

## DECOMPOSING TENSORS VIA RANK-ONE APPROXIMATIONS\*

ÁLVARO RIBOT<sup>†</sup>, EMIL HOROBET<sup>‡</sup>, ANNA SEIGAL<sup>†</sup>, AND ETTORE T. TURATTI<sup>§</sup>

**Abstract.** Matrices can be decomposed via rank-one approximations: the best rank-one approximation is a singular vector pair, and the singular value decomposition writes a matrix as a sum of singular vector pairs. The singular vector tuples of a tensor are the critical points of its best rank-one approximation problem. In this paper, we study tensors that can be decomposed via successive rank-one approximations: compute a singular vector tuple, subtract it off, compute a singular vector tuple of the new deflated tensor, and repeat. The number of terms in such a decomposition may exceed the tensor rank. Moreover, the decomposition may depend on the order in which terms are subtracted. We show that the decomposition is valid independent of order if and only if all singular vectors in the process are orthogonal in at least two factors. We study the variety of such tensors. We lower bound its dimension, showing that it is significantly larger than the variety of odecō tensors.

**Key words.** tensor decomposition, two-orthogonal tensors, singular vector tuples, odecō

**MSC codes.** 14N07, 15A18, 15A69

**DOI.** 10.1137/24M1715106

**1. Introduction.** For matrices, critical rank-one approximations combine to give the singular value decomposition: a matrix  $M$  can be decomposed via successive critical rank-one approximations as  $M - \lambda_1 u_1 \otimes v_1 - \cdots - \lambda_r u_r \otimes v_r = 0$ , where  $r = \text{rank}(M)$ . Each term  $\lambda_i u_i \otimes v_i$  is a critical rank-one approximation of both the original matrix  $M$  and the deflated matrix  $M - \sum_{j < i} \lambda_j u_j \otimes v_j$ . According to the Schmidt–Eckart–Young theorem [10], truncating to the  $k$  largest  $\lambda_i$ 's gives the best rank- $k$  approximation of  $M$ . Thus, the best rank- $k$  approximation problem reduces to iteratively solving best rank-one approximation problems.

Low-rank approximation of tensors can also be computed via rank-one updates. Though this approach may not lead to the best low-rank approximation [20, 31], it has proven successful in practice [38, 30, 1]. Numerically, a rank-one approximation can be computed efficiently with the power method and its variants [6, 21], for which there are convergence guarantees [34].

For tensors of order 3 or higher, a best rank- $r$  approximation may not even exist for  $r > 1$  [7], but a best rank-one approximation always exists because the set of rank-one tensors is closed. However, note that computing a best rank-one approximation

---

\*Received by the editors December 4, 2024; accepted for publication (in revised form) by A. Bernardi September 18, 2025; published electronically February 19, 2026.

<https://doi.org/10.1137/24M1715106>

**Funding:** The first author was supported by a fellowship from “la Caixa” Foundation (100010434), with fellowship LCF/BQ/EU23/12010097, and by an RCCHU fellowship. The second author was supported by the Project “Singularities and Applications” - CF 132/31.07.2023 funded by the European Union - NextGenerationEU - through Romania’s National Recovery and Resilience Plan. The third author was partially supported by the NSF (DMR-2011754). The fourth author was partially supported by the project Pure Mathematics in Norway, funded by Trond Mohn Foundation and Tromsø Research Foundation.

<sup>†</sup>School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (aribottarrado@g.harvard.edu, aseigal@seas.harvard.edu).

<sup>‡</sup>Mathematics and Computer Science, Sapientia Hungarian University of Transylvania, Targu-Mures, (Corunca), Mures, 547367 Hungary (horobetemil@ms.sapientia.ro).

<sup>§</sup>Department of Mathematics and Statistics, UIT - The Arctic University of Norway, Tromsø, 9037 Norway (ettore.t.turatti@uit.no).

---

**Algorithm 1.1.** Decompose  $\mathcal{T}$  via rank-one approximations.

---

**Input:** Tensor  $\mathcal{T} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ 

- 1:  $i = 1$
- 2:  $\mathcal{S} = \mathcal{T}$
- 3: **while**  $\mathcal{S} \neq 0$  **do**
- 4:     Compute a critical rank-one approximation  $x_i = x_i^{(1)} \otimes \cdots \otimes x_i^{(d)}$  of  $\mathcal{S}$
- 5:     Deflate  $\mathcal{S} \leftarrow \mathcal{S} - x_i$
- 6:      $i \leftarrow i + 1$
- 7: **end while**

**Output:** Decomposition  $\mathcal{T} = \sum_i x_i$ .

---

is NP-hard [17]. A global approach to computing a best rank-one approximation of a tensor is to compute the critical points of the distance function to the variety of rank-one tensors and to choose a critical point whose residual is smallest. A generic tensor has a finite number of such critical points [12]. These are the singular vector tuples of the tensor [25]. We use the terms critical rank-one approximation and singular vector tuple interchangeably.

