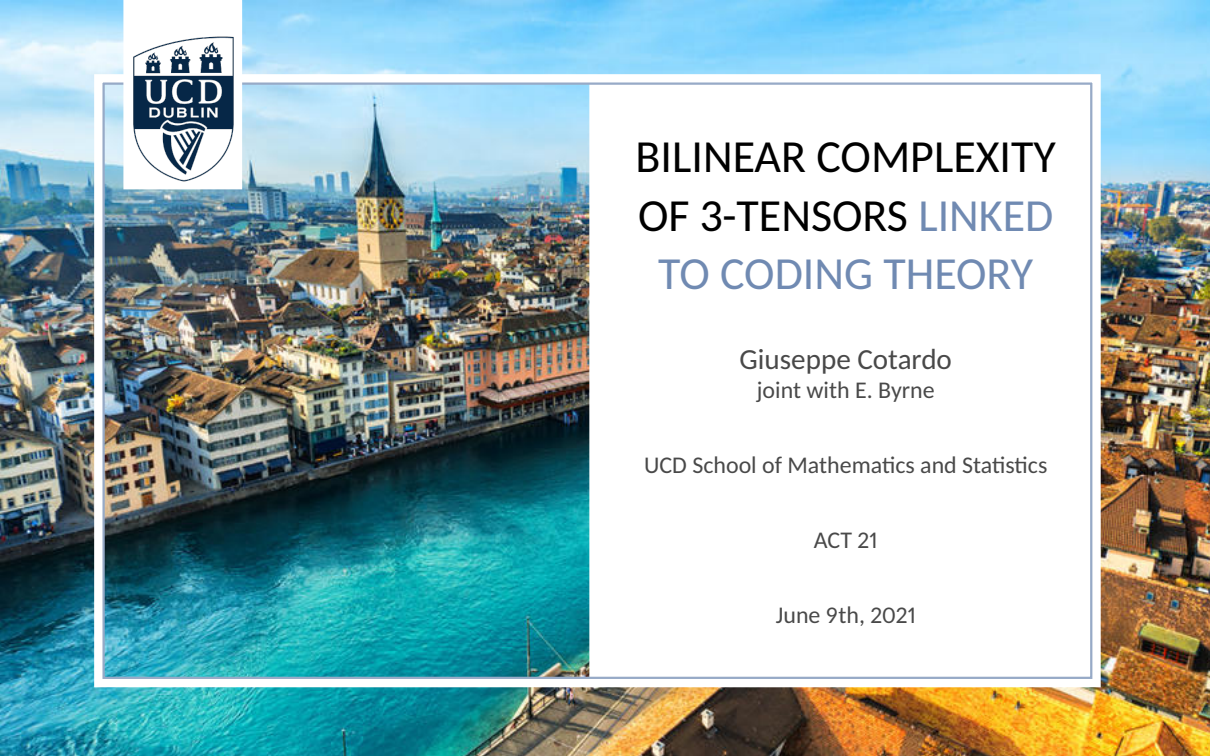


BILINEAR COMPLEXITY OF 3-TENSORS LINKED TO CODING THEORY

Giuseppe Cotardo
joint with E. Byrne

UCD School of Mathematics and Statistics

ACT 21

June 9th, 2021

WHAT IS COMPLEXITY ?

Definition

The **complexity** of a *problem* is the cost of the optimal procedure among all the ones that solve the *problem* and fit into a given model of computation.

- The cost of a *computation* that solves a problem is an **upper bound** on the complexity of that problem with respect to the given model.

- We are interested in the so-called **nonscalar model** where additions, subtractions and scalar multiplications are free of charge. The (**nonscalar**) **cost** of an algorithm is therefore the number of multiplications and divisions needed to compute the result.

Let A, B be 2×2 following matrices

$$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}.$$

The standard algorithm returns the matrix $C = AB$ by computing the following intermediate results:

$$\begin{aligned} c_1 &= a_1b_1 + a_2b_3, & c_2 &= a_1b_2 + a_2b_4, \\ c_3 &= a_3b_1 + a_4b_3, & c_4 &= a_3b_2 + a_4b_4. \end{aligned}$$

It requires 8 **multiplications** and 4 **additions**. Therefore, an **upper bound** for the complexity (in the nonscalar model) is 8.

