Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

OPTIMAL LOW-RANK APPROXIMATION USING TENSOR NETWORK STRUCTURE SEARCH

Autor: Aran Roig

Director: Dr. Nahuel Statuto

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, April 14, 2025

Abstract

TODO

Goldbach's weak conjecture asserts that every odd integer greater than 5 is the sum of three primes. We study that problem and the proof of it presented by H. A. Helfgott and D. Platt. We focus on the circle method. Finally, we describe a computation that confirms Goldbach's weak conjecture up to 10^{28}.

Resum

TODO

La conjectura feble de Goldbach afirma que tot nombre enter imparell major que 5 és la suma de tres nombres primers. En aquest treball estudiem aquest problema i la seva prova presentada per HA Helfgott i D. Platt. Ens centrem en el mètode del cercle. Finalment, describim un càlcul que confirma la conjectura feble de Goldbach fins a 10^{28}.

2020 Mathematics Subject Classification. 11G05, 11G10, 14G10

Agraïments

Vull agrair a ...

Contents

1 Introduction **1**
- 1.1 Objectives . . . 1
- 1.2 Thesis structure . . . 1
- 1.3 Preliminaries . . . 1

2 Tensors **3**
- 2.1 Penrose notation . . . 6

3 Tensor Networks **8**
- 3.1 Tensor networks and tensor network states . . . 8
- 3.2 Common Tensor network structures . . . 9
- 3.3 Approximating tensors to tensor network states . . . 10
- 3.4 The tensor tensor network structure space . . . 11
- 3.5 Finding the best structure . . . 11

4 Conclusions **13**

Bibliography **14**

A Chapter 1 **16**

Chapter 1

Introduction

TODO FALTA ACABAR TOT AQUEST CAPÍTOL

A well known problem in science and engineering is to retrieve a function given some data. It may be for example the solution of a partially differentiable equation given some boundary conditions or initial data or even a target function to be learned from some training set data. [10]

Explicar que normalment es que les funcions que resulten d'aixo segurament viuen en un espai molt gran, segurament amb una dimensió molt alta i que una cosa que s'hi sol fer es assumir per exemple que la nostra funció és pot reescriure com una de low-rank, és a dir, que es pot escriure com la suma d'unes altres funcions.

A partir d'això entren en joc un concepte originat de part de la física, les tensor networks. Les tensor networks no són més que representacions de tensors, normalment d'ordre alt, que es poden recuperar a partir de contraure diferents tensors d'ordres més petits. Resulta que les tensor networks, si representem les dades de la nostra funció com un tensor d'ordre alt, no acaben sent més que aproximacions low-rank del nostre problema.

L'objectiu final d'aquest treball es intentar donar una solució óptima a aquest problema utilitzant xarxes tensorials i després mostrar alguna aplicació en diferents camps de la física i del machine learning.

En resum hauria d'explicar el problema sencer

1.1 Objectives

1.2 Thesis structure

1.3 Preliminaries

Definition 1.1 (Graph)**.** [9] A **graph** G is defined as a tuple $G = (V, E)$ where $V = V(G)$ is a set of elements called **vertices** and $E = E(G) \subset \{\{u, v\} : u, v \in V\}$ is a set of elements called **edges**.

Definition 1.2 (Directed Graph)**.** *A* ***directed graph*** *D is a tuple $D = (V, \bar{E})$ where $V = V(D)$ are its vertices and $\bar{E} = \bar{E}(D) \subset \{(u, v) : u, v \in V\}$*

Definition 1.3. *Given a directed graph* $G = (V, \bar{E})$ *and a vertex* $i \in V$ *we define*

$$\mathrm{IN}(i) = \{j \in V : (j, i) \in \bar{E}\} \qquad \mathrm{OUT}(i) = \{j \in V : (i, j) \in \bar{E}\}$$

Chapter 2

Tensors

In this chapter we will introduce the basics of tensor algebra and we will present some results that will be useful in the following chapters

We will denote $\mathbb{V}_1, \dots, \mathbb{V}_n$ as finite vector spaces over a field $\mathbb{K}$ ($\mathbb{C}$ if unspecified) of dimension $\dim \mathbb{V}_i = N_i \; \forall i = 1, \dots, n$. We denote $\mathcal{B}_1, \dots, \mathcal{B}_n$ with $\mathcal{B}_i = \{e_1^{(i)}, \dots, e_{N_i}^{(i)}\}, i = 1, \dots, n$ the canonical basis of $\mathbb{V}_1, \dots, \mathbb{V}_n$

Definition 2.1 (Tensor). *A **tensor** is a multilineal map $T : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to \mathbb{K}$*

Definition 2.2 (Order of a tensor). *Given a tensor $T \in \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to \mathbb{K}$ We define the **order** of the tensor T as n.*

We will allow that some $\mathbb{V}_i$ are dual spaces of some other vector spaces $\mathbb{W}_i$, i.e $\mathbb{V}_i = \mathbb{W}_i^*$.

