The Subrank of Random Tensors

Jeroen Zuiddam

University of Amsterdam

Joint work with Harm Derksen and Visu Makam.

- We solved a problem in tensor theory about a notion called the subrank of tensors.

- We solved a problem in tensor theory about a notion called the subrank of tensors.
- The subrank was introduced by Strassen in 1987 to study fast matrix multiplication algorithms in CS
- and has connections to several problems in math and physics.

- We solved a problem in tensor theory about a notion called the subrank of tensors.
- The subrank was introduced by Strassen in 1987 to study fast matrix multiplication algorithms in CS
- and has connections to several problems in math and physics.
- Our result: We determine the subrank for "random tensors"/ "almost all tensors"/ generic tensors
- Improve on previous bounds of Strassen & Bürgisser from 1991

1. Subrank and Applications

2. Tensor Parameters and Their Value on Random Tensors

3. Subrank of Random Tensors

4. Upper bound

5. Lower bound

6. Tensor Space Decomposition

ingredient

1. Subrank and Applications

Two characterizations of rank of a matrix $M \in \mathbb{F}^{n \times n}$

1. Subrank and Applications

Two characterizations of rank of a matrix $M \in \mathbb{F}^{n \times n}$

Decomposition into simple matrices

minimize

$$M = \sum_{i=1}^{r} u_i \otimes v_i$$

Equiv:

$$M = A \, I_r \, B$$

"Create matrix from identity"

1. Subrank and Applications

Two characterizations of rank of a matrix $M \in \mathbb{F}^{n \times n}$

Decomposition into simple matrices

minimize

$$M = \sum_{i=1}^{r} u_i \otimes v_i$$

Equiv:

$$M = A \, I_r \, B$$

"Create matrix from identity"

Gaussian elimination

$$A M B = I_r$$

maximize

"Create identity from matrix"

Two different notions of rank of a tensor $T \in \mathbb{F}^{n \times n \times n}$

Two different notions of rank of a tensor $T \in \mathbb{F}^{n \times n \times n}$

Tensor rank

minimize

$$T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$$

Equiv:

$$T = U \otimes V \otimes W \cdot \sum_{i=1}^{r} e_i \otimes e_i \otimes e_i$$

$R(T)$

Two different notions of rank of a tensor $T \in \mathbb{F}^{n \times n \times n}$

Tensor rank

minimize

$$T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i \qquad R(T)$$

Equiv:

$$T = U \otimes V \otimes W \cdot \sum_{i=1}^{r} e_i \otimes e_i \otimes e_i$$

Subrank

maximize

$$\sum_{i=1}^{s} e_i \otimes e_i \otimes e_i = U \otimes V \otimes W \cdot T \qquad Q(T)$$

Two different notions of rank of a tensor $T \in \mathbb{F}^{n \times n \times n}$

Tensor rank

minimize

$$T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$$

Equiv:

$$T = U \otimes V \otimes W \cdot \sum_{i=1}^{r} e_i \otimes e_i \otimes e_i$$

$R(T)$

Applications

- Matrix multiplication
- Circuit complexity [Raz]

Subrank

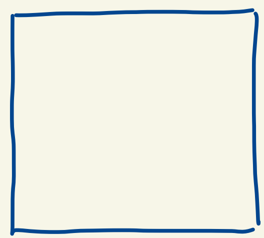
maximize

$$\sum_{i=1}^{s} e_i \otimes e_i \otimes e_i = U \otimes V \otimes W \cdot T$$

$Q(T)$

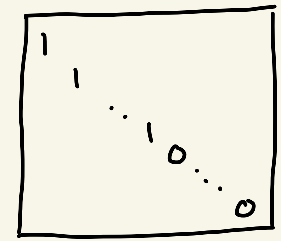
- Matrix Multiplication
- Additive Combinatorics