We can compute $C = AB$ using Strassen's algorithm, which gives

$$c_1 = S_1 + S_4 - S_5 + S_7, \qquad c_2 = S_2 + S_4, \qquad c_3 = S_3 + S_5, \qquad c_4 = S_1 + S_3 - S_2 + S_6$$

where the S_i's are the intermediate steps

$$S_1 = (a_1 + a_4)(b_1 + b_4), \qquad S_2 = (a_3 + a_4)b_1, \qquad S_3 = a_1(b_3 - b_4),$$
$$S_4 = a_4(b_3 - b_1), \qquad S_5 = (a_1 + a_2)b_4, \qquad S_6 = (a_3 - a_1)(b_1 + b_2),$$
$$S_7 = (a_2 - a_4)(b_3 + b_4).$$

It requires 7 **multiplications** and 18 **additions**.

Algorithm	# multiplication	# additions
standard	8	4
Strassen's	7	18

Remark

The complexity of multiplying 2×2 matrices (in the nonscalar model) is 7. The upper-bound is given by Strassen (1969), the lower bound was proved by Winograd (1971).

We assume n, m, k to be integers.

Definition

A 3-**tensor** $X := \sum_r a_r \otimes b_r \otimes c_r$ is an element of $\mathbb{K}^k \otimes \mathbb{K}^n \otimes \mathbb{K}^m$.

We assume n, m, k to be integers.

Definition

A 3-**tensor** $X := \sum_r a_r \otimes b_r \otimes c_r$ is an element of $\mathbb{K}^k \otimes \mathbb{K}^n \otimes \mathbb{K}^m$.

X can be seen as the representation of a bilinear map

$$\mathbb{K}^n \times \mathbb{K}^m \longrightarrow \mathbb{K}^k.$$

We assume n, m, k to be integers.

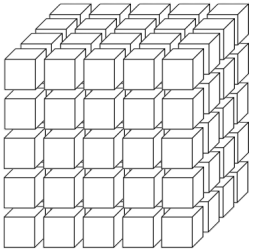
Definition

A 3-**tensor** $X := \sum_r a_r \otimes b_r \otimes c_r$ is an element of $\mathbb{K}^k \otimes \mathbb{K}^n \otimes \mathbb{K}^m$.

X is is related to the the 3-dimensional array

$$X_{ij\ell} = \sum_r (a_r)_\ell \cdot (b_r)_i \cdot (c_r)_j$$

which implies $\mathbb{K}^k \otimes \mathbb{K}^n \otimes \mathbb{K}^m \simeq \mathbb{K}^{k \times n \times m}$.



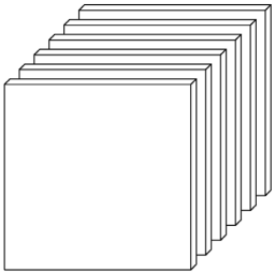
We assume n, m, k to be integers.

Definition

A 3-**tensor** $X := \sum_r a_r \otimes b_r \otimes c_r$ is an element of $\mathbb{K}^k \otimes \mathbb{K}^n \otimes \mathbb{K}^m$.

We can also identify the tensor X with the array of $n \times m$ matrices

$$X = (X_1 \mid \ldots \mid X_k).$$



Definition

X is said to be **simple** (or **rank one**) if there exist $a \in \mathbb{K}^k, b \in \mathbb{K}^n$ and $c \in \mathbb{K}^m$ such that $X = a \otimes b \otimes c$.

Definition

The **tensor rank** $\mathrm{trk}(X)$ of X is defined as the smallest R such that X can be expressed as sum of R simple tensors.

Definition

Let $\mathcal{A} := \{A_1, \ldots, A_R\} \subseteq \mathbb{K}^{n \times m}$ be a set of R l.i. rank-1 matrices. We say that $\mathcal{A}$ is a **perfect base** (or R-**base**) for the tensor X if $\langle X_1, \ldots, X_k \rangle \leq \langle A_1, \ldots, A_R \rangle$.

