# The Set of Orthogonal Tensor Trains

**Pardis Semnani** and Elina Robeva

Department of Mathematics, University of British Columbia

## Abstract

In this paper we study the set of tensors that admit a special type of decomposition called an orthogonal tensor train decomposition. Finding equations defining varieties of low-rank tensors is generally a hard problem, however, the set of orthogonally decomposable tensors is defined by appealing quadratic equations. The tensors we consider are an extension of orthogonally decomposable tensors, and we show that they are defined by similar quadratic equations, as well as some interesting additional equations.

## Introduction

- With the emergence of big data, it is more and more often the case that information is recorded in the form of a tensor (or multi-dimensional array).
- The importance of decomposing such a tensor is as follows:
  - It provides hidden information about the data at hand.
  - Having a concise decomposition of the tensor allows us to store it much more efficiently.
- Decomposing tensors is often computationally hard. For example, finding (the number of terms of) the CP-decomposition of a general tensor is NP-hard, and the set of tensors of CP rank at most $r$ is not closed for any $r \geq 2$, making the low-rank approximation problem impossible in some instances. It is also notoriously hard to describe all the defining equations of the set of tensors of rank at most $r$.
- On the other hand, orthogonally decomposable tensors can be decomposed efficiently via the slice method or via the tensor power method. Furthermore, the set of orthogonally decomposable tensors of bounded rank is closed, making the family of such tensors as appealing as the set of matrices.
- Tensor networks provide many other ways of decomposing tensors. They originate from quantum physics and are used to depict the structure of steady states of Hamiltonians of quantum systems. Many types of tensor network decompositions, like tensor trains, also known as matrix product states, are used in machine learning to decompose data tensors in meaningful ways.

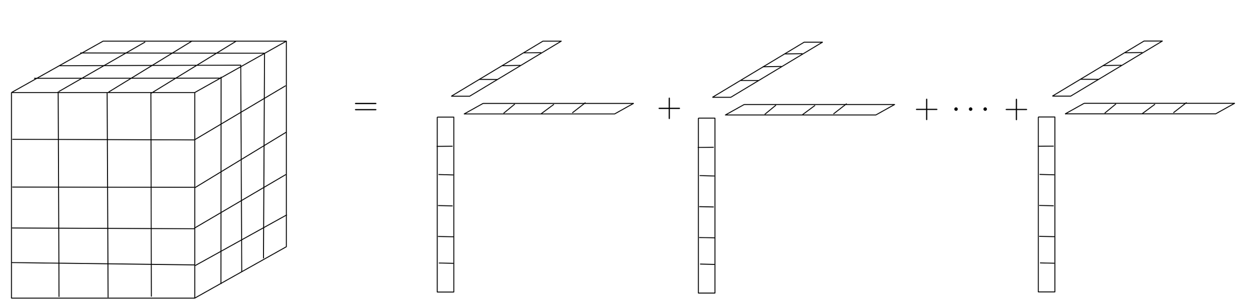


Figure 1: CP-decomposition of a tensor

## Contraction of Two Tensors

Let $T \in \mathbb{R}^{m_1} \otimes \cdots \otimes \mathbb{R}^{m_k}$ and $S \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_\ell}$ be two tensors such that $m_i = n_j$. The *contraction* of the tensors $T$ and $S$ along their $i$th and $j$th modes is the tensor $R \in \mathbb{R}^{m_1} \otimes \cdots \otimes \mathbb{R}^{m_{i-1}} \otimes \mathbb{R}^{m_{i+1}} \otimes \cdots \otimes \mathbb{R}^{m_k} \otimes \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_{j-1}} \otimes \mathbb{R}^{n_{j+1}} \otimes \cdots \otimes \mathbb{R}^{n_\ell}$ whose entries are

$$R_{a_1 \ldots a_{i-1} a_{i+1} \ldots a_k b_1 \ldots b_{j-1} b_{j+1} \ldots b_\ell} = \sum_{c=1}^{m_i} T_{a_1 \ldots a_{i-1} c a_{i+1} \ldots a_k} S_{b_1 \ldots b_{j-1} c b_{j+1} \ldots b_\ell}.$$

