Matrix multiplication via ~~Lie~~ matrix groups

Chris Umans

Caltech

Collaborators: Jonah Blasiak, Henry Cohn, Josh Grochow, Kevin Pratt

Introduction

A X B = C

- 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

The Group Algebra

- Given a finite group G
- The group algebra C[G] has elements

$$\sum_g a_g g$$

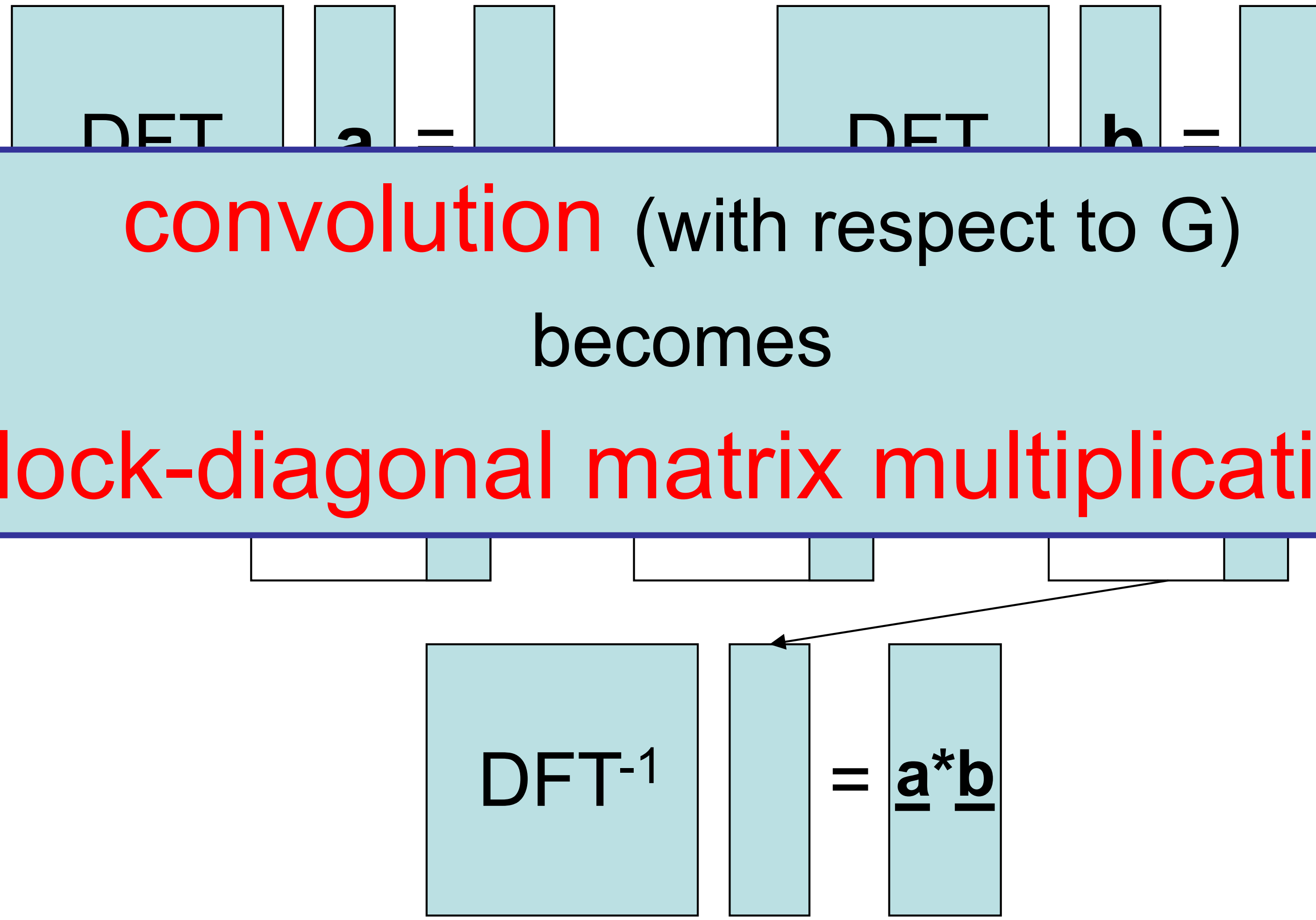
write as a vector in C^G

with multiplication

$$\left(\sum_g a_g g\right)\left(\sum_h b_h h\right) = \sum_f \left(\sum_{gh=f} a_g b_h\right) f$$