A minimal decomposition via rank-one approximations is known to exist only on a measure zero set; see the Schmidt–Eckart–Young decomposition in [35]. In this article, rather than producing a minimal decomposition via rank-one approximations, we are interested in tensors that are a finite sum of critical rank-one approximations. The decomposition need not be of minimal length. That is, we seek those tensors  $\mathcal{T}$  for which there exists a decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$  such that each  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i < j} x_i$ . These are the tensors for which the following algorithm terminates for some rank-one approximations  $x_i$ .

The varieties of tensors that admit such a decomposition are data loci in the sense of [18]. We denote by  $\text{DL}_r$  the variety of tensors that admit such a decomposition of length  $r$ . We obtain a chain of varieties  $\text{DL}_1 \subseteq \text{DL}_2 \subseteq \cdots$ , such that any tensor  $\mathcal{T} \in \text{DL}_r$  has a rank-one approximation  $x$  with  $\mathcal{T} - x \in \text{DL}_{r-1}$ . For symmetric tensors, this chain of varieties stabilizes and the limit is the variety of weakly odec tensors [19]. For matrices,  $\text{DL}_r$  is the variety of matrices of rank at most  $r$ , so the chain also stabilizes, filling the ambient space. In general, the rank-one tensor  $x_i$  need not be a singular vector tuple of  $\mathcal{T}$  for  $i \geq 2$ . That is, the decomposition is order dependent.

*Example 1.1.* Let  $\{e_0, e_1\}$  be an orthonormal basis of  $\mathbb{R}^2$ , and fix

$$\mathcal{T} = \underbrace{e_1 \otimes (e_0 + e_1) \otimes (e_0 - e_1)}_{x_1} + \underbrace{e_0 \otimes e_1 \otimes e_1}_{x_2} + \underbrace{e_0 \otimes e_0 \otimes e_0}_{x_3} \in (\mathbb{R}^2)^{\otimes 3}.$$

We observe that  $\mathcal{T} \in \text{DL}_3$ , and that  $x_1$  is a critical rank-one approximation of  $\mathcal{T}$ , but  $x_2$  and  $x_3$  are not critical rank-one approximations of  $\mathcal{T}$ .

We consider subvarieties of  $\text{DL}_r$  consisting of tensors for which the decomposition does not depend on the order. These are the tensors  $\mathcal{T} = \sum_{i=1}^r x_i$  in which, for any order of the summands,  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i < j} x_i$ . In particular, every  $x_i$  is a critical rank-one approximation of the original tensor  $\mathcal{T}$ . Put differently, for these tensors Algorithm 1.1 terminates with  $\mathcal{S}$  replaced by  $\mathcal{T}$  in line 4. While a generic tensor lies in the span of its singular vector tuples [9], one does not know the coefficient of each singular vector tuple, i.e., how much of each one to subtract off. For the tensors we consider, the coefficient of each singular vector tuple is its singular value, so the tensor is a sum of critical rank-one approximations.

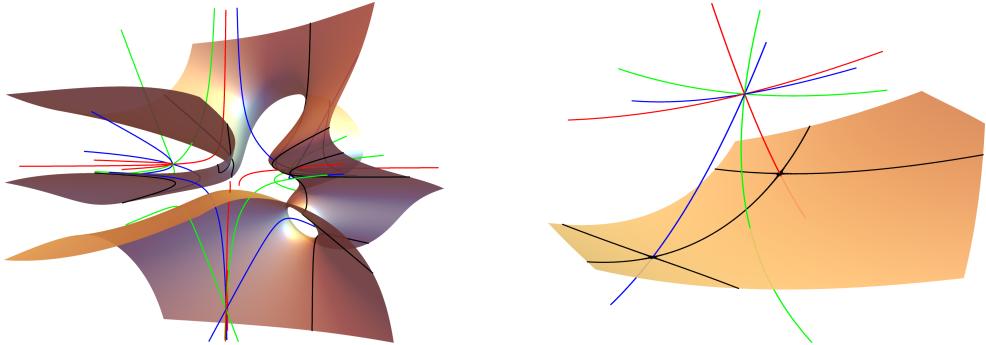


FIG. 1. The affine slice of the two-orthogonal variety in  $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$  when  $t_{01} = 1, t_{100} = 2t_{000}, t_{110} = 1, t_{101} = 2, t_{011} = 3$ . The image on the right is a zoomed-in and rescaled version of the image on the left. The surface is  $W_4 \setminus W_3$ . The black curve is  $W_3 \setminus W_2$ . The red, green, and blue curves are the three components of  $W_2$ . The intersection of these three curves is the odecovariety. These plots were made with *Mathematica* [37]. Note: color appears only in the online article.

