Maximal border subrank tensors

Chia-Yu Chang

Texas A&M University

IPAM, February 6
Notations

A, B, C: n-dimensional vector spaces over \( \mathbb{C} \)

\( \{a_i\}, \{b_i\}, \{c_i\} \): bases of A, B, C, respectively

\( \{\alpha_i\}, \{\beta_i\}, \{\gamma_i\} \): dual bases

\( \{e_i\} \): the standard basis of \( \mathbb{C}^s \), \( s \in \mathbb{N} \)

\( \langle s \rangle := \sum_{i=1}^{s} e_i \otimes e_i \otimes e_i \in \mathbb{C}^s \otimes \mathbb{C}^s \otimes \mathbb{C}^s \): the **unit tensor** of size \( s \)

A tensor \( T \in A \otimes B \otimes C \) can be viewed as a linear map \( T_A : A^* \to B \otimes C \).

Similarly, we have \( T_B \) and \( T_C \).

We say a tensor \( T \in A \otimes B \otimes C \) is **concise** if \( T_A \), \( T_B \), and \( T_C \) are injective.

In particular, the unit tensor \( \langle s \rangle \) is concise in \( \mathbb{C}^s \otimes \mathbb{C}^s \otimes \mathbb{C}^s \).
Definition

Let $T \in A \otimes B \otimes C$.

The **rank** of $T$, $R(T)$, is the minimal positive integer $r$ such that

$$T \in (\text{Hom}(\mathbb{C}^r, A) \times \text{Hom}(\mathbb{C}^r, B) \times \text{Hom}(\mathbb{C}^r, C)) \cdot \langle r \rangle.$$ 

The **border rank** of $T$, $R_b(T)$, is the minimal positive integer $r$ such that

$$T \in (\text{Hom}(\mathbb{C}^r, A) \times \text{Hom}(\mathbb{C}^r, B) \times \text{Hom}(\mathbb{C}^r, C)) \cdot \langle r \rangle.$$ 

The **subrank** of $T$, $Q(T)$, is the maximal positive integer $s$ such that

$$\langle s \rangle \in (\text{Hom}(A, \mathbb{C}^s) \times \text{Hom}(B, \mathbb{C}^s) \times \text{Hom}(C, \mathbb{C}^s)) \cdot T.$$ 

The **border subrank** of $T$, $Q_b(T)$, is the maximal positive integer $s$ such that

$$\langle s \rangle \in (\text{Hom}(A, \mathbb{C}^s) \times \text{Hom}(B, \mathbb{C}^s) \times \text{Hom}(C, \mathbb{C}^s)) \cdot T.$$
Properties of (Border) Rank and (Border) Subrank

\[ R(T) := \min \{ r : T \in (\text{Hom}(C^r, A) \times \text{Hom}(C^r, B) \times \text{Hom}(C^r, C)) \cdot \langle r \rangle \} \]

\[ \underline{R}(T) := \min \{ r : T \in (\text{Hom}(C^r, A) \times \text{Hom}(C^r, B) \times \text{Hom}(C^r, C)) \cdot \langle r \rangle \} \]

\[ Q(T) := \max \{ s : \langle s \rangle \in (\text{Hom}(A, C^s) \times \text{Hom}(B, C^s) \times \text{Hom}(C, C^s)) \cdot T \} \]

\[ \underline{Q}(T) := \max \{ s : \langle s \rangle \in (\text{Hom}(A, C^s) \times \text{Hom}(B, C^s) \times \text{Hom}(C, C^s)) \cdot T \} \]

For any \( T \in A \otimes B \otimes C \), we have

- \( Q(T) \leq \underline{Q}(T) \leq n \), where \( n = \dim(A) = \dim(B) = \dim(C) \)
  - \( T \) is of maximal (border) subrank if \( "= n" \)

- \( Q(T) \leq \underline{Q}(T) \leq \underline{R}(T) \leq R(T) \)

- \( Q(T) \leq \underline{Q}(T) \leq n \leq \underline{R}(T) \leq \underline{R}(T) \) if \( T \) is concise
Motivation from Complexity Theory

- The exponent of matrix multiplication is defined as
  \[
  \omega := \inf \{ h \in \mathbb{R} : R(M_{n,n,n}) = O(n^h) \},
  \]
  where \( M_{n,n,n} \) is the \( n \times n \times n \) matrix multiplication tensor.