TENSOR RANK FOR $(mn-2)$ - LAYER TENSORS

Theorem (Atkinson, Lloyd - 1983)

Let $\mathrm{char}(\mathbb{K}) \neq 2$ and let $X \in \mathbb{K}^{(mn-2)\times n\times m}$ be a tensor. We have $\mathrm{trk}(X) = mn-2$ unless X is such that $X_{j,1,1} + X_{j,2,2} = 0$ and $X_{j,1,2} = 0$ for all $1 \leq j \leq mn-2$.

TENSOR RANK FOR $(mn-2)$ - LAYER TENSORS

Theorem (Atkinson, Lloyd - 1983)

Let $\text{char}(\mathbb{K}) \neq 2$ and let $X \in \mathbb{K}^{(mn-2)\times n\times m}$ be a tensor. We have $\text{trk}(X) = mn-2$ unless X is such that $X_{j,1,1} + X_{j,2,2} = 0$ and $X_{j,1,2} = 0$ for all $1 \leq j \leq mn-2$.

Definition

The **dual** of $V \leq \mathbb{K}^{n\times m}$ is $V^{\perp} := \{N \in \mathbb{K}^{n\times m} : \text{Tr}(MN^t) = 0 \ \ \forall M \in V, M \neq 0\}$.

Definition (Atkinson, Lloyd - 1983)

A space of $n \times m$ matrices is said to be **perfect** if it is generated by rank-1 matrices.

Theorem (Byrne, C.)

Let $s \in \{1, \dots, m-1\}$, $|\mathbb{K}| \geq s+1$, $\mathcal{S} := \{1, \gamma_1, \dots, \gamma_{s-1}\}$ be a set of distinct elements of $\mathbb{K} \setminus \{0\}$ and $M \in \mathbb{K}^{m \times m}$ be a companion matrix of an irreducible polynomial of degree m. We have that $\left\langle I, M, \dots, M^{s-1} \right\rangle^{\perp} \leq \mathbb{K}^{m \times m}$ is perfect and an $(m^2 - s)$-base is

$$\mathcal{A}(\mathcal{S}) := \{J^i E_{1,j} \left(M^{-i}\right)^{\mathsf{t}} : s+1 \leq j \leq m, 0 \leq i \leq m-1\} \cup \{J^i \mathcal{E}(\gamma) \left(M^{-i}\right)^{\mathsf{t}} : 0 \leq i \leq m-2, \gamma \in \mathcal{S}\},$$

where $E_{1,j}$ is the matrix with 1 in position $(1, j)$ and zeros elsewhere,

$$J := \left(\begin{array}{c|c} 0 & 1 \\ \hline I_{m-1} & 0 \end{array}\right) \quad \text{and} \quad \mathcal{E}(\gamma) := \left(\begin{array}{ccccc} \gamma^m & \gamma^{m-1} & \cdots & \gamma & 1 \\ -\gamma^{m+1} & -\gamma^m & \cdots & -\gamma^2 & -\gamma \\ \hline & & 0 & & \end{array}\right).$$

Definition

A (**matrix rank-metric**) **code** is a subspace $\mathcal{C} \leq \mathbb{F}_q^{n \times m}$. The **minimum** (**rank**) **distance** of a non-zero code $\mathcal{C}$ is $d(\mathcal{C}) := \min(\{\mathrm{rk}(M) : M \in \mathcal{C}, M \neq 0\})$ and for $\mathcal{C} := \{0\}$, we define $d(\mathcal{C})$ to be $n + 1$.

Definition

A (**matrix rank-metric**) **code** is a subspace $\mathcal{C} \leq \mathbb{F}_q^{n \times m}$. The **minimum** (**rank**) **distance** of a non-zero code $\mathcal{C}$ is $d(\mathcal{C}) := \min(\{\mathrm{rk}(M) : M \in \mathcal{C}, M \neq 0\})$ and for $\mathcal{C} := \{0\}$, we define $d(\mathcal{C})$ to be $n + 1$.

It is well-know that the dual $\mathcal{C}^\perp$ of $\mathcal{C}$ is a code.

Definition

A (**matrix rank-metric**) **code** is a subspace $\mathcal{C} \leq \mathbb{F}_q^{n \times m}$. The **minimum** (**rank**) **distance** of a non-zero code $\mathcal{C}$ is $d(\mathcal{C}) := \min(\{\mathrm{rk}(M) : M \in \mathcal{C}, M \neq 0\})$ and for $\mathcal{C} := \{0\}$, we define $d(\mathcal{C})$ to be $n + 1$.