Multiplication in Group Algebra

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



convolution (with respect to G)

becomes

block-diagonal matrix multiplication

The basic idea: a reduction

Find a group G that permits an embedding

matrix A → **<u>A</u>** ∈ C[G], matrix B → **<u>B</u>** ∈ C[G]

so that we can read off entries of AB from

<u>A</u>*<u>B</u>

The embedding:

Subgroups X, Y, Z of G satisfy the

triple product property (TPP)

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

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

A = $\sum_{x,y} A[x,y](xy^{-1})$

B = $\sum_{y,z} B[y,z](yz^{-1})$

(AB)[x,z] = coefficient on xz^{-1} in A·B

The embedding:

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

Subsets X, Y, Z of G satisfy the

triple product property (TPP)

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

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

$\underline{A} = \sum_{x,y} A[x,y](xy^{-1})$

$\underline{B} = \sum_{y,z} B[y,z](yz^{-1})$

(AB)[x,z] = coefficient on xz^{-1} in $\underline{A} \cdot \underline{B}$

Character degrees

- if |X|=|Y|=|Z|=k, this is *reduction* from k × k mat. mult. to block-diagonal mat. mult.

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

$$k^\omega \leq \sum_i d_i^\omega$$

need $k > d_{\max}$
and $k \approx |G|^{1/2}$

- Usually use: $k^\omega \leq d_{\max}^{\omega-2} \cdot |G|$

If $d_{\max} \approx |G|^{1/2}$, prove nothing until prove $\omega = 2$.

Which groups can prove $\omega = 2$?

- no abelian group
- no group G with $|G|^{\epsilon}$ -size abelian normal subgroup with bounded exponent [BCCGNSU 2017]
- no group G with with $|G|^{\epsilon}$ -size normal p-subgroup with mild extra conditions [BCCGU 2017]
- simple groups may be good candidates
  - no 3 Young subgroups in alt. group [BCCGU 2017]
  - this work: matrix groups

Matrix groups

- GL(n, F), SL(n, F)
  - F can be finite, or **C**, **R**
  - also orthogonal, unitary, symplectic...
- These groups, and nice subgroups of them, have a notion of dimension:
  - e.g. dim of GL_n is n^2, dim of subgroup of lower-unitriangular matrices is $(n^2 - n)/2$

Recall TPP goal: subgroups of sqrt size
⇔ subgroups of half dimension

Key relaxation: continuous setting

- We will use matrix groups over **R**
  - "sum of squares = 0 $\Rightarrow$ each summand = 0" is powerful and enables good constructions
  - First challenge: obtain an analog of $\omega = 2$

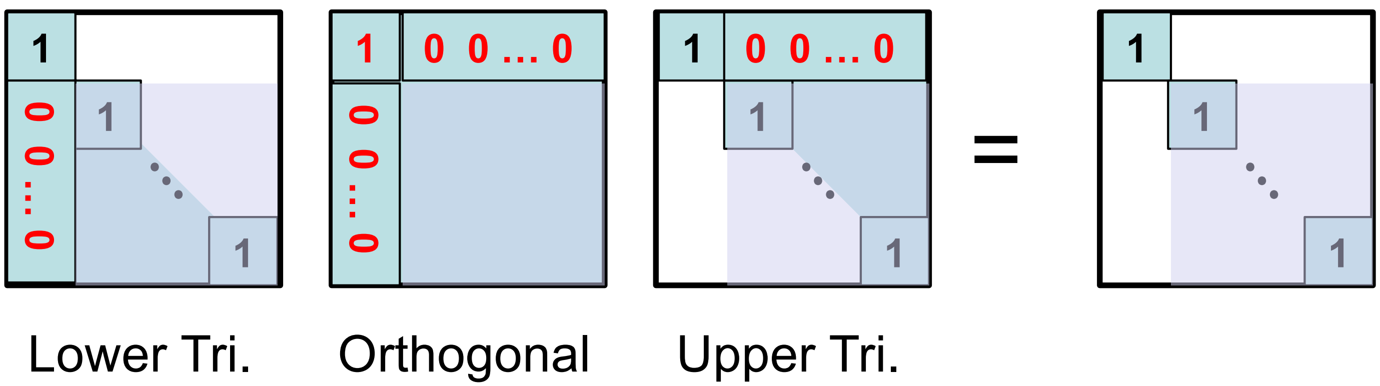
> In a matrix group over **R**, can we get TPP with X, Y, Z, having 1/2 the dimension ?