Definition 2.3. *Let $T \in \mathbb{V}_1 \times \cdots \times \mathbb{V}_p \times \mathbb{W}_1^* \times \cdots \times \mathbb{W}_q^* \to \mathbb{K}$. We say that the tensor T is p-times covariant and q-times contravariant.*

Definition 2.4 (Tensor product). *Let L be the vector space generated by the base $V_1 \times \cdots \times V_n$, i.e the set of linear combinations of the elements $(v_1, \dots, v_n), v_i \in \mathbb{V}_i$. Let $\mathcal{R}$ be the linear subspace of L generated by the relation R defined by:*

$$(v_1, \dots, \alpha v_i, \dots, v_n) \sim \alpha(v_1, \dots, v_n) \; \forall i = 1, \dots, n, \forall \alpha \in \mathbb{K}$$

$$(v_1, \dots, v_i + u_i, \dots, v_n) \sim (v_1, \dots, v_i \dots, v_n) + (v_1, \dots, u_i, \dots, v_n) \; \forall i = 1, \dots, n$$

*The tensor product $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ is defined as the quotient $L/\mathcal{R}$ and its called **tensor product space**. The image of $(v_1, \dots, v_n)$ by the quotient is denoted by $v_1 \otimes \cdots \otimes v_n$*

The following theorem gives us a correspondance between each tensor $T : \mathbb{V}_1 \times \cdots \times \mathbb{V}_n \to \mathbb{K}$ and each element of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$:

Theorem 2.5 (Universal property of the tensor product). [6] The tensor product of two vector spaces $\mathbb{V} \otimes \mathbb{W}$ for every bilinear map $h : \mathbb{V} \times \mathbb{W} \to X$ there exists an unique bilinear map $\tilde{h} : \mathbb{V} \otimes \mathbb{W} \to X$ such that the following diagram commutes:

$$\begin{tikzcd} \mathbb{V} \times \mathbb{W} \arrow[r, "\varphi"] \arrow[dr, "h"'] & \mathbb{V} \otimes \mathbb{W} \arrow[d, "\tilde{h}"] \\ & X \end{tikzcd}$$

We can construct a basis for $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$. We define

$$\mathcal{B}_{\otimes} = \{e^{(1)}_{i_1} \otimes \cdots \otimes e^{(n)}_{i_n} : 1 \leqslant i_j \leqslant N_j, 1 \leqslant j \leqslant n\}$$

Where $\{e^{(i)}_1, e^{(i)}_2, \ldots, e^{(i)}_{N_i}\}$ is the canonical basis for $\mathbb{V}_i$. Constructed this way, $\mathcal{B}_{\otimes}$ is a (canonical) basis of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$.

Remark 2.6. The dimension of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ is $\dim \mathbb{V}_1 \cdot \dim \mathbb{V}_2 \cdots \dim \mathbb{V}_n$ and its elements can be expressed as

$$T = \sum_{s_1,\ldots,s_n}^{N_1,\ldots,N_n} T_{s_1,\ldots,s_n} \cdot e^{(1)}_{s_1} \otimes \cdots \otimes e^{(n)}_{s_n} \tag{2.0.1}$$

Definition 2.7 (Tensor product)**.** *Given two tensors $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ and $U \in \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$ with $\dim \mathbb{V}_i = N_i$, $i = 1, \ldots, n$ and $\dim \mathbb{W}_j = M_j$, $j = 1, \ldots, m$ and canonical basis $\{v^{(i)}_1, \ldots, v^{(i)}_{N_i}\}$ for each $\mathbb{V}_i$ and $\{w^{(j)}_1, \ldots, w^{(j)}_{M_j}\}$ for each $\mathbb{W}_j$ we define the tensor product $T \otimes U$ as*

$$T \otimes U = \sum_{i_1,\ldots,i_n}^{N_1,\ldots,N_n} \sum_{j_1,\ldots,j_m}^{M_1,\ldots,M_m} T_{i_1,\ldots,i_n} U_{j_1,\ldots,j_m} \cdot v^{(1)}_{i_1} \otimes \cdots \otimes v^{(n)}_{i_n} \otimes w^{(1)}_{j_1} \otimes \cdots \otimes w^{(m)}_{j_m}$$

Note that $T \otimes U \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n \otimes \mathbb{W}_1 \otimes \cdots \otimes \mathbb{W}_m$

From remark 2.6, a tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ can be identified as a "n-dimensional array" $\mathcal{T}$, i.e, a discrete function

$$\begin{aligned} \mathcal{T} : \prod_{i=1}^{n} \{1, \ldots, N_i\} &\longrightarrow \mathbb{K} \\ T(i_1, \ldots, i_n) &\longmapsto T_{i_1,\ldots,i_n} \end{aligned}$$

From now on, we will write the discrete function $\mathcal{T}$ as an element of $\mathbb{K}^{N_1 \times \cdots \times N_n}$ [11]. of T as $\mathcal{T}$. Now we will define some definitions from the underlying definition of the tensor viewed as a discrete function:

Now we will define some definitions that will help us establish a formalized method of mapping n-dimensional arrays with vectors and matrices:

Definition 2.8 (Linearization)**.** *Fixed $N_1, \ldots, N_n \in \mathbb{N}$, given $i_1, \ldots, i_n \in \mathbb{N}$ such that $1 \leqslant i_1, \leqslant N_1, \ldots, 1 \leqslant i_n \leqslant N_n$, we define the Linearization of the indices $i_1, \ldots, i_n$ as the mapping $\prod_{i=1}^{n} \{1, \ldots, N_i\} \to \{1, \ldots, \prod_{i=1}^{n} N_i\}$:*

$$\overline{i_1, i_2, \ldots, i_n} = \sum_{j=2}^{n} \left((i_j - 1) \prod_{k=1}^{j} N_k \right) + i_1$$