One class of tensors known to possess such a property are the odecovariety (orthogonally decomposable) tensors, those that have a decomposition  $\mathcal{T} = \sum_{i=1}^r x_i^{(1)} \otimes \cdots \otimes x_i^{(d)}$  such that for all  $i \neq j$  we have the orthogonality  $x_i^{(k)} \perp x_j^{(k)}$  in all factors  $1 \leq k \leq d$  [28, 29]. We relax this notion by requiring only that for all  $i \neq j$ , the  $i$ th and  $j$ th summands are orthogonal in at least two factors. We call such a decomposition *two-orthogonal*. We define  $W_r$  to be the Zariski closure of the set of tensors that have a two-orthogonal decomposition of length at most  $r$ . The *two-orthogonal variety* is  $W = \overline{\bigcup_{r \geq 1} W_r}$ . See Figure 1.

The notion of a two-orthogonal tensor also appears in [35, Definition 3.2], where it is called weakly two-orthogonal. We drop “weakly” to avoid confusion with the weakly odecovariety tensors, which appear in [19]. Two-orthogonality is necessary for having a Schmidt–Eckart–Young decomposition [35, Theorem 3.3]. The singular value decomposition implies that every matrix is two-orthogonal. For symmetric decompositions, two-orthogonality implies orthogonality in all factors, so it leads to an odecovariety tensor. We will see that a generic  $2 \times 2 \times 2$  tensor is not two-orthogonal in section 5. We believe this is the general behavior for tensors of order at least three, as two-orthogonality imposes closed conditions on the singular vector tuples.

Now we state our main results. First, we show that the two-orthogonal tensors are the tensors that can be decomposed via rank-one approximations.

**THEOREM 1.2.** *The set of two-orthogonal tensors in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  coincides with the set of tensors with a decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$  such that each  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i \in \mathcal{I}} x_i$  for all  $\mathcal{I} \subseteq \{1, \dots, r\} \setminus \{j\}$ .*

Two-orthogonal decompositions cannot have arbitrarily many summands. We count the maximum number of possible summands in the following.

**THEOREM 1.3.** *The maximal length of a two-orthogonal decomposition in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  is  $N = \min_{1 \leq k \leq d} \prod_{j \neq k} n_j$ . In particular, the two-orthogonal variety is  $W = W_N$ .*

By constructing two-orthogonal decompositions of maximal length, we parametrize a family of two-orthogonal tensors with a dimension-preserving map. This gives a lower bound on the dimension of the two-orthogonal variety.

**THEOREM 1.4.** *The dimension of the two-orthogonal variety in  $(\mathbb{R}^n)^{\otimes d}$  is at least*

$$n^{d-1} + d \binom{n}{2}.$$

In  $V = (\mathbb{R}^n)^{\otimes d}$ , the odecovariety has dimension  $n + d \binom{n}{2}$ . Hence, the dimension of the two-orthogonal variety exceeds that of the odecovariety by at least  $n(n^{d-2} - 1)$ . Finally, we show that two-orthogonal decompositions are generically identifiable for  $2 \times 2 \times 2$  tensors. We suspect that this is also true for generic two-orthogonal tensors in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ .

**THEOREM 1.5.** *A generic tensor in  $W \subseteq (\mathbb{R}^2)^{\otimes 3}$  has a unique two-orthogonal decomposition.*

The article is organized as follows. In section 2 we study the main properties of two-orthogonal decompositions. We prove Theorems 1.2 and 1.3, we introduce the notion of two-orthogonal rank and how it relates to the usual rank, and we extend some of our results to partially symmetric tensors. We also show that truncating two-orthogonal decompositions may lead to best low-rank approximations and answer an open question posed in [35]. In section 3 we prove Theorem 1.4. In section 4 we provide a parametric description of the two-orthogonal variety using graphs, with a focus on binary tensors. There is an interplay between algebraic geometry and combinatorics, with open directions for future work. Finally, in section 5 we provide an algebraic description of the two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$  and prove Theorem 1.5.

**2. Two-orthogonal decompositions.** Let  $V = \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  with  $d \geq 2$  and each  $n_k \geq 2$ . Let  $X = \{v_1 \otimes \cdots \otimes v_d \mid v_k \in \mathbb{R}^{n_k}\} \subset V$  denote the cone over the Segre variety, the set of rank-one tensors. Let  $\langle \cdot, \cdot \rangle_k$  be nondegenerate bilinear forms in  $\mathbb{R}^{n_k}$  for  $k = 1, \dots, d$ . They induce a nondegenerate bilinear form in  $V$ , called the *Bombieri–Weyl inner product*, defined on two rank-one tensors as

$$\langle v_1 \otimes \cdots \otimes v_d, w_1 \otimes \cdots \otimes w_d \rangle = \prod_{k=1}^d \langle v_k, w_k \rangle_k$$