## Tensor Networks

Consider a graph $G = (V, E, D)$ where

- $V$ is the set of vertices.
- $E$ is the set of edges adjacent to 2 vertices.
- $D$ is the set of edges adjacent to just 1 vertex (also called *dangling* edges).
- Each edge in $e \in E \cup D$ is assigned a positive integer $n_e$.
- Each vertex $v \in V$ is assigned a tensor $T_v \in \otimes_{v \in e} \mathbb{R}^{n_e}$.
- The resulting *tensor network* state $T \in \otimes_{e \in D} \mathbb{R}^{n_e}$ is obtained by contracting the tensors $T_v$ along all edges $e \in E$.

For instance, in the graph below,

- $V = \{1, 2\}$.
- The only member of $E$ is the dashed edge.
- $D$ consists of blue and red edges.
- $T_1 \in \mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2} \otimes \mathbb{R}^{n_3} \otimes \mathbb{R}^{n_6}$.
- $T_2 \in \mathbb{R}^{n_4} \otimes \mathbb{R}^{n_5} \otimes \mathbb{R}^{n_6}$.
- The graph represents the tensor obtained by the contraction of $T_1$ and $T_2$ along their 4th and 3rd modes respectively.

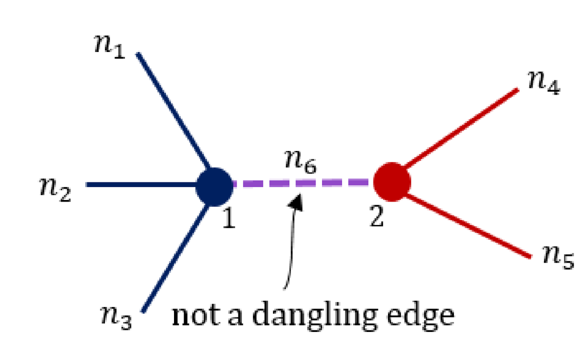


Figure 2: A tensor network

## Tensor Trains

- A *tensor train* of length $l$ is a tensor network consisting of $l$ vertices arranged in a line, each of which has 3 adjacent edges.

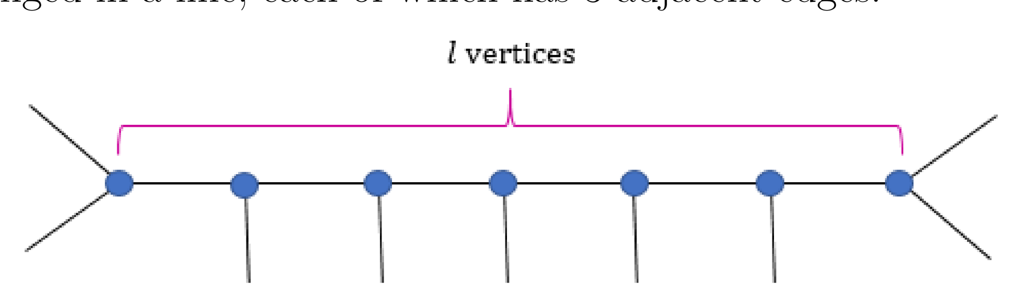


Figure 3: A tensor train of length $l$

- A tensor train of length $l$ is said to be *symmetrically orthogonally decomposable* (symmetrically odeco) if all the $l$ order-3 tensors generating the tensor train are symmetrically odeco.
- $SOT_{l,n}$ represents the set of symmetrically odeco tensor trains of length $l$ which are generated by $n \times n \times n$ tensors.