- [Str69]: \( 2 \leq \omega \leq \log_2 7 < 2.81 < 3 \)

- A well-known method to find upper bounds on \( \omega \) is the laser method
  [Str87]: study an intermediate tensor \( T \) which is
  
  1. of small border rank (low cost)
  2. close to being a matrix multiplication tensor (high value)

- The intermediate tensors of large (asymptotic) subrank are good to get bounds for \( \omega \).
Motivation

For a generic tensor $T$, $R(T) = R(T) = \text{maximum border rank } \sim \frac{n^2}{3}$.

How about $Q(T)$ and $Q(T)$?

Unknown!

Theorem (Derksen, Makam, Zuiddam, 2022)

The generic subrank of tensors in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ has bounds

$$3(\lceil \sqrt{\frac{n}{3}} + \frac{1}{4} - \frac{1}{2} \rceil) \leq Q(n) \leq \lfloor \sqrt{3n} - 2 \rfloor.$$ 

In particular, the generic subrank is not maximal.

Proposition

The border subrank of generic tensors in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ is at most $n - 1$ for $n \geq 3$.

Main result today: A lower bound of the dimension of the set of maximal border subrank tensors
Maximal subrank tensors

View $\langle n \rangle = \sum_{i=1}^{n} a_i \otimes b_i \otimes c_i \in A \otimes B \otimes C$, since $A, B, C$: $n$-dimensional.

Note $Q(\langle n \rangle) = Q(\langle n \rangle) = n$.

**Proposition**

The orbit of the unit tensor $(\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)) \cdot \langle n \rangle$ consists of all maximal subrank tensors.

**Proof.**

If $Q(T) = n$, then there exist $X \in \text{End}(A)$, $Y \in \text{End}(B)$, and $Z \in \text{End}(C)$ such that

$$\langle n \rangle = (X \otimes Y \otimes Z) \cdot T \in \text{im}(X) \otimes \text{im}(Y) \otimes \text{im}(Z).$$

Since $\langle n \rangle$ is concise, we get that $X, Y, Z$ are invertible.
Maximal border subrank tensors

\[ Q(T) = n \text{ if and only if} \]
\[ \langle n \rangle \in (\text{End}(A) \times \text{End}(B) \times \text{End}(C)) \cdot T = (\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)) \cdot T \]

Define \( Q_{\text{max}} := \{ T \in A \otimes B \otimes C : Q(T) = n \} \).

Then \( (\text{GL}(A) \times \text{GL}(B) \times \text{GL}(C)) \cdot \langle n \rangle \subset Q_{\text{max}} \).

Write \( G = \text{GL}(A) \times \text{GL}(B) \times \text{GL}(C) \).

Hence, \( \dim(Q_{\text{max}}) \geq \dim(G) - \dim(G_{\langle n \rangle}) = 3n^2 - 2n \),

where \( G_{\langle n \rangle} := \{ g \in G : g \cdot \langle n \rangle = \langle n \rangle \} \) is the symmetry group of \( \langle n \rangle \).

**Main Theorem**

\[ \dim(Q_{\text{max}}) \geq \frac{2n^3 + 3n^2 - 2n}{3} \sim \frac{2}{3} n^3. \]
The nullcone by the symmetry group of the unit tensor

Define the nullcone \( \mathcal{N}_{G^{\langle n \rangle}} := \{ w \in A \otimes B \otimes C : 0 \in G^{\langle n \rangle} \cdot w \} \), and let

\[
\text{Cone}(\langle n \rangle, \mathcal{N}_{G^{\langle n \rangle}}) := \{ v + w : v \in \mathbb{C} \cdot \langle n \rangle \text{ and } w \in \mathcal{N}_{G^{\langle n \rangle}} \} \subset \mathbb{Q}_{\text{max}}.
\]

Then we have \( G \cdot \text{Cone}(\langle n \rangle, \mathcal{N}_{G^{\langle n \rangle}}) \subset \mathbb{Q}_{\text{max}}. \)

**Proposition**

The symmetry group of the unit tensor is \( G^{\langle n \rangle} = \mathfrak{S}_n \ltimes T \), where

\[
T := \{ (\lambda, \mu, \nu) \in G : \lambda, \mu, \nu: \text{diagonal, } \lambda \mu \nu = \text{Id}_n \}
\]

is a maximal torus and \( \mathfrak{S}_n \) is the symmetric group on \( n \) elements.