Lemma 2.9. *The linearization mapping is bijective*

Definition 2.10 (Vectorization)**.** *Given a tensor $\mathcal{T} \in \mathbb{K}^{N_1 \times \cdots \times N_n}$, we define the **vectorization** of $\mathcal{T}$ as the first order tensor $\mathcal{V} \in \mathbb{K}^{N_1 N_2 \cdots N_n}$ defined by:*

$$\mathcal{V}(\overline{i_1 i_2 \ldots i_n}) = \mathcal{T}(i_1, i_2, \ldots, i_n)$$

We will write $\mathcal{V} = \operatorname{vec} \mathcal{T}$

Definition 2.11 (Tensor unfolding)**.** *Let* $\mathcal{T} \in \mathbb{R}^{N_1 \times \cdots \times N_n}$, $n \geqslant 2$, $1 \leqslant d \leqslant n$ *and* p *a permutation of the vector* $(1, 2, \ldots, n)$. *We define the* ***generalized unfolding*** *of the tensor* $\mathcal{T}$ *as the 2nd-order tensor* $\mathcal{U} \in \mathbb{R}^{\prod_{i=1}^{d} N_{p_i} \times \prod_{i=d+1}^{n} N_{p_i}}$:

$$\mathcal{U}(\overline{i_{p_1}, \ldots, i_{p_d}}, \overline{i_{p_{d+1}}, \ldots, i_{p_n}}) = \mathcal{T}(i_1, \ldots, i_n)$$

We will write $\mathcal{U} = \operatorname{unfold}(\mathcal{T}, (p_1, \ldots, p_d), (p_{d+1}, \ldots, p_n))$. *We also define* $\operatorname{unfold}_d T := \operatorname{unfold}(\mathcal{T}, (1, \ldots, d), (d+1, \ldots, n))$

Definition 2.12 (Tensor slices)**.** *Let* $\mathcal{T} \in \mathbb{K}^{N_1 \times \cdots \times N_n}$. *Consider* $S \subset \{1, \ldots, n\}$ *a subset of modes (dimensions) of* $\mathcal{T}$. *Let* $\mathbf{i}_S = (i_k)_{k \in S}$. *We define the* ***slice of*** $\mathcal{T}$ *as the tensor* $\mathcal{T}_{\mathbf{i}_S}$ *of order* $n - \#S$

$$\begin{aligned} \mathcal{T}_{\mathbf{i}_S} : \prod_{k \notin S} \{1, \ldots, N_k\} &\longrightarrow \mathbb{K} \\ \mathcal{T}_{\mathbf{i}_S}(i_{j_1}, \ldots, i_{j_m}) &\longmapsto \mathcal{T}(i_1, \ldots, i_n) \end{aligned}$$

Where $\{j_1, \ldots, j_m\} = \{1, \ldots, n\} \setminus S$.

Sometimes we will also implicitly specify $\mathbf{i}_S$ *by writing* $\mathcal{T}(a_1, \ldots, a_n)$ *and replacing* a_j *with* i_j *if* $j \in S$ *and ":" otherwise*

Example 2.13. Consider $\mathcal{M} \in \mathbb{K}^{N_1 \times N_2}$ a second order tensor. We can see this tensor as a bidimensional array (matrix). The slice $\mathcal{M}(i, :)$ results in the i-th row of $\mathcal{M}$. The slice $\mathcal{M}(:, j)$ results in the j-th column of $\mathcal{M}$

Example 2.14. Consider a 4th-order tensor $\mathcal{T} \in \mathbb{K}^{N_1 \times N_2 \times N_3 \times N_4}$. Fixed i_2, i_3, the tensor slice of 2th-order $\mathcal{T} = \mathcal{A}(:, i_2, i_3, :) \in \mathbb{K}^{N_1 \times N_4}$ with its entries defined by $\mathcal{T}(i_1, i_4) = \mathcal{A}(i_1, i_2, i_3, i_4)$

Definition 2.15 (Rank of a tensor)**.** *We say that a tensor is of rank* r *and we write* $\operatorname{rank} T = r$ *with* $r \in \mathbb{N}$ *being the minimum value such that we can write* T *as*

$$T = \sum_{p=1}^{r} v_p^{(1)} \otimes \cdots \otimes v_p^{(n)}$$

where $v_1^{(i)}, \ldots, v_r^{(i)} \in \mathbb{V}_i, i = 1, \ldots, n$

One can easily see that $\operatorname{rank} T \leqslant \prod_{i=1}^{n} N_i$. Unlike matrices, determining the rank of a tensor is an NP-hard problem. [2]. Even finding the maximum rank, (i.e determining $\max_{T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n} \operatorname{rank} T$) this is still an unresolved problem. We will now present an slightly better upper bound for the tensor rank:

Proposition 2.16.

$$\operatorname{rank} T \leqslant \left\lfloor \frac{\prod_{i=1}^{n} N_i}{\sum_{i=1}^{n} N_i} \right\rfloor \tag{2.0.2}$$

