_i + B_i)$ Å $(A_j + B_k)$ = ; if $j \neq k$

Example

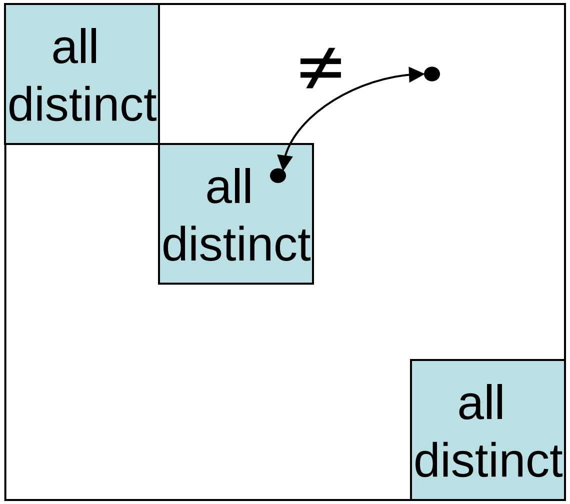
- $H = A^k$ for Abelian group A
  - $A_1 = (\neq 0, \neq 0, \neq 0, 0, 0, 0)$
  - $B_1 = (0, 0, 0, \neq 0, \neq 0, \neq 0)$

  - $A_2 = (\neq 0, \neq 0, 0, \neq 0, 0, 0)$
  - $B_2 = (0, 0, \neq 0, 0, \neq 0, \neq 0)$
  - … all equipartitions of k coords



- $|A_i + B_i| = |A_i|$¢$|B_i|$ for all i
- Elements of $(A_i + B_i)$ have nonzero in every coordinate
- Elements of $(A_j + B_k)$ have zero in some coordinate

Parameters



- **Optimal**:
  - $n = |H|^{\mathbf{1/2 - o(1)}}$
  - $|A_i| = |B_i| = |H|^{\mathbf{1/2 - o(1)}}$
  - yields $\omega = 2$
- Best construction we know (on previous slide):
  - $n = |H|^{0.3868\ldots}$
  - $|A_i| = |B_i| = |H|^{0.4491\ldots}$
  - yields $\omega \cdot 2.48\ldots$

Conjecture: exists optimal construction

Outline

Lecture I:

- crash course on main ideas from Strassen 1969 through Le Gall 2014
- conjectures implying ! = 2

Lecture II:

- group-theoretic approach

Tomorrow: extending to coherent configurations