and extended to all of  $V$  by bilinearity. We write  $v_k \perp w_k$  to denote orthogonality:  $\langle v_k, w_k \rangle_k = 0$ . Given a positive integer  $n$ , let  $[n] = \{1, \dots, n\}$ , and let  $\{e_1, \dots, e_n\}$  be an orthonormal basis of  $\mathbb{R}^n$ . We use the same letters for the bases of different factors of  $V$ . That is, the set  $\{e_{i_1} \otimes \cdots \otimes e_{i_d} \mid i_k \in [n_k]\}$  is an orthonormal basis of  $V$ . Let  $\|\cdot\|$  denote the Frobenius norm on  $V$  induced by the inner product.

**DEFINITION 2.1.** *A two-orthogonal decomposition of  $\mathcal{T}$  is an expression  $\mathcal{T} = \sum_{i=1}^r x_i$ , where  $x_i = x_i^{(1)} \otimes \cdots \otimes x_i^{(d)} \neq 0$  and all pairs of summands  $x_i$  and  $x_j$  with  $i \neq j$  are orthogonal in at least two factors: there exist indices  $k_1 \neq k_2$  such that  $x_i^{(k_1)} \perp x_j^{(k_1)}$  and  $x_i^{(k_2)} \perp x_j^{(k_2)}$ . A tensor  $\mathcal{T}$  is two-orthogonal if it has a two-orthogonal decomposition for some  $r$ .*

**Remark 2.2.** The summands in a two-orthogonal decomposition are orthogonal to each other. Hence, a two-orthogonal decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$  has  $\|\mathcal{T}\|^2 = \sum_{i=1}^r \|x_i\|^2$ .

**DEFINITION 2.3.** *Given a tensor  $\mathcal{T} \in V$ , a rank-one tensor  $x = x^{(1)} \otimes \cdots \otimes x^{(d)}$  is a singular vector tuple of  $\mathcal{T}$  if for all  $k \in [d]$  there exists  $\lambda_k \in \mathbb{R}$  such that*

$$\mathcal{T}(x^{(1)}, \dots, x^{(k-1)}, \cdot, x^{(k+1)}, \dots, x^{(d)}) = \lambda_k x^{(k)},$$

where  $\mathcal{T}$  is viewed as a multilinear map and  $\mathbb{R}^{n_k}$  is identified with its dual  $(\mathbb{R}^{n_k})^*$  via  $\langle \cdot, \cdot \rangle_k$ . If  $\langle x^{(k)}, x^{(k)} \rangle_k = 1$  for all  $k$ , then  $\lambda_k = \lambda$  for all  $k$ , and  $\lambda$  is called the singular value of  $x$ .

*Remark 2.4.* We consider singular vector tuples defined over  $\mathbb{R}$ , though they are usually defined over  $\mathbb{C}$ . We focus on  $\mathbb{R}$  because approximating tensors by rank-one tensors is done primarily on the real numbers in applications. To define them over  $\mathbb{C}$ , one considers the bilinear form  $\langle \cdot, \cdot \rangle_k$  as an extension of a real inner product to a complex product instead of opting for a Hermitian inner product. However, such extensions introduce isotropic vectors:  $v \in \mathbb{C}^{n_k} \setminus \{0\}$  with  $\langle v, v \rangle_k = 0$ . While a general tensor has no isotropic singular vector tuples [9, Proposition 2.6], our tensors are special. In fact, self-orthogonality plays a role in the study of symmetric odecor tensors [19]. We leave the study of two-orthogonal tensors in  $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$  for future work.

We view singular vector tuples as points in the corresponding Segre product [12]. Hence, the normalization  $\langle x^{(k)}, x^{(k)} \rangle_k = 1$  is without loss of generality, although it is necessary for the singular value to be well defined. Let  $\mathcal{T} = \sum_{i=1}^r x_i$  be a two-orthogonal decomposition. Then, each  $x_i$  is a singular vector tuple for  $\mathcal{T}$  and, after normalizing each factor, the corresponding singular value is  $\|x_i\|$ . We recall the following facts.

**PROPOSITION 2.5.** Fix  $\mathcal{T} \in V$  and  $x = x^{(1)} \otimes \cdots \otimes x^{(d)} \in X \setminus \{0\} \subset V$ .