  - Later: a way to get *bona fide* matrix mult. algorithms from such constructions

TPP in Lie groups with subgroups of ½ the dimension

Example construction

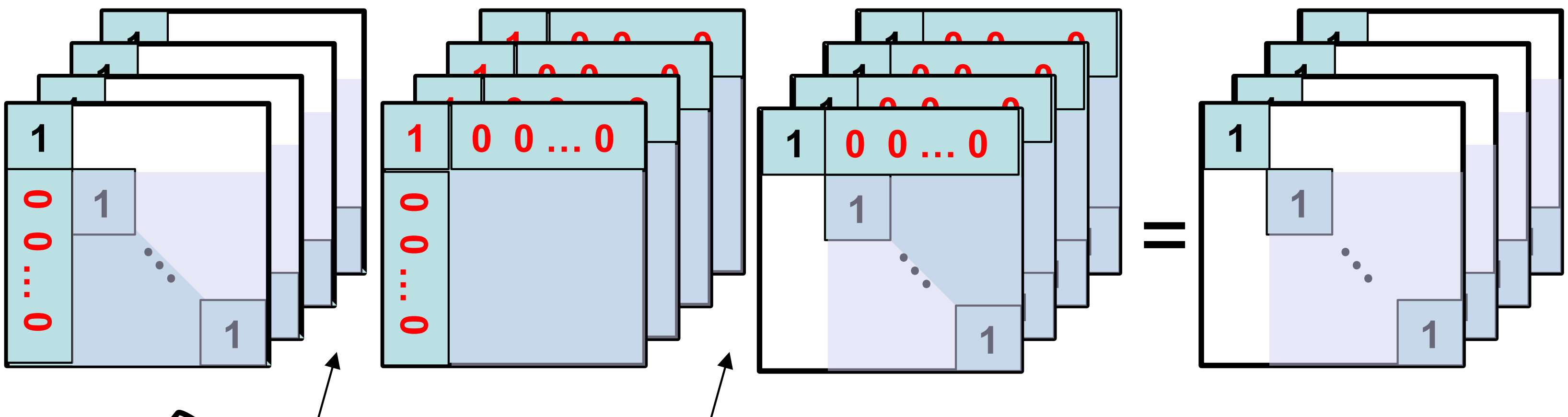
- Three subgroups in GL(n, **R**):
  - lower uni-triangular, orthogonal, upper uni-tri.

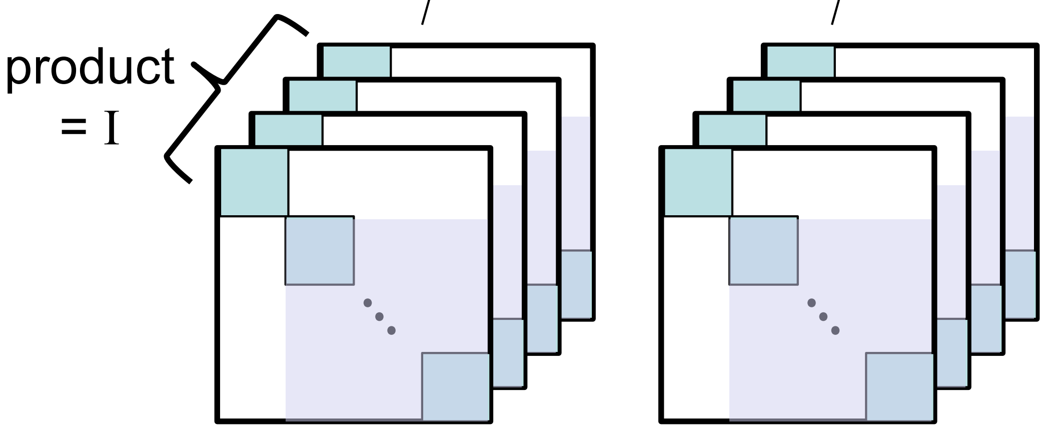


dimensions $\frac{n^2-n}{2}$ in group of dimension n^2

Construction achieving ½ dim

- Three subsets in GL(n, **R**)^k:





Subgroup D

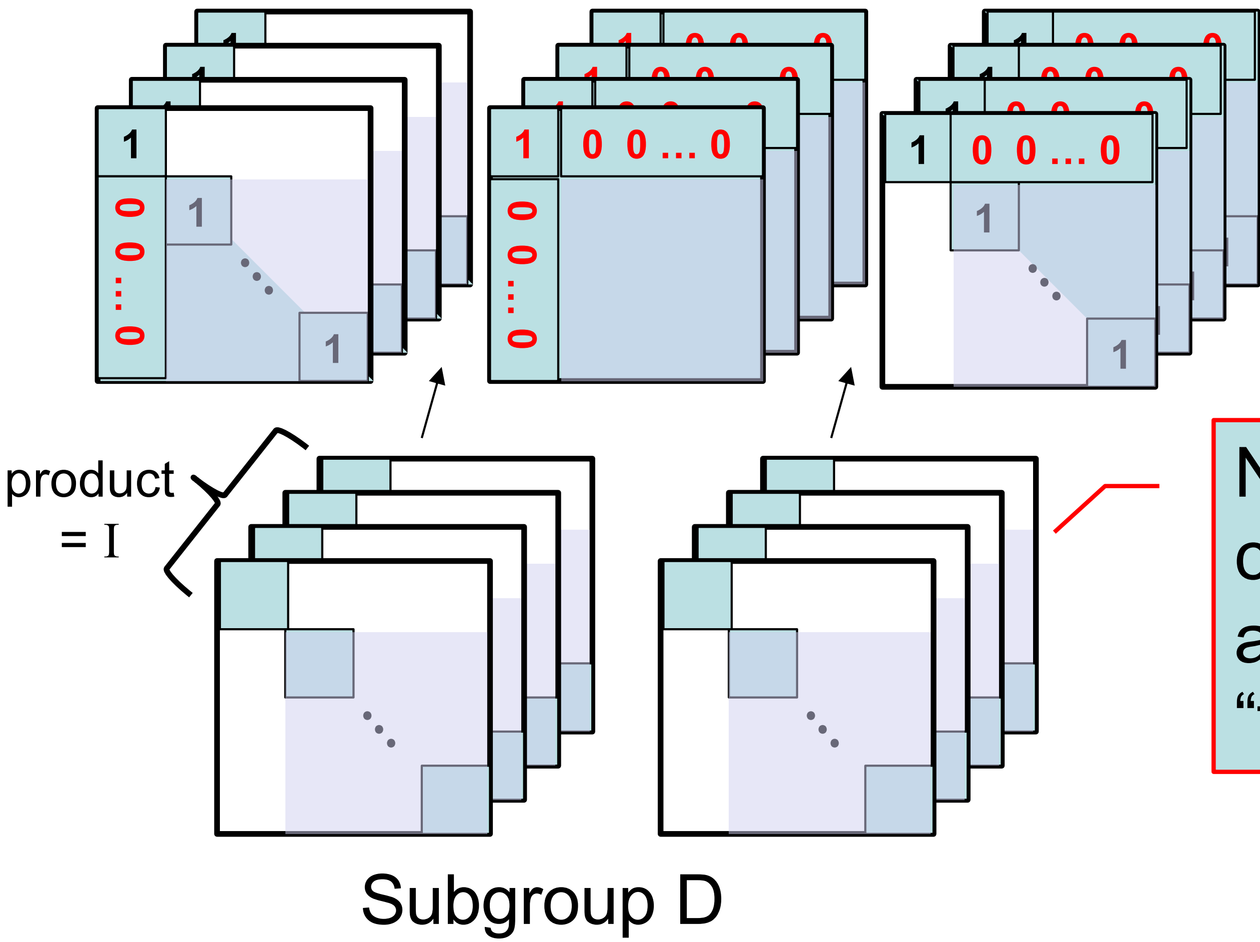
$(\text{Lower Tri})^k \cdot D$

$(\text{Orthog})^k \cdot D$

$(\text{Upper Tri})^k \cdot D$

Construction achieving ½ dim

- Three subsets in $GL(n, \mathbf{R})^k$:



Not able to conclude that these are the identity: "failure at diagonal"

Dimensions of construction

G = GL(n, **R**) $\{(D_1, \ldots, D_k) : \prod_i D_i = I\} = D \subseteq \mathrm{G}^k$

	Lower Tri.	Orthog.	Upper Tri.	G
dim:	$(n^2 - n)/2$	$(n^2 - n)/2$	$(n^2 - n)/2$	n^2
	$(\mathrm{LT})^k \cdot D$	$(\mathrm{Orth})^k \cdot D$	$(\mathrm{UT})^k \cdot D$	G^k
dim/k:	$(n^2 - n)/2 + n - o_k(1)$	$(n^2 - n)/2 + n - o_k(1)$	$(n^2 - n)/2 + n - o_k(1)$	n^2

Fixing "failure at diagonal"

G = GL(n, **R**) $\{(D_1, \ldots, D_k) : \prod_i D_i = I\} = D \subseteq G^k$

$H = \{M \in G : Mv = v\}$ for v = all-ones vector

key: $D \cap H^k$ = {identity}

	$(\mathrm{LT})^k \cdot D \cap H^k$	$(\mathrm{Orth})^k \cdot D \cap H^k$	$(\mathrm{UT})^k \cdot D \cap H^k$	H^k
dim/k:	$(n^2 - n)/2 + n - o_k(1) - n$	$(n^2 - n)/2 + n - o_k(1) - n$	$(n^2 - n)/2 + n - o_k(1) - n$	$n^2 - n$

Success! But… **Thm** [BCGPU23]: no analog in $GL(n, F_q)$.