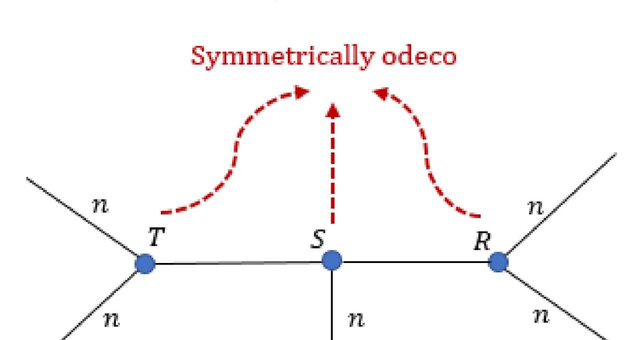


Figure 4: Symmetrically odeco tensor train of length 3 (A tensor in $SOT_{3,n}$)

- It has been shown in [1] that if a tensor train of length 2 has a symmetrically odeco decomposition, one can find the decomposition efficiently.

## Results

- Let $n \geq 2$ and let $\mathcal{P}_n$ denote the set of polynomials
$$f_{a,b,c,d,\sigma_1,\sigma_2} := p_{abcd} - p_{\sigma_1(a)\sigma_1(b)\sigma_2(c)\sigma_2(d)},$$
where $a, b, c, d \in \{1, \ldots, n\}$, $\sigma_1$ is a permutation on $\{a, b\}$, and $\sigma_2$ is a permutation on $\{c, d\}$. Then $\mathcal{P}_n$ vanishes on $SOT_{2,n}$.

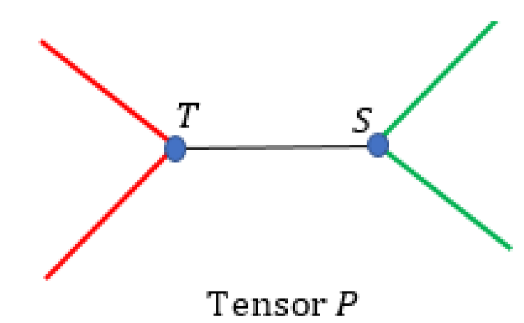


Figure 5: The fact that $\mathcal{P}_n$ vanishes on $SOT_{2,n}$ shows that $P$ is symmetric with respect to any of the two pairs of modes indicated in red and green.

- Let $n \geq 2$ and $\mathcal{Q}_n$ denote the set of polynomials
$$g^{(1)}_{a,b,c,d,e,f} := \sum_{t=1}^{n} p_{abet}\, p_{cdft} - p_{abft}\, p_{cdet},$$
and
$$g^{(2)}_{a,b,c,d,e,f} := \sum_{t=1}^{n} p_{etab}\, p_{ftcd} - p_{ftab}\, p_{etcd},$$
where $a, b, c, d, e, f \in \{1, \ldots, n\}$. Then $\mathcal{Q}_n$ vanishes on $SOT_{2,n}$.