Let \( \mathcal{N}_T := \{ w \in A \otimes B \otimes C : 0 \in T \cdot w \} \subset \mathcal{N}_{G^{\langle n \rangle}}. \)

By Hilbert-Mumford criterion, \( \mathcal{N}_{G^{\langle n \rangle}} = G^{\langle n \rangle} \cdot \mathcal{N}_T. \)

Since \( T \) is normal in \( G^{\langle n \rangle} \), we have that \( \mathcal{N}_{G^{\langle n \rangle}} = \mathcal{N}_T. \)
The nullcone defined by the torus

Let \( x_{ijk} = \alpha_i \otimes \beta_j \otimes \gamma_k \). The coordinate ring of \( A \otimes B \otimes C \) is \( \mathbb{C}[x_{ijk}] \).

\[
\mathcal{N}_{G\langle n \rangle} = \mathcal{N}_T = \text{Zeros}(\{f \in \mathbb{C}[x_{ijk}] : f: \text{homogeneous}, \deg(f) > 0, g \cdot f = f \quad \forall g \in T\}) \\
= \text{Zeros}(\{f \in \mathbb{C}[x_{ijk}] : f: \text{monomials}, \deg(f) > 0, g \cdot f = f \quad \forall g \in T\})
\]
since \( T \) is a torus and monomials span the weight vectors.

Thus \( \mathcal{N}_{G\langle n \rangle} \) is a union of linear spaces.

In particular,

\[
\mathcal{N}_{G\langle n \rangle} \subset \text{Zeros}(\{x_{iii}, x_{ijj}, x_{iji}, x_{ijx}, x_{ijkx}, x_{jikx} : \text{distinct } 1 \leq i, j, k \leq n\}) \\
= \bigcup \text{of linear spaces with } \dim = n^3 - \left( n + 3 \binom{n}{2} + 2 \binom{n}{3} \right)
\]

\( W := \langle a_i \otimes b_j \otimes c_k : \text{at least one of } j, k \text{ less than } i \rangle \) is one of the linear spaces.
**Claim:** \( W \subset \mathcal{N}_{G_{\langle n \rangle}} := \{ w \in A \otimes B \otimes C : 0 \in G_{\langle n \rangle} \cdot w \} \)

**Proof of claim:** Let \( c(t) \in G_{\langle n \rangle} \) be defined as

\[
c(t) := \begin{pmatrix}
(t^{\lambda_1} & 0) \quad (t^{\mu_1} & 0) \quad (t^{\nu_1} & 0) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & t^{\lambda_n} & 0 & \cdots & t^{\mu_n} & 0 & \cdots & t^{\nu_n}
\end{pmatrix}
\]

where \( \lambda_k = 2^n - 2^{n-k+1} \) and \( \mu_k = \nu_k = 2^{n-k} - 2^{n-1} \) for \( k = 1, \ldots, n \).

Then

\[
c(t) \cdot (a_i \otimes b_j \otimes c_k) = t^{\lambda_i+\mu_j+\nu_k} (a_i \otimes b_j \otimes c_k)
\]

\[
= t^{2^{n-j}+2^{n-k} - 2^{n-i+1}} (a_i \otimes b_j \otimes c_k)
\]

which tends to zero as \( t \to 0 \) when \( a_i \otimes b_j \otimes c_k \in W \).
Thus we can focus on $W$ rather than on the whole nullcone $N_{G\langle n \rangle}$.

Consider the cone over $W$ with vertex $\langle n \rangle$.

$$\text{Cone}(\langle n \rangle, W) := \{ v + w : v \in \mathbb{C} \cdot \langle n \rangle \text{ and } w \in W \} \subset \text{Cone}(\langle n \rangle, N_{G\langle n \rangle}) \subset \mathbb{Q}_{\text{max}},$$

which has dimension

$$\text{dim}(W) + 1 = \frac{4n^3 - 3n^2 - n}{6} + 1.$$

The orbit closure of the cone $G \cdot \text{Cone}(\langle n \rangle, W)$ is also a subset of $\mathbb{Q}_{\text{max}}$. 