Matrix rank

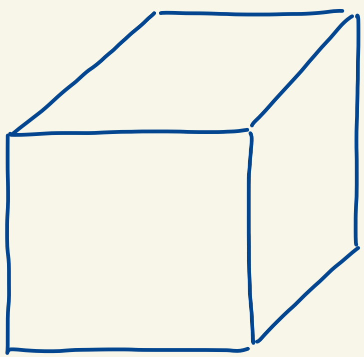
$M =$ 

linear combinations
of rows and columns

$\rightsquigarrow$

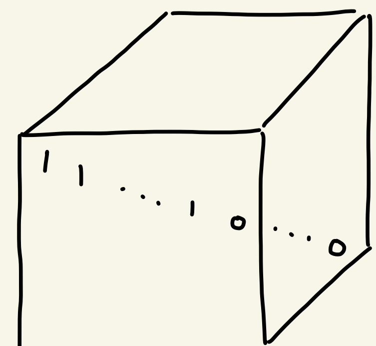


Subrank

$T =$ 

linear combinations
of slices in all
three directions

$\rightsquigarrow$



Applications of Subrank

- Complexity Theory

$T \in \mathbb{F}^{n \times n \times n} \quad \leftrightsquigarrow \quad$ bilinear map $T : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$

$Q(T) \quad \leftrightsquigarrow \quad$ number of independent scalar multiplications that can be reduced to T

Applications of Subrank

- Complexity Theory

$T \in \mathbb{F}^{n \times n \times n}$ $\rightsquigarrow$ bilinear map $T : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$

$Q(T)$ $\rightsquigarrow$ number of independent scalar multiplications that can be reduced to T

- Quantum Information

$T \in \mathbb{C}^{n \times n \times n}$ $\rightsquigarrow$ Tripartite quantum state

$Q(T)$ $\rightsquigarrow$ largest "GHZ" state obtainable from T by SLOCC

Applications of Subrank

- Complexity Theory

$T \in \mathbb{F}^{n \times n \times n}$ $\leftrightsquigarrow$ bilinear map $T : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$

$Q(T)$ $\leftrightsquigarrow$ number of independent scalar multiplications that can be reduced to T

- Quantum Information

$T \in \mathbb{C}^{n \times n \times n}$ $\leftrightsquigarrow$ Tripartite quantum state

$Q(T)$ $\leftrightsquigarrow$ largest "GHZ" state obtainable from T by SLOCC

- Combinatorics

$H \subseteq [n] \times [n] \times [n]$ hypergraph, independence number $\alpha(H) \leq Q(T)$ for any T that "fits" H. E.g. cap sets, sunflowers, corners, ...

2. Tensor Parameters and Their Value on Random Tensors

$T \in \mathbb{F}^{n \times n \times n}$

$$0 \leq Q(T) \qquad \leq \qquad R(T) \leq n^2$$

2. Tensor Parameters and Their Value on Random Tensors

$T \in \mathbb{F}^{n \times n \times n}$

$$0 \leq Q(T) \leq SR(T) \leq n \leq R(T) \leq n^2$$

$$AR(T)$$

$$GR(T)$$

$$R^G(T)$$

Slice rank

$$T = \sum_{i=1}^{a} \sum_{j} u_i \otimes v_{ij} \otimes w_{ij} + \sum_{i=1}^{b} \sum_{j} u'_{ij} \otimes v'_i \otimes w'_{ij}$$

$$+ \sum_{i=1}^{c} \sum_{j} u''_{ij} \otimes v''_{ij} \otimes w''_i$$

minimize $a+b+c$

$SR(T)$

Geometric rank

$$\operatorname{codim} \{ (u,v) \in \mathbb{F}^n \times \mathbb{F}^n : \forall w \ \ T(u,v,w) = 0 \}$$

$GR(T)$

$\vdots$

$T \in \mathbb{F}^{n \times n \times n}$

Generally :

$$0 \leq Q(T) \leq SR(T) \leq n \leq R(T) \leq n^2$$

$AR(T)$

$GR(T)$

$R^G(T)$

$T \in \mathbb{F}^{n \times n \times n}$

$$\begin{array}{lcccccccccc}
\text{Generally:} & 0 & \le & Q(T) & \le & SR(T) & \le & n & \le & R(T) & \le n^2 \\
 & & & & & AR(T) & & & & & \\
 & & & & & GR(T) & & & & & \\
 & & & & & R^{G}(T) & & & & & \\
\text{On random tensors } T: & & & \underbrace{\qquad}_{?} & & \underbrace{\qquad\qquad}_{\approx n} & & & & \underbrace{\qquad\qquad}_{\approx n^2} &
\end{array}$$