Proof. Let $r = \operatorname{rank} T$. We can write $T = \sum_{p=1}^{r} v_p^{(1)} \otimes \cdots \otimes v_p^{(n)}$. Now, each term of this sum has $\sum_{i=1}^{n} N_i$ adjustable parameters, since each $v_p^{(i)}$ is a vector of $\mathbb{V}_i$ with its dimension being N_i. So, in total we will have $r \sum_{i=1}^{n} N_i$ adjustable parameters in our decomposition. Since our tensor T is completly determined by $\prod_{i=1}^{n} N_i$ parameters, we can impose that $r \sum_{i=1}^{n} N_i \leqslant \prod_{i=1}^{n} N_i$ $\square$

Definition 2.17. [1] Let $T \in \left(\bigotimes_{i=1}^{k-1} \mathbb{V}_i \otimes \mathbb{V}_k \otimes \bigotimes_{i=k+1}^{p} \mathbb{V}_i\right) \otimes \left(\bigotimes_{i=1}^{l-1} \mathbb{W}_i^* \otimes \mathbb{W}_l^* \otimes \bigotimes_{i=l+1}^{q} \mathbb{W}_i^*\right)$ With $\mathbb{V}_k = \mathbb{W}_l$. Consider the mapping

$$\mathcal{C}_k^l : \left(\bigotimes_{i=1}^{p} \mathbb{V}_i\right) \otimes \left(\bigotimes_{i=1}^{q} \mathbb{W}_i^*\right) \longrightarrow \left(\bigotimes_{i=1}^{k-1} \mathbb{V}_i \otimes \bigotimes_{i=k+1}^{p} \mathbb{V}_i\right) \otimes \left(\bigotimes_{i=1}^{l-1} \mathbb{W}_i^* \otimes \bigotimes_{i=l+1}^{q} \mathbb{W}_i^*\right)$$

$$\mathcal{C}_k^l \left(\bigotimes_{i=1}^{p} v_i \otimes \bigotimes_{i=1}^{q} f_q\right) = \left(\bigotimes_{i=1}^{k-1} v_i \otimes \bigotimes_{i=k+1}^{p} v_i \otimes \bigotimes_{i=1}^{l-1} f_i \otimes \bigotimes_{i=l+1}^{q} f_i\right) f_l(v_k)$$

$\mathcal{C}_k^l$ is defined as the tensor contraction mapping of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_p \otimes \mathbb{V}_{p+1}^* \otimes \cdots \otimes \mathbb{V}_q^*$ by the indices (k, l). We call $\mathcal{C}_k^l(T)$ the contraction of T by indices (k, l)

Definition 2.18. *Let* $X \in \left(\bigotimes_{i=1}^{k-1} \mathbb{V}_i \otimes \mathbb{V}_k \otimes \bigotimes_{i=k+1}^{p} \mathbb{V}_i\right)$ *and* $Y \in \left(\bigotimes_{i=1}^{l-1} \mathbb{W}_i \otimes \mathbb{W}_l^* \otimes \bigotimes_{i=l+1}^{q} \mathbb{W}_i\right)$ *with* $\mathbb{V}_k = \mathbb{W}_l$*. We define the contraction between* X *and* Y *by the indices* (k, l) *as*

$$X \times_k^l Y := \mathcal{C}_k^l(X \otimes Y)$$

Fixing bases for $\mathbb{V}_1, \ldots, \mathbb{V}_p, \mathbb{W}_1, \ldots, \mathbb{W}_q$ and representing X, Y as discrete functions by its representations in those basis, we get a way for computing $\mathcal{C}_k^l(X \otimes Y)$:

Definition 2.19 (Base dependant tensor contraction)**.** [13] Given $\mathcal{X}, \mathcal{Y}$ as the discrete function representations of X, Y from earlier, we can write $\mathcal{C}_k^l(X \otimes Y)$ element-wise as:

$$\begin{aligned}
&\mathcal{C}_k^l(i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_p, j_1, \ldots, j_{l-1}, j_{l+1}, \ldots, j_q) \\
&= \sum_{s=1}^{N_k} \mathcal{X}(i_1, \cdots, i_{k-1}, s, i_{k+1}, \cdots, i_p) \mathcal{Y}(j_1, \cdots, j_{l-1}, s, j_{l+1}, \cdots, j_q)
\end{aligned}$$

Since the notation for making a single contraction is already very tedious to write, from now on we will use the Penrose notation.

2.1 Penrose notation

The Penrose an intuitive graphical language to represent tensor contractions that dates back from at least the early 1970s [5]

Given an nth-order tensor $\mathcal{T} \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ the way we draw it using the Penrose notation is as a circle with as many edges as the order of the tensor, as seen in Fig. 2.1

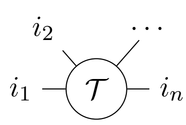

Figure 2.1: Representation of a tensor $\mathcal{T} \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ using the Penrose notation

A lot of times we will not write explicitly the name of the indexes, since we only care about their order. The order of the indexes will be determined by their orientation respect to the circle: to get the order we will start from the left and then following a clockwise rotation. The order in which we encounter the edges will be the order of the indexes. For example, in Fig. 2.1, the order would be $i_1, i_2, \ldots, i_n$