1. The tangent space to  $X$  at  $x$  is

$$T_x X = \sum_{k=1}^d x^{(1)} \otimes \cdots \otimes x^{(k-1)} \otimes \mathbb{R}^{n_k} \otimes x^{(k+1)} \otimes \cdots \otimes x^{(d)}.$$

2. Suppose that  $\langle x^{(k)}, x^{(k)} \rangle_k = 1$  for all  $k \in [d]$ , then  $x$  is a singular vector tuple of  $\mathcal{T}$  with singular value  $\lambda$  if and only if  $T - \lambda x \in N_x X = (T_x X)^\perp$ , the normal space to  $X$  at  $x$ .
3. The singular vector tuples correspond to the critical points of the distance function  $d_{\mathcal{T}}(y) := \langle \mathcal{T} - y, \mathcal{T} - y \rangle$ , on  $y \in X$ . A singular vector tuple  $x$  with nonzero singular value  $\lambda$  corresponds to the critical rank-one approximation  $\lambda x$ , and vice versa.

*Proof.* Part (1) follows by applying the Leibniz rule to the parametrization of  $X$ . For part (2), the equations in Definition 2.3 are equivalent to

$$\langle \mathcal{T}, x^{(1)} \otimes \cdots \otimes x^{(k-1)} \otimes v^{(k)} \otimes x^{(k+1)} \otimes \cdots \otimes x^{(d)} \rangle = \lambda \langle v^{(k)}, x^{(k)} \rangle_k$$

for all  $v^{(k)} \in \mathbb{R}^{n_k}$ . Since  $\langle x^{(k)}, x^{(k)} \rangle_k = 1$  for all  $k \in [d]$ , this can be rewritten as

$$\langle \mathcal{T} - \lambda x, x^{(1)} \otimes \cdots \otimes x^{(k-1)} \otimes v^{(k)} \otimes x^{(k+1)} \otimes \cdots \otimes x^{(d)} \rangle = 0$$

for all  $v^{(k)} \in \mathbb{R}^{n_k}$ , so the statement follows. Part (3) follows from (2); see [25, section 3].  $\square$

The following proposition shows that every tensor can be approximated as a sum of critical rank-one approximations: subtracting off best rank-one approximations decreases the norm of a tensor, and the norm of the residual tends to zero in the limit. This has been studied before [11, 27], but we will provide a proof for completeness.

**PROPOSITION 2.6.** Fix a tensor  $\mathcal{T} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ . For each  $j \in \mathbb{N}$ , let  $x_j$  be a best rank-one approximation of  $\mathcal{T} - \sum_{i < j} x_i$ . Then  $\sum_{i=1}^n x_i \rightarrow \mathcal{T}$  as  $n \rightarrow \infty$ .

*Proof.* We show that  $\|\mathcal{T} - \sum_{i=1}^n x_i\| \rightarrow 0$  as  $n \rightarrow \infty$ . Let  $\|\cdot\|_\sigma$  denote the spectral norm. We have  $\|x_1\| = \|\mathcal{T}\|_\sigma$ , by the definition of best rank-one approximation. Moreover,  $\|\mathcal{T}\|^2 = \|\mathcal{T} - x_1\|^2 + \|x_1\|^2$ , since a critical rank-one approximation  $x_1$  and its residual  $\mathcal{T} - x_1$  are orthogonal. Hence

$$\|\mathcal{T} - x_1\|^2 = \|\mathcal{T}\|^2 - \|\mathcal{T}\|_\sigma^2 \leq (1 - c^2) \|\mathcal{T}\|^2,$$

where the constant  $c > 0$  relates the spectral and Frobenius norms, i.e.,  $c\|\mathcal{S}\| \leq \|\mathcal{S}\|_\sigma$  for all tensors  $\mathcal{S} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ . This constant exists because two norms in a finite-dimensional vector space are equivalent. We have  $0 < c < 1$ , since  $\|\mathcal{S}\|_\sigma < \|\mathcal{S}\|$  if  $\mathcal{S}$  is not a rank-one tensor. Repeating the above for iterated best rank-one approximations yields

$$\|\mathcal{T} - \sum_{i \leq j} x_i\|^2 \leq (1 - c^2) \|\mathcal{T} - \sum_{i < j} x_i\|^2 \leq \cdots \leq (1 - c^2)^j \|\mathcal{T}\|^2.$$

Since  $0 < c < 1$ , we have  $\|\mathcal{T} - \sum_{i \leq n} x_i\| \rightarrow 0$  as  $n \rightarrow \infty$ .  $\square$

The previous result shows that one can use Algorithm 1.1 for *any* input tensor  $\mathcal{T}$  and get arbitrarily good approximations from a numerical perspective. We study when these decompositions are exact, not just arbitrarily good. Moreover, as shown in Example 1.1, these decompositions may be order dependent. We are interested in decompositions that are valid independent of the order. The following result characterizes these decompositions.