3. Subrank of Random Tensors

Theorem For almost all $T \in \mathbb{F}^{n \times n \times n}$ we have $Q(T) = \theta(\sqrt{n})$

Remarks:

- "Almost all" = "random" = generic
- that is, there is a nonempty Zarisk-open $\mathcal{U} \subseteq \mathbb{F}^{n \times n \times n}$ such that for all $T \in \mathcal{U}$ we have $Q(T) = \theta(\sqrt{n})$

3. Subrank of Random Tensors

Theorem For almost all $T \in \mathbb{F}^{n \times n \times n}$ we have $Q(T) = \theta(\sqrt{n})$

Remarks:

- "Almost all" = "random" = generic
- that is, there is a nonempty Zarisk-open $\mathcal{U} \subseteq \mathbb{F}^{n \times n \times n}$ such that for all $T \in \mathcal{U}$ we have $Q(T) = \theta(\sqrt{n})$
- Very precise bounds: $\sqrt{3n-2} - 5 \leq Q(T) \leq \sqrt{3n-2}$
- Previously: $Q(T) \leq n^{2/3 + o(1)}$

3. Subrank of Random Tensors

> **Theorem** For almost all $T \in \mathbb{F}^{n \times n \times n}$ we have $Q(T) = \theta(\sqrt{n})$

Remarks:

- "Almost all" = "random" = generic
- that is, there is a nonempty Zarisk-open $\mathcal{U} \subseteq \mathbb{F}^{n \times n \times n}$ such that for all $T \in \mathcal{U}$ we have $Q(T) = \theta(\sqrt{n})$
- Very precise bounds: $\sqrt{3n-2} - 5 \leq Q(T) \leq \sqrt{3n-2}$
- Previously: $Q(T) \leq n^{2/3 + o(1)}$
- Also for higher-order tensors
- Application: Subrank is not additive under direct sum.

Upper bound

$Q(n) :=$ subrank of a generic tensor in $\mathbb{F}^{n\times n\times n}$

To prove: $Q(n) \leq \sqrt{3n-2}$

Upper bound

$Q(n) :=$ subrank of a generic tensor in $\mathbb{F}^{n\times n\times n}$

To prove: $Q(n) \leq \sqrt{3n-2}$

$C_r := \{$ tensors in $\mathbb{F}^{n\times n\times n}$ with subrank $\geq r \}$

<u>Lemma 1</u> $Q(n) =$ largest r such that $\dim C_r = \underbrace{\dim \mathbb{F}^{n\times n\times n}}_{n^3}$.

Upper bound

$Q(n) :=$ subrank of a generic tensor in $\mathbb{F}^{n \times n \times n}$

To prove: $Q(n) \leq \sqrt{3n-2}$

$C_r := \{$ tensors in $\mathbb{F}^{n \times n \times n}$ with subrank $\geq r \}$

Lemma 1 $Q(n) =$ largest r such that $\dim C_r = \underbrace{\dim \mathbb{F}^{n \times n \times n}}_{n^3}$.

Lemma 2 $\dim C_r \leq n^3 - r(r^2 - 3n + 2)$

Upper bound

$Q(n) :=$ subrank of a generic tensor in $\mathbb{F}^{n \times n \times n}$

To prove: $Q(n) \leq \sqrt{3n-2}$

$C_r := \{$ tensors in $\mathbb{F}^{n \times n \times n}$ with subrank $\geq r \}$

<u>Lemma 1</u> $Q(n) =$ largest r such that $\dim C_r = \underbrace{\dim \mathbb{F}^{n \times n \times n}}_{n^3}$.

<u>Lemma 2</u> $\dim C_r \leq n^3 - r(r^2 - 3n + 2)$

Let $t = Q(n)$

Then $n^3 = \dim C_t \leq n^3 - t(t^2 - 3n + 2)$.