Dimension of the orbit closure of the cone

Let $v = \langle n \rangle + w \in \text{Cone}(\langle n \rangle, W)$ be a general point, where $w \in W$, and let $\text{Tran}_G(v, \text{Cone}(\langle n \rangle, W)) := \{ g \in G : g \cdot v \in \text{Cone}(\langle n \rangle, W) \}$.

The orbit closure of the cone $G \cdot \text{Cone}(\langle n \rangle, W)$ has dimension

$$\dim(G) + \dim(\text{Cone}(\langle n \rangle, W)) - \dim(\text{Tran}_G(v, \text{Cone}(\langle n \rangle, W)))$$

We can compute $\dim(\text{Tran}_G(v, \text{Cone}(\langle n \rangle, W)))$ by considering its tangent space at the identity element of $G$

$$\{(x, y, z) \in g : (x, y, z).(\langle n \rangle + w) \in \text{Cone}(\langle n \rangle, W)\},$$

where $g = \text{End}(A) \oplus \text{End}(B) \oplus \text{End}(C)$. 
Lower bound

The dimension of the tangent space of $\text{Tan}_G(\nu, \text{Cone}(\langle n \rangle, W))$ is

$$\frac{3n^2 + n + 2}{2}.$$

Thus the dimension of $G \cdot \text{Cone}(\langle n \rangle, W)$ is

$$\dim(G) + \dim(\text{Cone}(\langle n \rangle, W)) - \dim(\text{Tan}_G(\nu, \text{Cone}(\langle n \rangle, W)))$$

$$= 3n^2 + \frac{4n^3 - 3n^2 - n}{6} + 1 - \frac{3n^2 + n + 2}{2}$$

$$= \frac{2n^3 + 3n^2 - 2n}{3}.$$

This gives a lower bound of the dimension of $\mathbf{Q}_{\text{max}}$. 
Future Questions

(1) The nullcone $\mathcal{N}_{G_{\langle n \rangle}}$ contains the union of $W$ and its permutations. We have examples showing that the nullcone is larger, but not necessarily larger dimensional. What are all components of the nullcone?

(2) A tensor $T \in A \otimes B \otimes C$ is of maximal border subrank if the unit tensor $\langle n \rangle$ lies in the orbit closure $G \cdot \overline{T}$. Can we apply a one parameter subgroup of $G$ on the tensor $T$ to approach the unit tensor?

(3) Tensors in the orbit closure of the cone $\text{Cone}(\langle n \rangle, \mathcal{N}_{G_{\langle n \rangle}})$ are of maximal border subrank. Do all maximal border subrank tensors lie in the orbit closure of the cone? Do we have this if (2) is true?
Future Questions

(4) [Baiggi, C., Draisma, Rupniewski]: An upper bound on dimension of $Q_{\text{max}}$ is

$$n^3 - \lfloor n/3 \rfloor^3 + 6n^2.$$ 

Can we reduce the gap between the upper and the lower bound?

(5) For a generic tensor $T$, there are nontrivial upper and lower bounds on its subrank [DMZ22]. Can we find nontrivial upper or lower bounds on border subrank of a generic tensor?

(6) We have methods for finding upper bounds of subrank and border subrank that rely on other notion of rank, for example, geometric rank, slice rank, and $G$-stable rank. Can we develop a strategy to find the subrank and border subrank of a tensor or upper bounds on them that does not rely on other notions of rank?
Thank you!
Harm Derksen, Visu Makam, and Jeroen Zuiddam.
Subrank and optimal reduction of scalar multiplications to generic tensors.
In *37th Computational Complexity Conference*, volume 234 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 9, 23. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.

Volker Strassen.
Gaussian elimination is not optimal.
*Numer. Math.*, 13:354–356, 1969.

Volker Strassen.
Relative bilinear complexity and matrix multiplication.
*J. Reine Angew. Math.*, 375/376:406–443, 1987.