**THEOREM 1.2.** *The set of two-orthogonal tensors in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  coincides with the set of tensors with a decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$  such that each  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i \in \mathcal{I}} x_i$  for all  $\mathcal{I} \subseteq \{1, \dots, r\} \setminus \{j\}$ .*

*Proof.* For a rank-one tensor  $x = x^{(1)} \otimes \cdots \otimes x^{(d)} \in X \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  we have that  $x$  is a critical rank-one approximation of a tensor  $\mathcal{T}$  if and only if  $\mathcal{T} - x \in N_x X$ , and  $T_x X = \sum_{k=1}^d x^{(1)} \otimes \cdots \otimes \mathbb{R}^{n_k} \otimes \cdots \otimes x^{(d)}$ ; see Proposition 2.5.

If  $\mathcal{T} = \sum_{i=1}^r x_i$  is two-orthogonal, then for all  $i \neq j$  we have  $x_i \in N_{x_j} X$ . Therefore, any linear combination  $\sum_{i \neq j} \alpha_i x_i$  lies in  $N_{x_j} X$ . Hence,  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i \in \mathcal{I}} x_i$  for all  $\mathcal{I} \subseteq \{1, \dots, r\} \setminus \{j\}$ .

Conversely, consider a tensor  $\mathcal{T} = \sum_{i=1}^r x_i$  such that  $x_j$  is a critical rank-one approximation of  $\mathcal{T} - \sum_{i \in \mathcal{I}} x_i$  for all  $\mathcal{I} \subseteq \{1, \dots, r\} \setminus \{j\}$ . Then, for any  $i \neq j$ ,  $x_j$  is a critical rank-one approximation of  $x_i + x_j$ . Hence,  $x_i \in N_{x_j} X$ . So, for all  $k \in [d]$  and all  $v_k \in \mathbb{R}^{n_k}$  we have

$$\langle x_i^{(1)} \otimes \cdots \otimes x_i^{(d)}, x_j^{(1)} \otimes \cdots \otimes v^{(k)} \otimes \cdots \otimes x_j^{(d)} \rangle = \langle x_i^{(k)}, v^{(k)} \rangle_k \prod_{l \neq k} \langle x_i^{(l)}, x_j^{(l)} \rangle_l = 0.$$

These equations are satisfied if and only if there exist two indices  $k_1 \neq k_2 \in [d]$  such that  $x_i^{(k_1)} \perp x_j^{(k_1)}$  and  $x_i^{(k_2)} \perp x_j^{(k_2)}$ . This holds for all pairs of summands, so  $\mathcal{T}$  is a two-orthogonal tensor.  $\square$

**DEFINITION 2.7.** Define  $W_r$  to be the variety of length-at-most- $r$  two-orthogonal tensors

$$W_r = \overline{\left\{ \sum_{i=1}^r x_i^{(1)} \otimes \cdots \otimes x_i^{(d)} \mid x_i^{(k_1)} \perp x_j^{(k_1)}, x_i^{(k_2)} \perp x_j^{(k_2)} \text{ for at least two indices } k_1 \neq k_2 \right\}}.$$

We allow zero summands here, so the length can be smaller than  $r$ . For  $r = 1$ , no conditions are imposed, so  $W_1 = X$  is the set of rank-one tensors. The overline denotes Zariski closure. We define the two-orthogonal variety to be  $W = \overline{\bigcup_r W_r}$ .

*Remark 2.8.* The two-orthogonal variety  $W \subseteq V = \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  is invariant under the action of the product of special orthogonal groups  $G = \mathrm{SO}(n_1) \times \cdots \times \mathrm{SO}(n_d)$  on  $V$ , acting by change of basis in each factor.

*Remark 2.9.* The equations defining some components of the two-orthogonal variety relate to the odecō equations [2, Theorem 9]. For example, consider a two-orthogonal tensor  $\mathcal{T} = x_1 + x_2 \in W_2 \subset (\mathbb{R}^2)^{\otimes d}$  such that  $x_1$  and  $x_2$  are orthogonal in the  $k_1$ th and  $k_2$ th factors, then  $\mathcal{T} \star_{k_i} \mathcal{T} \in \bigotimes_{l \neq k_i} S^2(\mathbb{R}^2)$  for  $i = 1, 2$ , where  $\star_{k_i}$  denotes contraction along the  $k_i$ th factor and  $S^2(\mathbb{R}^2)$  denotes the space of  $2 \times 2$  symmetric matrices. One can check that these equations define the components of  $W_2$  for  $(\mathbb{R}^2)^{\otimes 3}$  (see section 5) and  $(\mathbb{R}^2)^{\otimes 4}$ . A similar argument may be used for some components of  $W_n \subset (\mathbb{R}^n)^{\otimes d}$ . Finding the implicit equations that describe the two-orthogonal variety is left to future work.

**2.1. Maximal length of two-orthogonal decompositions.** From the definition of  $W_r$ , we have a chain of varieties  $X = W_1 \subseteq W_2 \subseteq \cdots \subseteq W_r \subseteq \cdots$ . The following result shows that this chain stabilizes in finitely many steps. That is, there exists  $N \in \mathbb{N}$  such that  $W_N = W_{N+1} = \cdots$ .