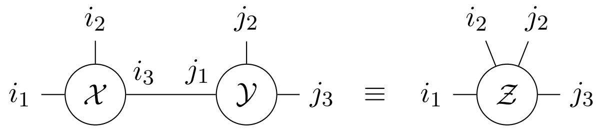

Figure 2.2: Representation in the Penrose notation of the contraction between two tensors $\mathcal{X} \in \mathbb{K}^{N_1 \times N_2 \times N_3}, \mathcal{Y} \in \mathbb{K}^{M_1 \times M_2 \times M_3}$ by their indices i_3 and j_1 with $N_1, N_2, N_3, M_1, M_2, M_3 \in \mathbb{N}$ and $N_3 = M_1$

We represent the contraction between two tensors $\mathcal{X} \in \mathbb{K}^{N_1 \times \cdots \times N_n}, \mathcal{Y} \in \mathbb{K}^{M_1 \times \cdots \times M_m}$ as their representation in the Penrose notation with the edges that represent the indexes that are contracting by joining them. For example, the contraction shown in Fig. 2.2, results in the tensor $\mathcal{Z} = \mathcal{X} \times_3^1 \mathcal{Y} \in \mathbb{K}^{N_1 \times N_2 \times M_2 \times M_3}$

Chapter 3

Tensor Networks

The concept of tensor networks originated from a physics background. Roger Penrose described how its diagrammatic language could be used in various applications of physics [5].

Later, in 1992, Steven R. White developed de Density Matrix Renormalization Group (DRMG) algorithm for quantum lattice systems. It was considered the first successfull tensor network application [8].

The commonly use terminology "tensor decomposition" (TD) is equivalent to "tensor network" to some extent. After several years of progress accross different research fields, there is no significant distinction between tensor decomposition and tensor networks. TD was employed primarily in signal processing fields [7]. Traditional TD models can be viewed as basic kinds of tensor networks. In this thesis we will study some of the properties of different tensor decomposition methods, and their effectivity.

In this chapter, we will define a mathematic definition of a tensor decomposition based on [10], we will see some common examples of tensor decompositions and we will define the tensor network structure space.

3.1 Tensor networks and tensor network states

Informally, the way we construct a tensor network consists of picking a directed graph $G = (V, \bar{E})$, and for each vertex $i \in V$ we assign a vector space $\mathbb{V}_i$ and for each edge $(i, j) \in \bar{E}$ we assign a vector space $\mathbb{E}_i$ to the tail of the edge and its dual covector space $\mathbb{E}_i^*$ to the head of the edge. We will also demand that the graph G is connected.

More formally, let $\mathbb{V}_1, \ldots, \mathbb{V}_d$ be vector spaces with $\dim \mathbb{V}_i = N_i, i = 1, \ldots, d$. Let $\mathbb{E}_1, \ldots, \mathbb{E}_c$ be finite vector spaces with $\dim \mathbb{E}_i = R_i, i = 1, \ldots, c$. For each $i \in V$ we associate the tensor product space

$$\left(\bigotimes\nolimits_{j\in \mathrm{IN}(i)} \mathbb{E}_j\right) \otimes \mathbb{V}_i \otimes \left(\bigotimes\nolimits_{j\in \mathrm{OUT}(i)} \mathbb{E}_j^*\right)$$

and a contraction map κ_G defined by contracting factors in $\mathbb{E}_j$ with factors of $\mathbb{E}_j^*$

$$\kappa_G : \bigotimes\nolimits_{i=1}^{d} \left[\left(\bigotimes\nolimits_{j\in \mathrm{IN}(i)} \mathbb{E}_j\right) \otimes \mathbb{V}_i \otimes \left(\bigotimes\nolimits_{j\in \mathrm{OUT}(i)} \mathbb{E}_j^*\right)\right] \to \bigotimes\nolimits_{i=1}^{d} \mathbb{V}_i$$

Since every directed edge (i, j) must point out of a vertex i and point into a vetex j, each copy of $\mathbb{E}_j$ is paired with one copy of $\mathbb{E}^*$, so the contraction κ_G is well defined.

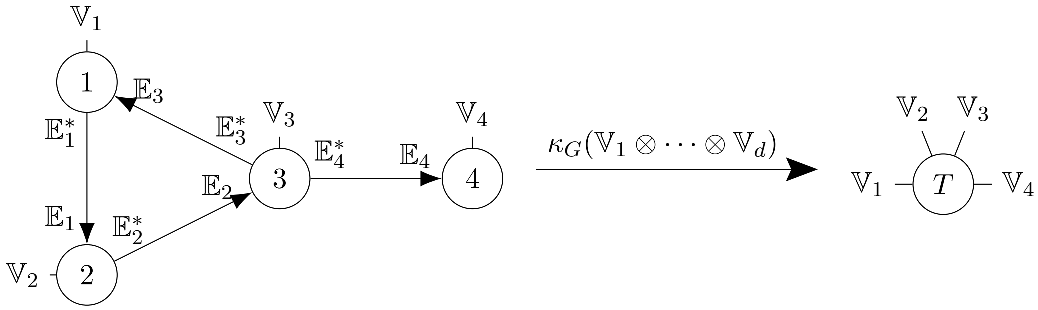

Figure 3.1: Example of the correspondance between some directed graph G and the vector spaces $\mathbb{V}_1, \ldots, \mathbb{V}_d, \mathbb{E}_1, \ldots, \mathbb{E}_c, \mathbb{E}_1^*, \ldots, \mathbb{E}_c^*$ using the Penrose notation