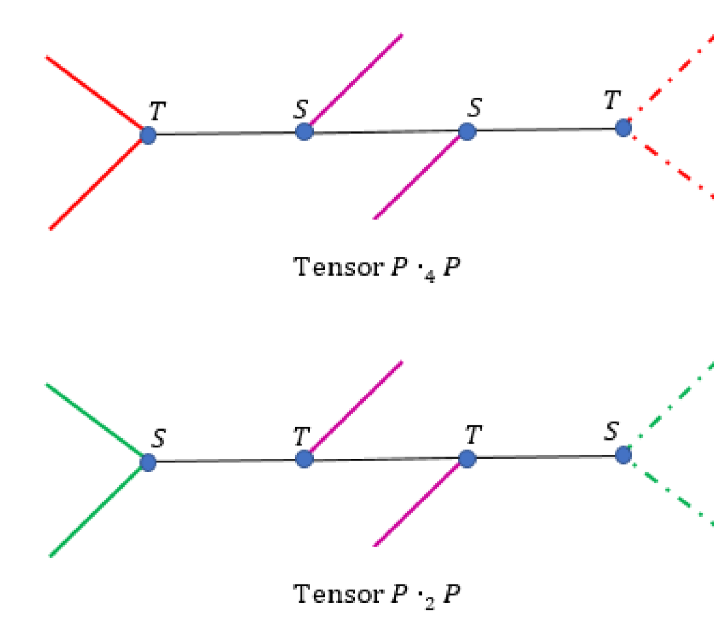


Figure 6: The fact that $\mathcal{Q}_n$ vanishes on $SOT_{2,n}$ shows that $P \bullet_4 P \in S^2(\mathbb{R}^n) \otimes S^2(\mathbb{R}^n) \otimes S^2(\mathbb{R}^n)$, where the first two $S^2(\mathbb{R}^n)$ correspond to the red edges and the red dashed edges respectively, and the last $S^2(\mathbb{R}^n)$ corresponds to the commuting of the indices coming from the purple edges and is the one that gives rise to half of the equations in $\mathcal{Q}_n$. Similarly, $P \bullet_2 P \in S^2(\mathbb{R}^n) \otimes S^2(\mathbb{R}^n) \otimes S^2(\mathbb{R}^n)$, where the first two $S^2(\mathbb{R}^n)$ correspond to the green and green dashed edges respectively, and the last $S^2(\mathbb{R}^n)$ corresponds to the fact that the indices along the purple edges commute and gives rise to the other half of the equations in $\mathcal{Q}_n$.

- The symmetries that the equations in $\mathcal{Q}_n$ depict for $P \bullet_2 P$ and $P \bullet_4 P$ are very similar to those that have been proved for symmetrically orthogonally decomposable and ordinary orthogonally decomposable tensors in [2].

## Results

- Let $n \geq 2$ and define
$$h_n := \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \sum_{\sigma \in S_n} \sum_{\gamma \in S_n} \operatorname{sgn}(\sigma) \operatorname{sgn}(\gamma)\, p_{k_1\sigma(1)k_n\gamma(1)}\, p_{k_2\sigma(2)k_1\gamma(2)} \cdots p_{k_n\sigma(n)k_{n-1}\gamma(n)}.$$
Then $h_n$ vanishes on $SOT_{2,n}$.
- If $n$ is even, then in fact, $h_n$ vanishes on the set of all tensor trains of length 2 in which the two tensors at the vertices are symmetric and have rank at most $n$.
- When $n \leq 5$, we are able to computationally show that $h_n$ does not lie in the ideal generated by $\mathcal{P}_n \cup \mathcal{Q}_n$.
- For all $n \geq 2$, the dimension of the Zariski closure of $SOT_{2,n}$ equals $n(n+1) - 1$. This follows from counting the dimensions of the parametrization of $SOT_{2,n}$ and from noting that the decomposition of a tensor in $SOT_{2,n}$ can be found uniquely in a sense according to [1] .
- When $n = 2$, the dimension of the ideal cut out by $\mathcal{P}_n$, $\mathcal{Q}_n$, and the polynomial $h_n$ equals $5 = n(n+1) - 1$, which was checked numerically using Macaulay2 and Maple.

## Conjecture

We conjecture that the Zariski closure of the set $SOT_{2,n}$ is cut out by the vanishing of $\mathcal{P}_n, \mathcal{Q}_n$, and the polynomial $h_n$ for $n \geq 2$, and that the ideal defined by these equations is prime.

## References

[1] K. Halaseh, T. Muller, and E. Robeva.
Orthogonal decomposition of tensor trains.
*arXiv:2010.04202*, 2020.

[2] A. Boralevi, J. Draisma, E. Horobet, and E. Robeva.
Orthogonal and unitary tensor decomposition from an algebraic perspective.
*Israel Journal of Mathematics*, 222(1):223–260, 2017.

[3] C. Hillar and L.-H. Lim.
Most tensor problems are NP-hard.
*Journal of the ACM*, 60(6), 2013.

[4] A. Anandkumar, R. Ge, D. Hsu, S. Kakade, and M. Telgarsky.
Tensor decompositions for learning latent variable models.
*Journal of Machine Learning Research*, 15(80):2773–2832, 2014.

## Contact Information

- Pardis Semnani: psemnani@math.ubc.ca
- Elina Robeva: erobeva@math.ubc.ca
- You can read our paper here.