It is well-know that the dual $\mathcal{C}^\perp$ of $\mathcal{C}$ is a code.

Proposition (Kruskal - 1977)

We have that $\mathrm{trk}(\mathcal{C}) \geq \dim_{\mathbb{F}_q}(\mathcal{C}) + d(\mathcal{C}) - 1$.

Codes meeting this bound are called **MTR** (**Minimal Tensor Rank**).

Let $\Gamma := \{\gamma_1, \ldots, \gamma_m\}$ be a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and $v \in \mathbb{F}_{q^m}^n$ and we define the map

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \xmapsto{\Gamma} \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}.$$

This map is an $\mathbb{F}_q$-isomorphism.

Definition

A **vector** (**rank-metric**) **code** is a subspace $C \leq \mathbb{F}_{q^m}^n$. The **minimum distance** $d(C)$ of C is the minimum distance of $\Gamma(C)$ for any choice of a basis Γ of $\mathbb{F}_{q^m}/\mathbb{F}_q$.

Definition

Let $1 \leq k \leq n$ and $\beta_1, \ldots, \beta_n \in \mathbb{F}_{q^m}$ be l.i. over $\mathbb{F}_q$. The $\mathbb{F}_{q^m}$**-linear Gabidulin code** $\mathcal{G}_k(\beta_1, \ldots, \beta_n)$ is defined as

$$\left\{ (f(\beta_1), \ldots, f(\beta_n)) : f \in \left\{ f_0 x + f_1 x^q + \cdots + f_{k-1} x^{q^{k-1}} : f_0, \ldots, f_{k-1} \in \mathbb{F}_{q^m} \right\} \right\},$$

Definition

Let $1 \leq k \leq n$ and $\beta_1, \ldots, \beta_n \in \mathbb{F}_{q^m}$ be l.i. over $\mathbb{F}_q$. The $\mathbb{F}_{q^m}$**-linear Gabidulin code** $\mathcal{G}_k(\beta_1, \ldots, \beta_n)$ is defined as

$$\left\{(f(\beta_1), \ldots, f(\beta_n)) : f \in \left\{f_0 x + f_1 x^q + \cdots + f_{k-1} x^{q^{k-1}} : f_0, \ldots, f_{k-1} \in \mathbb{F}_{q^m}\right\}\right\},$$

Proposition (Byrne, Neri, Ravagnani, Sheekey - 2019)

Let $q \geq m + n - 2$, α be primitive element of $\mathbb{F}_{q^m}$ and $\lambda \in \mathbb{F}_{q^m}$. For any integer $j \in \{0, \ldots, m-1\}$, we have

$$\mathrm{trk}\left(\mathcal{G}_1\left(\lambda, \lambda\alpha^{q^j}, \ldots, \lambda\alpha^{nq^j}\right)\right) = m + n - 1$$

and, in particular, the code is MTR.

- Some 1-dimensional Gabidulin codes corresponds to the multiplication in $\mathbb{F}_{q^m}$. This is well studied problem in complexity theory.

- The tensor rank is invariant under equivalence.

- The tensor rank does not dualize.

Theorem (Byrne, C.)

Let $q \geq m$, α be primitive element of $\mathbb{F}_{q^m}$ and $\lambda \in \mathbb{F}_{q^m}$. For any integer $j \in \{0, \ldots, m-1\}$, we have

$$\mathrm{trk}\left(\mathcal{G}_1\left(\lambda, \lambda\alpha^{q^j}, \ldots, \lambda\alpha^{nq^j}\right)^{\perp}\right) = nm - m + 1.$$

and, in particular, the code is MTR.

Theorem (Byrne, C.)

Let $q \geq m$, α be primitive element of $\mathbb{F}_{q^m}$ and $\lambda \in \mathbb{F}_{q^m}$. For any integer $j \in \{0, \ldots, m-1\}$, we have

$$\mathrm{trk}\left(\mathcal{G}_1\left(\lambda, \lambda\alpha^{q^j}, \ldots, \lambda\alpha^{nq^j}\right)^{\perp}\right) = nm - m + 1.$$

and, in particular, the code is MTR.

THANK YOU