Then $t^2 - 3n + 2 \leq 0$

So $t \leq \sqrt{3n-2}$

$C_r :=$ { tensors in $\mathbb{F}^{n \times n \times n}$ with subrank $\geq r$ }

<u>Lemma 2</u> $\dim C_r \leq n^3 - r(r^2 - 3n + 2)$

$C_r := \{ \text{tensors in } \mathbb{F}^{n \times n \times n} \text{ with subrank} \geq r \}$

Lemma 2 $\dim C_r \leq n^3 - r(r^2 - 3n + 2)$

Proof idea

- Non-injective parametrization of C_r
- Compute dimension of parameter space
- Subtract dimension of "over-count" (fiber dimension)

$C_r := \{ \text{tensors in } \mathbb{F}^{n \times n \times n} \text{ with subrank} \geq r \}$

Lemma 2 $\dim C_r \leq n^3 - r(r^2 - 3n + 2)$

Proof idea

- Non-injective parametrization of C_r
- Compute dimension of parameter space
- Subtract dimension of "over-count" (fiber dimension)

$X_r = \{ \text{tensors in } \mathbb{F}^{n \times n \times n} \text{ with } [r] \times [r] \times [r] \text{ subtensor arbitrary diag.} \}$

$$\Psi_r : GL_n \times GL_n \times GL_n \times X_r \to \mathbb{F}^{n \times n \times n}$$
$$(A, B, C, T) \longmapsto (A \otimes B \otimes C) T \qquad \text{has image } C_r$$

$\square$

Lower bound

$X_r = \{$ tensors in $\mathbb{F}^{n \times n \times n}$ with $[r] \times [r] \times [r]$ subtensor arbitrary diag. $\}$

$\Psi_r : GL_n \times GL_n \times GL_n \times X_r \to \mathbb{F}^{n \times n \times n}$

$(A, B, C, T) \mapsto (A \otimes B \otimes C) T$ has image $\mathcal{C}_r$

Lower bound

$$X_r = \{ \text{tensors in } \mathbb{F}^{n\times n\times n} \text{ with } [r]\times[r]\times[r] \text{ subtensor arbitrary diag.} \}$$

$$\Psi_r : GL_n \times GL_n \times GL_n \times X_r \to \mathbb{F}^{n\times n\times n}$$

$$(A, B, C, T) \longmapsto (A\otimes B\otimes C)T \qquad \text{has image } C_r$$

Proof idea

- Find condition that imply image of Ψ_r has full dimension
- Use notion of differential $d\Psi_r$

Lower bound

$X_r = \{ \text{tensors in } \mathbb{F}^{n \times n \times n} \text{ with } [r] \times [r] \times [r] \text{ subtensor arbitrary diag.} \}$

$$\begin{aligned} \psi_r : \; & GL_n \times GL_n \times GL_n \times X_r \to \mathbb{F}^{n \times n \times n} \\ & (A, B, C, T) \mapsto (A \otimes B \otimes C) T \end{aligned}$$

has image C_r

Proof idea

- Find condition that imply image of ψ_r has full dimension
- Use notion of differential $d\psi_r$

tangent space of X_r

$$\begin{aligned} (d\psi_r)_{(g_1, g_2, g_3, T)} : \; & Mat_{n \times n} \times Mat_{n \times n} \times Mat_{n \times n} \times Y_r \to \mathbb{F}^{n \times n \times n} \\ & (A, B, C, T) \mapsto ((A \otimes g_2 \otimes g_3) + (g_1 \otimes B \otimes g_3) + (g_1 \otimes g_2 \otimes C)) T \\ & \qquad + (g_1 \otimes g_2 \otimes g_3) S. \end{aligned}$$

Tensor space decompositions

Goal: write tensor space $\mathbb{F}^{n \times n \times n}$ as a sum of tensor subspaces, as efficiently as possible such that each subspace has the form of an $n \times n$ matrix subspace tensored with $\mathbb{F}^n$

Tensor space decompositions

Goal: write tensor space $\mathbb{F}^{n \times n \times n}$ as a sum of tensor subspaces, as efficiently as possible such that each subspace has the form of an $n \times n$ matrix subspace tensored with $\mathbb{F}^n$

$X \subseteq \mathrm{Mat}_{n \times n} = \mathbb{F}^n \otimes \mathbb{F}^n$

$X[1] = \mathbb{F}^n \otimes X \subseteq \mathbb{F}^{n \times n \times n}$

$X[2] =$

$X[3] = X \otimes \mathbb{F}^n$

Tensor space decompositions

Goal: write tensor space $\mathbb{F}^{n \times n \times n}$ as a sum of tensor subspaces, as efficiently as possible such that each subspace has the form of an $n \times n$ matrix subspace tensored with $\mathbb{F}^n$