Obtaining bounds on ω from Lie group constructions

Original framework: computing AB

- Given X, Y, Z in finite G, satisfying TPP:
  - for each irrep $\rho: \mathrm{G} \to C^{d \times d}$ compute:

$$\rho\left(\Sigma_{x,y} A[x,y](xy^{-1})\right) \cdot \rho\left(\Sigma_{y',z} B[y',z](y'z^{-1})\right)$$
$$= \Sigma_{x,y,y',z} A[x,y]B[y',z]\,\rho(xy^{-1}y'z^{-1})$$

  - the $\rho_{i,j}: G \to C$ form a basis for *all* $f: G \to C$.
  - "read off AB[x,z]" means take the linear combination for fn. f that is 1 only on xz^{-1}

New framework for Lie groups

- Given finite subsets $\mathbf{X} \subseteq X, \mathbf{Y} \subseteq Y, \mathbf{Z} \subseteq Z$ in Lie group G, satisfying TPP:
  - for *some* irreps $\rho\colon \mathrm{G} \to C^{d\times d}$ compute

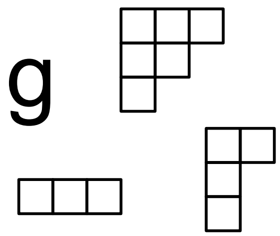
$$\rho\left(\Sigma_{x,y} A[x,y](xy^{-1})\right) \cdot \rho\left(\Sigma_{y',z} B[y',z](y'z^{-1})\right)$$
$$= \Sigma_{x,y,y',z} A[x,y]B[y',z]\, \rho(xy^{-1}y'z^{-1})$$

  - to "read off AB[x,z]" find linear combo of $\rho_{i,j}$ equal to f(M) = 1 if $M = xz^{-1}$

    0 if $M =$ any other $xy^{-1}y'z^{-1}$

Separating polynomials

- Irreps of GL(n, **R**) indexed by Young diagrams.



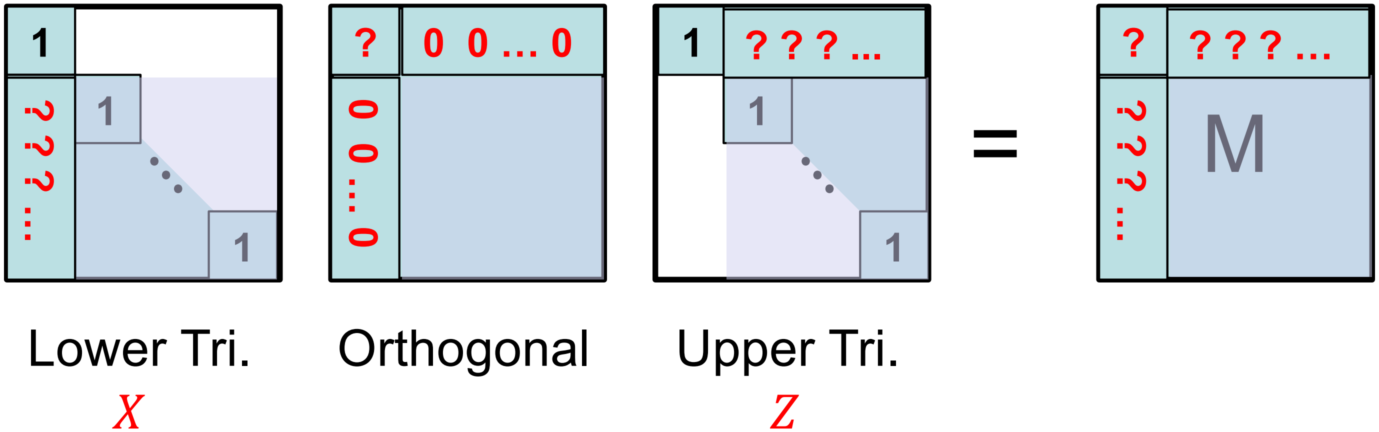
  - the $\rho_{i,j}$ for irreps up to size D span exactly the set of total-degree D polynomials
  - cut off at size D; now to “read off AB[x,z]”:
  - find “separating polynomial of deg D”:

$f_{x,z}$(M) = 1 if $M = xz^{-1}$

0 if $M =$ any other x$y^{-1}y'z^{-1}$

Separating polynomials example

- Three subgroups in GL(n, **R**):



$$f_{X,Z}(M) = \delta_1(M[1,1])$$
$$\cdot\, \delta_{(z_1,z_2,\ldots)}(M[\textbf{top row}])$$
$$\cdot\, \delta_{(x_1,x_2,\ldots)}(M[\textbf{left col}]) \quad \ldots$$

ind. degree equals # possible values in each entry of M

Separating polynomials

- Given finite subsets $\mathbf{X} \subseteq X, \mathbf{Y} \subseteq Y, \mathbf{Z} \subseteq Z$ in Lie group G, satisfying TPP:
  - each of size $\mathrm{q}^{\text{dim of subgroup}}$
  - separating polynomials of total degree $O(q)$

target degree

(example on previous slide: degree $\mathrm{O}(q^2)$)

yields same inequality on ω we would get if group was GL(n, F_q); if subgroups are ½ the ambient dimension then $\omega = 2$

Two ideas for designing separating polynomials

Setup so far

- X, Y, Z subgroups in Lie group G satisfying the Triple Product Property
- design finite subsets $\mathbf{X} \subseteq X, \mathbf{Y} \subseteq Y, \mathbf{Z} \subseteq Z$
  - each of size $\mathrm{q}^{\text{dim of subgroup}}$
- design separating polynomials of deg $O(q)$
  - argument M = $xy^{-1}y'z^{-1}$
  - poly

$f_{x,z}$(M) = 1 if $M = xz^{-1}$

= 0 if M = any other $xy^{-1}y'z^{-1}$

Setup so far

- design finite subsets $\mathbf{X} \subseteq X, \mathbf{Y} \subseteq Y, \mathbf{Z} \subseteq Z$
  - each of size $\mathrm{q}^{\text{dim of subgroup}}$
- design separating polynomials of deg $O(q)$
  - argument M = $xy^{-1}y'z^{-1}$
  - poly

$f_{x,z}$(M) = 1 if $M = xz^{-1}$
= 0 if M = any other $xy^{-1}y'z^{-1}$

Idea #1:

design $f_0(xy^{-1}y'z^{-1})$ = 1 if $y^{-1}y' = I$
= 0 if $y^{-1}y' \neq I$

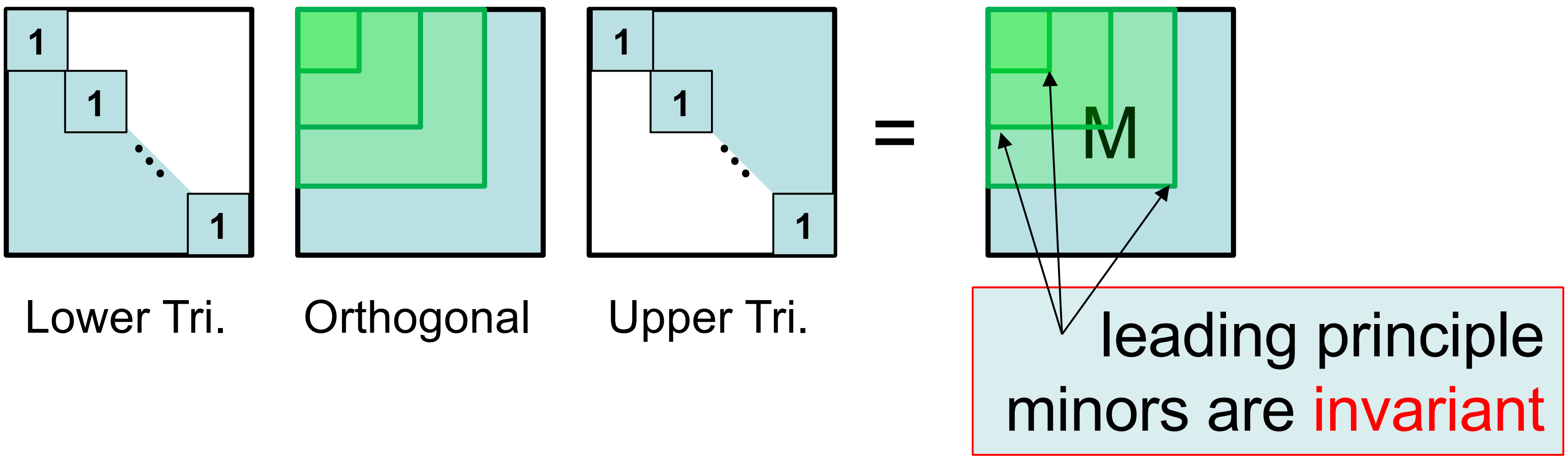
Invariant polynomials

Select f_0 from ring of invariant polynomials

– under left-multiplication by X

– under right-multiplication by Z

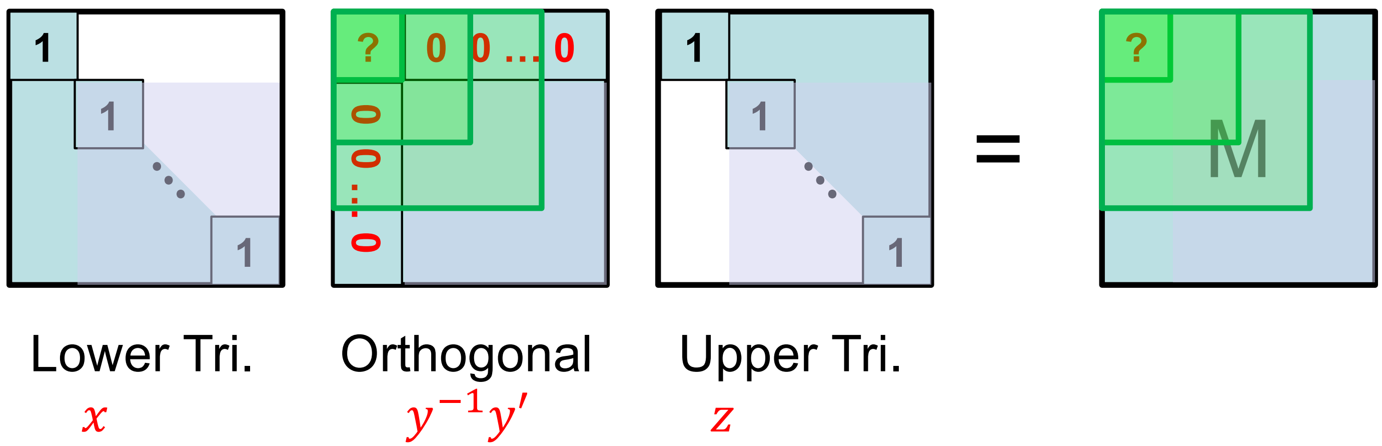
- Example: subgroups in GL(n, **R**)



Invariant polynomials

- subgroups in GL(n, **R**):

leading principle minors are invariant

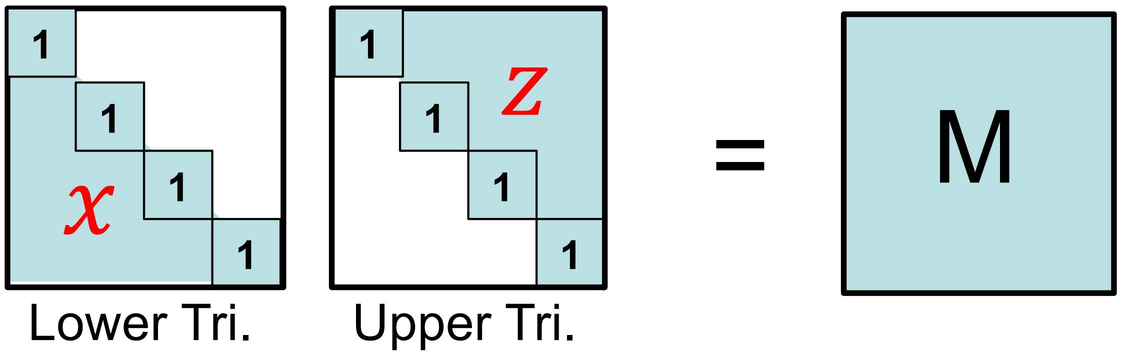


$$f_0(M) = \delta_1(lpm_1(M)) \cdot \delta_1(lpm_2(M)) \cdots$$

<u>**Claim**</u>: $f_0(xy^{-1}y'z^{-1}) = f_0(y^{-1}y') = 1$ implies $y^{-1}y' = I$

Remaining task:

- finite subsets of *2 subgroups* in GL(n, **R**):



  - find "separating polynomials" (to be multiplied with f_0)