Definition 3.1 (Tensor network state)**.** *If a tensor $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ can be written as $T = \kappa_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n)$ where*

$$\mathcal{G}_i \in \left(\bigotimes\nolimits_{j \in \mathrm{IN}(i)} \mathbb{E}_j\right) \otimes \mathbb{V}_i \otimes \left(\bigotimes\nolimits_{j \in \mathrm{OUT}(i)} \mathbb{E}_j^*\right)$$

then we will say that T is a ***tensor network state*** *associated to the graph G with cores $\mathcal{G}_i \in \mathbb{V}_i, i = 1, \ldots, n$*

Definition 3.2 (Tensor network)**.** *We will define all the resulting tensors that are possible by varying $\mathcal{G}_1, \ldots, \mathcal{G}_n$ and then contracting through κ_G as the tensor network associated to G and the vector spaces $\mathbb{V}_1, \ldots, \mathbb{V}_n, \mathbb{E}_1, \ldots, \mathbb{E}_c$ and we will write this set as $\mathrm{TNS}(G; \mathbb{E}_1, \ldots, \mathbb{E}_c, \mathbb{V}_1, \ldots, \mathbb{V}_n)$, i.e*

$$\mathrm{TNS}(G; \mathbb{E}_1, \ldots, \mathbb{E}_c, \mathbb{V}_1, \ldots, \mathbb{V}_n) := \Bigg\{ \kappa_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n) \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n : \\ \mathcal{G}_i \in \left(\bigotimes\nolimits_{j \in \mathrm{IN}(i)} \mathbb{E}_j\right) \otimes \mathbb{V}_i \otimes \left(\bigotimes\nolimits_{j \in \mathrm{OUT}(i)} \mathbb{E}_j^*\right) \Bigg\}$$

Now, since all vector spaces are determined up to isomorphism by its dimension, when the vector spaces $\mathbb{E}_1, \ldots, \mathbb{E}_c, \mathbb{V}_1, \ldots, \mathbb{V}_n$ are unimportant, we will write the tensor network as $\mathrm{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_n)$. We will also make this substitution on the representation of the tensor network in the Penrose notation.

Now we will give some examples of common tensor network structures:

3.2 Common Tensor network structures

Example 3.3. [Tensor Train decomposition] [4] Let $\mathcal{T} \in \mathbb{R}^{N_1 \times \cdots \times N_n}$. A tensor train decomposition or matrix product state of $\mathcal{T}$ are a set of 3th-order tensors $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$ with $\mathcal{G}_i \in \mathbb{R}^{R_{i-1} \times N_i \times R_i}$ and $R_0 = R_n = 1$ such that every element of $\mathcal{T}$ is written in the form

$$\mathcal{T}(i_1, i_2, \ldots, i_n) = \sum_{r_0, \ldots, r_n}^{R_0, \ldots, R_n} \mathcal{G}(r_0, i_1, r_1)\mathcal{G}(r_1, i_2, r_2) \cdots \mathcal{G}(r_{n-1}, i_n, r_n) \tag{3.2.1}$$

We denote $R_0, R_1, \ldots, R_n$ as the ranks of the tensor train decomposition, or TT-ranks.

We can easily see that the tensor train decomposition (or TT) is obtained by our definition of a tensor network when G is a path, also the contraction of the whole network yields (3.2.1)

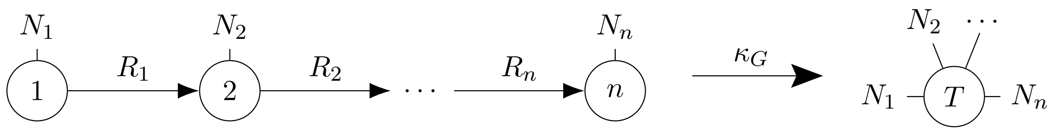

Figure 3.2: Tensor Train decomposition

Example 3.4. [Tensor Ring decomposition] [12] Tensor ring decomposition (or TR) or also known a matrix product state with periodic boundary conditions, is obtained when G is a cycle.

Tensor Ring decomposition is considered generalization of Tensor Train decomposition, it's contraction is the same as (3.2.1) but removing the condition $R_0 = R_1 = 1$.