$X \subseteq \mathrm{Mat}_{n \times n} = \mathbb{F}^n \otimes \mathbb{F}^n$

$X[1] = \mathbb{F}^n \otimes X \subseteq \mathbb{F}^{n \times n \times n}$

$X[2] =$

$X[3] = X \otimes \mathbb{F}^n$

Theorem there are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3 \times 3 \times 3} = X_1[1] + X_2[2] + X_3[3].$$

Theorem There are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3\times 3\times 3} = X_1[1] + X_2[2] + X_3[3].$$

Note: dimensions left and right are equal.

Theorem There are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3\times 3\times 3} = X_1[1] + X_2[2] + X_3[3].$$

Note: dimensions left and right are equal.

Remark Not possible with matrices: there are no subspaces $X_1 \subseteq \mathbb{F}^n$ of dimension $n/2$ each such that $\mathbb{F}^{n\times n} = X_1[1] + X_2[2]$

Theorem There are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3\times 3\times 3} = X_1[1] + X_2[2] + X_3[3].$$

Note: dimensions left and right are equal.

Remark Not possible with matrices: there are no subspaces $X_1 \subseteq \mathbb{F}^n$ of dimension $n/2$ each such that $\mathbb{F}^{n\times n} = X_1[1] + X_2[2]$

Theorem There are subspaces $X_i \subseteq (\mathbb{F}^n)^{\otimes n-1}$ of dim n^{n-2} each, such that

$$(\mathbb{F}^n)^{\otimes n} = X_1[1] + X_2[2] + \cdots + X_n[n].$$

Again: dimensions match.

Theorem there are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3 \times 3 \times 3} = X_1[1] + X_2[2] + X_3[3].$$

Theorem there are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

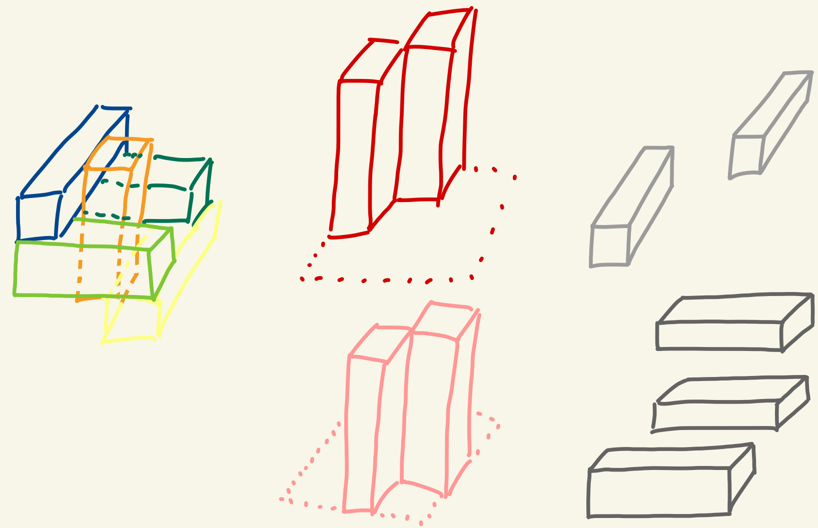
$$\mathbb{F}^{3\times 3\times 3} = X_1[1] + X_2[2] + X_3[3].$$

$$\begin{array}{ccc}
X_1 & X_2 & X_3 \\[1em]
\begin{matrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} &
\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} &
\begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \\[1em]
\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{matrix} &
\begin{matrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} &
\begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \\[1em]
\begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{matrix} &
\begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{matrix} &
\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{matrix}
\end{array}$$