**LEMMA 2.10.** *A two-orthogonal decomposition of a tensor in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  has at most  $N = \min_{k \in [d]} \prod_{j \neq k} n_j$  summands. In particular, the two-orthogonal variety is  $W = W_N$ .*

*Proof.* Fix  $\mathcal{T} = x_1 + \cdots + x_r \in W_r \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ , where  $x_i = x_i^{(1)} \otimes x_i^{(2)} \otimes \cdots \otimes x_i^{(d)} \neq 0$  for all  $i \in [r]$  and these summands are orthogonal in at least two factors. Without loss of generality, suppose that  $n_1 \leq \cdots \leq n_d$ . Consider the tensor  $\tilde{\mathcal{T}} = \tilde{x}_1 + \cdots + \tilde{x}_r \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_{d-1}}$ , where  $\tilde{x}_i = x_i^{(1)} \otimes \cdots \otimes x_i^{(d-1)}$  for all  $i \in [r]$ . The tensors  $\tilde{x}_i$  and  $\tilde{x}_j$  are orthogonal (in at least one factor) for all  $i \neq j$ , since  $x_i$  and  $x_j$  are orthogonal in at least two factors. In particular,  $\tilde{x}^{(1)}, \dots, \tilde{x}^{(r)}$  are linearly independent. Indeed, if we had  $\tilde{x}_i = \sum_{j \neq i} \alpha_j \tilde{x}_j$ , taking the inner product with  $\tilde{x}_i$  leads to  $\|\tilde{x}_i\|^2 = 0$ , a contradiction. In conclusion,

$$r = \dim \left( \text{span}\{\tilde{x}^{(1)}, \dots, \tilde{x}^{(r)}\} \right) \leq \dim (\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_{d-1}}) = n_1 n_2 \cdots n_{d-1}. \quad \square$$

**DEFINITION 2.11** (see [26]). *A Latin square  $L$  is an  $n \times n$  matrix with elements in  $[n]$  such that every row and every column of  $L$  are a permutation of  $[n]$ . Put differently, for every  $i \in [n]$  the maps  $k \mapsto L(i, k)$  and  $k \mapsto L(k, i)$  are permutations of  $n$ . More generally, a Latin hypercube is an array  $L$  indexed by  $[n]^d$  satisfying the following: for every  $j \in [d]$  and every  $(i_1, \dots, i_{j-1}, i_{j+1}, \dots, i_d) \in [n]^{d-1}$  the map  $k \mapsto L(i_1, \dots, i_{j-1}, k, i_{j+1}, \dots, i_d)$  is a permutation of  $[n]$ .*

There exists a Latin hypercube for every  $n$  and  $d$ , e.g.,  $L(i_1, \dots, i_d) = i_1 + \cdots + i_d \bmod n$ , where the sum is understood to be an element in  $[n]$ . A Latin square  $L$  gives two-orthogonal tensors of order three, as follows. Consider the set  $\mathcal{I} = \{(i, j, L(i, j)) \mid i, j \in [n]\} \subset [n]^3$ . We can identify each tuple  $\mathbf{i} = (i_1, i_2, i_3) \in \mathcal{I}$  with the rank-one tensor  $e_{i_1} \otimes e_{i_2} \otimes e_{i_3}$ . Every pair of distinct tuples  $\mathbf{i}, \mathbf{j} \in \mathcal{I}$  differs in at least two indices. Therefore, the tensors  $e_{i_1} \otimes e_{i_2} \otimes e_{i_3}$  and  $e_{j_1} \otimes e_{j_2} \otimes e_{j_3}$  are orthogonal in at least two factors.

*Example 2.12.* Let  $V = (\mathbb{R}^3)^{\otimes 3}$ . A  $3 \times 3$  Latin square corresponds to a two-orthogonal tensor in  $V$  with 9 summands. For example,

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 2 & 3 & 1 \\ \hline 3 & 1 & 2 \\ \hline \end{array} \leftrightarrow \begin{aligned} & e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_3 \otimes e_3 \\ & + e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_2 \otimes e_3 + e_2 \otimes e_3 \otimes e_1 \\ & + e_3 \otimes e_1 \otimes e_3 + e_3 \otimes e_2 \otimes e_1 + e_3 \otimes e_3 \otimes e_2. \end{aligned}$$

Lemma 2.10 implies that the maximum number of summands of a two-orthogonal tensor in  $V$  is nine. Each summand can be scaled arbitrarily and two-orthogonality is preserved.