$f_{x,z}$(M) = 1 if $M = xz^{-1}$

0 if $M =$ any other $x'z'^{-1}$

q values in entries of x, z ⇒ $O(q^2)$ values in entries of M

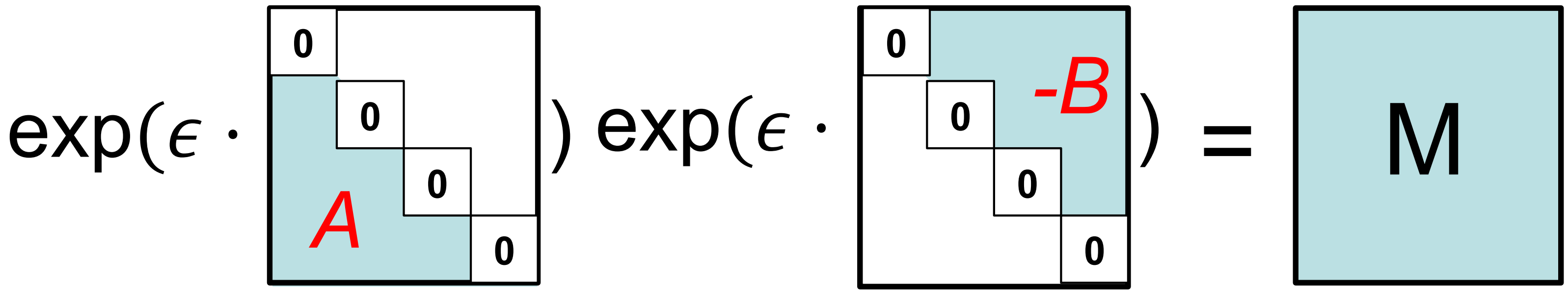
Idea #2: use Lie algebra

- Lie Group G has associated Lie Algebra **g**
  - **g** is a vectorspace
  - for any $A \in$ **g**, we have $exp(\epsilon A) \in G$

  (e.g. Orthogonal Group ⇒ skew-symmetric matrices)

- finite subsets of X, Y, Z can be defined via finite subsets of associated Lie algebras
  - the ϵ means the matrices have ϵ's in their entries, and irreps have ϵ's in their entries
  - final bound is on border-rank rather than rank!

Remaining task now easier



$$M = I + \epsilon(A - B) + O(\epsilon^2)$$

- choose entries of A, B in $\{0,1,2,\dots,q\}$
- now, only $O(q)$ values in $(M - I)/\epsilon$, *up to* $O(\epsilon)$
- separating polynomials of deg. $O(q)$:

  $f_{x,z}$(M) = $h_{A,B}\left(\frac{M-I}{\epsilon}\right)$, where

  $h_{A,B}(M') = 1$ if $M' = A - B$; otherwise 0

Lie algebra trick works in general

- Lie subgroups X, Y, Z that satisfy the TPP, with Lie algebras **x**, **y**, **z** (note: **x** ∩ **z** $=\ \{0\}$)
- fix a basis for **x**, **z**
- Choose finite subsets:
  - **X** = {exp(ϵA) : $A \in$ **x**, coefficients in $\{1 \ldots\ q\}$ }
  - **Z** = {exp(ϵB) : $B \in$ **z**, coefficients in $\{1 \ldots\ q\}$ }

$$M = I + \epsilon(A - B) + O(\epsilon^2)$$

  - $O(q)$ values per coefficient in $(M - I)/\epsilon$, *up to* $O(\epsilon)$

Putting it all together

- X, Y, Z subgroups in Lie group G satisfying the Triple Product Property
- determine the ring of polynomials invariant under left-mult. by X, right-mult by Z
- design subset $\mathbf{Y} \subseteq Y$ of size $\mathrm{q}^{\text{dim of subgroup}}$
- design sep. polynomial in ring, of deg $O(q)$

$f_0(y^{-1}y') = 1$ if $y^{-1}y' = I$
$= 0$ if $y^{-1}y' \neq I$

subgroups ½ the ambient dimension $\Rightarrow \omega = 2.$

Conclusions

- We know of two other constructions. Both come with separating polynomials, currently degree $O(q^2)$ rather than $O(q)$

- Open: find a construction that achieves TPP with half-dimensional subgroups, and finite subsets with separating polynomials having degree $O(q)$. Then $\omega = 2$.

Thank you!

So far…

- Achieved goal of TPP construction with subgroups half the dimension
  - if in GL(n, **R**), would imply a precise analog of $\omega = 2$ in the sense that if the construction moved to GL(n, F_q) it would prove $\omega = 2$.
  - but in Aff(n, **R**), no: Aff(n, F_q) has $d_{max} \approx q^{dim/2}$ instead of $q^{\text{bounded away from dim/2}}$

<u>Challenge</u>: as-good construction in GL(n, R).