Theorem there are subspaces $X_i \subseteq \mathrm{Mat}_{3,3}$ of dim 3 each, such that

$$\mathbb{F}^{3\times3\times3} = X_1[1] + X_2[2] + X_3[3].$$

X_1	X_2	X_3
$\begin{matrix}1&0&0\\0&0&0\\0&0&0\end{matrix}$	$\begin{matrix}0&0&0\\0&0&1\\0&0&0\end{matrix}$	$\begin{matrix}0&0&0\\0&1&0\\0&0&0\end{matrix}$
$\begin{matrix}0&0&0\\0&0&0\\0&0&1\end{matrix}$	$\begin{matrix}0&0&0\\1&0&0\\0&0&0\end{matrix}$	$\begin{matrix}0&1&0\\1&0&0\\0&0&0\end{matrix}$
$\begin{matrix}0&0&1\\0&0&0\\1&0&0\end{matrix}$	$\begin{matrix}0&0&1\\0&0&0\\1&0&1\end{matrix}$	$\begin{matrix}0&0&0\\0&0&1\\0&1&0\end{matrix}$



Application: Subrank is not additive under direct sum

Theorem There are tensors $S, T \in \mathbb{F}^{n \times n \times n}$ such that $Q(S), Q(T) \leq \sqrt{3n-2}$ while $Q(S \oplus T) \geq n$.

Application: Subrank is not additive under direct sum

Theorem There are tensors $S, T \in \mathbb{F}^{n\times n\times n}$ such that $Q(S), Q(T) \leq \sqrt{3n-2}$ while $Q(S \oplus T) \geq n$.

Proof idea

- Let T be "random".
- Let $S = I_n - T$. Then S is "random".
- Then $Q(S), Q(T) \leq \sqrt{3n-2}$ by our theorem.
- On the other hand, $Q(S \oplus T) \geq Q(S+T) = Q(I_n) = n$. $\square$

Selected Open Problems

1. Our upper bound $Q(T) \leq \lfloor \sqrt{3n-2} \rfloor$ for generic $T \in \mathbb{F}^{n \times n \times n}$ is tight for $n \leq 100$. Is this always true?

2. Determine all possible tensor space decompositions

3. What is the largest gap between $Q(S \oplus T)$ and $Q(S) + Q(T)$?

New directions: asymptotic gap in the subrank

$S \in V_1 \otimes V_2 \otimes V_3$

$T \in W_1 \otimes W_2 \otimes W_3$

Kronecker product: $S \boxtimes T \in (V_1 \otimes W_1) \otimes (V_2 \otimes W_2) \otimes (V_3 \otimes W_3)$

New directions: asymptotic gap in the subrank

$S \in V_1 \otimes V_2 \otimes V_3$

$T \in W_1 \otimes W_2 \otimes W_3$

Kronecker product: $S \boxtimes T \in (V_1 \otimes W_1) \otimes (V_2 \otimes W_2) \otimes (V_3 \otimes W_3)$

Subrank is super-mult: $Q(S \boxtimes T) \geq Q(S)\, Q(T)$

New directions: asymptotic gap in the subrank

$S \in V_1 \otimes V_2 \otimes V_3$

$T \in W_1 \otimes W_2 \otimes W_3$

Kronecker product: $S \boxtimes T \in (V_1 \otimes W_1) \otimes (V_2 \otimes W_2) \otimes (V_3 \otimes W_3)$

Subrank is super-mult: $Q(S \boxtimes T) \geq Q(S)\,Q(T)$

Question How does $Q(T^{\boxtimes n})$ behave when $n \to \infty$?

New directions: asymptotic gap in the subrank

$S \in V_1 \otimes V_2 \otimes V_3$

$T \in W_1 \otimes W_2 \otimes W_3$

Kronecker product: $S \boxtimes T \in (V_1 \otimes W_1) \otimes (V_2 \otimes W_2) \otimes (V_3 \otimes W_3)$

Subrank is super-mult: $Q(S \boxtimes T) \geq Q(S)\, Q(T)$

Question How does $Q(T^{\boxtimes n})$ behave when $n \to \infty$?

Theorem (Christandl, Gesmundo, Z) Let $T \in V_1 \otimes V_2 \otimes V_3$ be any tensor.
Exactly one of the following is true:

i) $T = 0$

ii) $Q(T^{\boxtimes n}) = 1$ for all n

iii) $Q(T^{\boxtimes n}) = 1.88^{n - o(n)}$ for all n

iv) $Q(T^{\boxtimes n}) \geq 2^n$ for all n