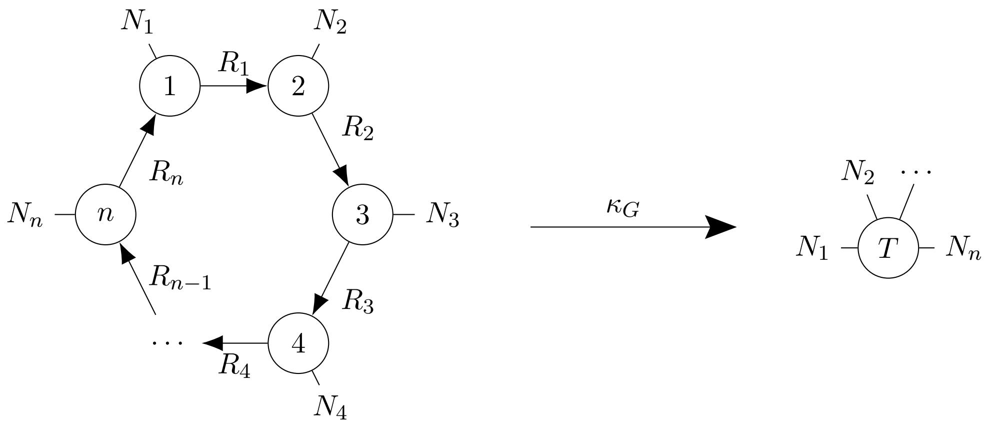

Figure 3.3: Tensor Ring (TR) decomposition

Theorem 3.5 (Circular dimensional permutation invariance)**.** *Let $\mathcal{T} \in \mathbb{R}^{N_1 \times \cdots \times N_n}$ be a nth-order tensor with its corresponding tensor ring decomposition $\mathcal{T} = \mathcal{R}(\mathcal{U}^{(1)}, \mathcal{U}^{(2)}, \ldots, \mathcal{U}^{(n)})$*

Example 3.6. [Fully connected tensor network] The fully connected tensor network decomposition is obtenied when G is a complete graph.

Figure 3.4: Fully connected tensor network decomposition (FCTN)

3.3 Approximating tensors to tensor network states

Now, one natural question that might occur is, given some tensor $T \in \mathbb{V}_1 \otimes \cdots \times \mathbb{V}_n$ and a tensor network $X = \mathrm{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_n)$, can we ensure that $T \in X$? (i.e T is an state of X) And if not, given all $\mathcal{X} \in X$, can we find an $\mathcal{X}_0 \in X$ such that $\|T - \mathcal{X}_0\|_F$ is minimal?

The following theorem gives us that by some $R_1, \ldots, R_c$, every tensor T can be a state of X:

Theorem 3.7. [10] Let $T \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_n$ and let G be a connected graph with n vertices and c edges. There exists $R_1, \ldots, R_c \in \mathbb{N}$ such that

$$T \in \mathrm{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_d)$$

in fact, we can choose $R_1 = \cdots = R_c = \operatorname{rank} T$

Proof. Let $r = \operatorname{rank} T$. Then there exist $v_1^{(i)}, \ldots, v_r^{(i)} \in \mathbb{V}_i, i = 1, \ldots, n$ such that

$$T = \sum_{p=1}^{r} v_1^{(p)} \otimes \cdots \otimes v_n^{(p)}$$

We take $R_1 = \cdots = R_c = r$ we take for each $i = 1, \ldots, n$

$$\mathcal{G}_i = \sum_{p=1}^{r} \left(\bigotimes\nolimits_{j \in \mathrm{IN}(i)} e_p^{(j)} \right) \otimes v_p^{(i)} \otimes \left(\bigotimes\nolimits_{j \in \mathrm{OUT}(i)} e_p^{(j)*} \right)$$

Now observe that for each $i = 1, \ldots, n$ there exists an unique h such that whenever $j \in \mathrm{IN}(i) \cap \mathrm{OUT}(i)$, $e_p^{(j)}$ and $e_p^{(j)*}$ contract and give δ_{pq}, therefore the summand vanishes except when $p = q$. This together with the assumption that G is a connected graph implies that $\kappa_G(\mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_n)$ reduces to a sum of terms of the form $v_p^{(1)} \otimes \cdots \otimes v_p^{(d)}$ for $p = 1, \ldots, r$, which is of course T $\square$

This theorem serves as an upper bound for our problem. Suppose that $T \in \mathbb{K}^{N_1 \times \cdots \times N_n}$ and $\operatorname{rank} T = \prod_{i=1}^{n} N_i$. This theorem says that given any connected directed graph G we can express of We will see that in fact, we can find some graphs G where $T \in \mathrm{TNS}(G; R_1, \ldots, R_c, N_1, \ldots, N_d)$ with $R_1, \ldots, R_c < \operatorname{rank} T$

From now on, suppose that $G, R_1, \ldots, R_c$ are fixed. Now we will discuss some algorithms for finding cores such that, when contracting, approximates T as best as possible, that means finding

$$\min_{\mathcal{G}_1, \ldots, \mathcal{G}_n} \|T - \kappa_G(\mathcal{G}_1, \ldots, \mathcal{G}_n)\|_F$$

3.4 The tensor tensor network structure space

3.5 Finding the best structure

TODO:

- Descriure G-ranks
- Algorismes per aproximar TNS per G-ranks propers i mínims si es pot fer
- Algun algorisme per trobar heuristicament els G-ranks adequats? (suposo q depen de compressió ratio i l'error relatiu)
- Com podem trobar un G adequat?
- Estratègies per contraure tensors més ràpidament? (DRMG?)
- Algorismes, part pràctica en C/C++
- Fer moltes gràfiques

- Fer aplicacions per machine learning, etc.
- Fixar la mathematical subject classification

Chapter 4

Conclusions

TODO

Fent servir un símil geomètrico-cartogràfic, aquesta memòria constitueix un mapa a escala planetària de la demostració de la conjectura feble de Goldbach presentada per Helfgott i un mapa a escala continental de la verificació numèrica d'aquesta. Estudis posteriors i més profunds haurien de permetre elaborar mapes de menor escala.

La naturalesa dels nombres primers ens ha portat per molts racons diferents de les Matemàtiques; en no imposar-nos restriccions en la forma de pensar, hem pogut gaudir del viatge i assolir els objectius que ens vam plantejar a l'inici del projecte i anar més enllà, sobretot en el camp de la computació i la manipulació de grans volums de dades numèriques.

Una gran part dels coneixements bàsics que hem hagut de fer servir han estat treballats en les assignatures de Mètodes analítics en teoria de nombres i d'Anàlisi harmònica i teoria del senyal, que són optatives de quart curs del Grau de Matemàtiques. Altres els hem hagut d'aprendre durant el desenvolupant del projecte. S'ha realitzat una tasca de recerca bibliogràfica important, consultant recursos antics i moderns, tant en format digital com en format paper.

Bibliography

[1] Christopher Terence John Dodson and Timothy Poston. "Tensors and Multilinear Forms". In: *Tensor Geometry: The Geometric Viewpoint and Its Uses*. Ed. by Christopher Terence John Dodson and Timothy Poston. Berlin, Heidelberg: Springer, 1991, pp. 98–113. ISBN: 978-3-642-10514-2. DOI: 10.1007/978-3-642-10514-2_6. URL: https://doi.org/10.1007/978-3-642-10514-2_6 (visited on 03/28/2025).

[2] Christopher Hillar and Lek-Heng Lim. *Most Tensor Problems Are NP-hard*. July 1, 2013. DOI: 10.48550/arXiv.0911.1393. arXiv: 0911.1393 [cs]. URL: http://arxiv.org/abs/0911.1393 (visited on 03/31/2025). Pre-published.

[3] Chao Li et al. *Permutation Search of Tensor Network Structures via Local Sampling*. June 14, 2022. DOI: 10.48550/arXiv.2206.06597. arXiv: 2206.06597 [cs]. URL: http://arxiv.org/abs/2206.06597 (visited on 03/24/2025). Pre-published.

[4] I. V. Oseledets. "Tensor-Train Decomposition". In: *SIAM Journal on Scientific Computing* 33.5 (Jan. 2011), pp. 2295–2317. ISSN: 1064-8275. DOI: 10.1137/090752286. URL: https://epubs.siam.org/doi/10.1137/090752286 (visited on 03/24/2025).

[5] Roger Penrose. "Applications of Negative Dimensional Tensors". In: *Combinatorial Mathematics and its Applications* (1971), pp. 221–244. URL: https://homepages.math.uic.edu/~kauffman/Penrose.pdf (visited on 03/04/2025).

[6] Steven Roman. "Tensor Products". In: *Advanced Linear Algebra*. Ed. by Steven Roman. New York, NY: Springer, 2008, pp. 355–409. ISBN: 978-0-387-72831-5. DOI: 10.1007/978-0-387-72831-5_14. URL: https://doi.org/10.1007/978-0-387-72831-5_14 (visited on 03/25/2025).

[7] Maolin Wang et al. *Tensor Networks Meet Neural Networks: A Survey and Future Perspectives*. May 8, 2023. DOI: 10.48550/arXiv.2302.09019. arXiv: 2302.09019 [cs]. URL: http://arxiv.org/abs/2302.09019 (visited on 03/09/2025). Pre-published.

[8] Steven R. White. "Density Matrix Formulation for Quantum Renormalization Groups". In: *Physical Review Letters* 69.19 (Nov. 9, 1992), pp. 2863–2866. DOI: 10.1103/PhysRevLett.69.2863. URL: https://link.aps.org/doi/10.1103/PhysRevLett.69.2863 (visited on 03/27/2025).

[9] Robin J. Wilson. *Introduction to Graph Theory*. 4. ed., [Nachdr.] Harlow Munich: Prentice Hall, 2009. 171 pp. ISBN: 978-0-582-24993-6.

[10] Ke Ye and Lek-Heng Lim. *Tensor Network Ranks*. Feb. 9, 2019. DOI: 10.48550/arXiv.1801.02662. arXiv: 1801.02662 [math]. URL: http://arxiv.org/abs/1801.02662 (visited on 03/24/2025). Pre-published.

[11] Tatsuya Yokota. *Very Basics of Tensors with Graphical Notations: Unfolding, Calculations, and Decompositions.* Nov. 25, 2024. DOI: 10.48550/arXiv.2411.16094. arXiv: 2411.16094 [cs]. URL: http://arxiv.org/abs/2411.16094 (visited on 02/24/2025). Pre-published.

[12] Qibin Zhao et al. *Tensor Ring Decomposition.* June 17, 2016. DOI: 10.48550/arXiv.1606.05535. arXiv: 1606.05535 [cs]. URL: http://arxiv.org/abs/1606.05535 (visited on 03/04/2025). Pre-published.

[13] Yu-Bang Zheng et al. "Fully-Connected Tensor Network Decomposition and Its Application to Higher-Order Tensor Completion". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 35.12 (12 May 18, 2021), pp. 11071–11078. ISSN: 2374-3468. DOI: 10.1609/aaai.v35i12.17321. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17321 (visited on 03/04/2025).

Appendix A

Chapter 1