**LEMMA 2.13.** *There exists a two-orthogonal decomposition in  $V = \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  with  $N = \min_{k \in [d]} \prod_{j \neq k} n_j$  summands.*

*Proof.* We construct a two-orthogonal decomposition of maximal length using Latin hypercubes. First, suppose  $n_1 = \cdots = n_d = n$ . Let  $L$  be a Latin hypercube indexed by  $[n]^{d-1}$ . Consider the set  $\mathcal{I} = \{(i_1, \dots, i_{d-1}, L(i_1, \dots, i_{d-1})) \mid i_k \in [n]\}$ . Then  $|\mathcal{I}| = n^{d-1}$  and every pair of tuples in  $\mathcal{I}$  differ in at least two indices. This gives the family of two-orthogonal tensors in  $(\mathbb{R}^n)^{\otimes d}$ :

$$\sum_{\mathbf{i} \in \mathcal{I}} \lambda_{\mathbf{i}} e_{i_1} \otimes \cdots \otimes e_{i_d}$$

for any  $\lambda_{\mathbf{i}} \in \mathbb{R}$ . Let  $V = \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ , with  $n_1 \leq \cdots \leq n_d$ . Consider a Latin hypercube indexed by  $[n_{d-1}]^{d-1}$  and choose a subarray of format  $n_1 \times \cdots \times n_{d-1}$ . This gives a two-orthogonal decomposition of length  $n_1 \cdots n_{d-1}$  in  $V$ , following the same reasoning.  $\square$

**THEOREM 1.3.** *The maximal length of a two-orthogonal decomposition in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  is  $N = \min_{1 \leq k \leq d} \prod_{j \neq k} n_j$ . In particular, the two-orthogonal variety is  $W = W_N$ .*

*Proof.* Lemma 2.10 implies that a two-orthogonal decomposition cannot have more than  $N = \min_{k \in [d]} \prod_{j \neq k} n_j$  summands. Lemma 2.13 shows how to construct two-orthogonal decompositions with  $N$  summands.  $\square$

Decompositions obtained from Latin hypercubes satisfy the following property.

**DEFINITION 2.14.** *A two-orthogonal decomposition  $\mathcal{T} = \sum_{i=1}^r x_i^{(1)} \otimes \cdots \otimes x_i^{(d)}$  is basis-aligned if for all  $k \in [d]$  and all  $i, j \in [r]$ ,  $x_i^{(k)}$  and  $x_j^{(k)}$  are either collinear or orthogonal.*

Not all two-orthogonal decompositions are basis-aligned. For example, consider the decomposition  $\mathcal{T} = e_1 \otimes e_1 \otimes e_1 + (e_1 + e_2) \otimes e_2 \otimes e_2 \in W_2$ . We explore two-orthogonal decompositions that are not basis-aligned in section 4.

**2.2. Two-orthogonal rank.** Two-orthogonal decompositions are not unique, in general.

*Example 2.15.* Let  $d \geq 2$ , let  $V = (\mathbb{R}^2)^{\otimes d}$ , and let  $\mathcal{I}$  be the set of binary strings of length  $d$  with an even number of ones:  $\mathcal{I} = \{(i_1, \dots, i_d) \in \{0, 1\}^d \mid \sum_{k=1}^d i_k = 0 \bmod 2\}$ . Let  $\{e_0, e_1\}$  be an orthogonal basis of  $\mathbb{R}^2$ . The following tensor admits two different two-orthogonal decompositions:

$$\mathcal{T} = \sum_{(i_1, \dots, i_d) \in \mathcal{I}} e_{i_1} \otimes \cdots \otimes e_{i_d} = \frac{1}{2}(e_0 + e_1)^{\otimes d} + \frac{1}{2}(e_0 - e_1)^{\otimes d}.$$

The first decomposition has maximal length, but the second one shows that  $\mathcal{T} \in W_2$ .

While Lemma 2.13 constructs two-orthogonal decompositions of maximal length, it is unclear whether these tensors may have two-orthogonal decompositions with fewer terms. We do not know whether the stabilization of the chain  $W_1 \subseteq W_2 \subseteq \dots$  occurs exactly in  $N = \min_{k \in [d]} \prod_{j \neq k} n_j$  steps or before. That is, we do not know whether  $W_{N-1} = W_N$ . The previous example motivates the following definition.

**DEFINITION 2.16.** *The two-orthogonal rank of a two-orthogonal tensor  $\mathcal{T} \in \mathbb{R}^{n_1} \otimes \dots \otimes \mathbb{R}^{n_d}$  is the smallest number  $r$  such that  $\mathcal{T}$  admits a two-orthogonal decomposition of length  $r$ .*

The rank of a tensor  $\mathcal{T}$  is the smallest number  $r$  such that  $\mathcal{T}$  can be expressed as the sum of  $r$  rank-one tensors. The two-orthogonal rank and rank coincide for matrices, due to the singular value decomposition. For higher-order tensors, we can only say that the two-orthogonal rank of  $\mathcal{T}$  is at least the rank of  $\mathcal{T}$ . We will see in section 5 that for  $2 \times 2 \times 2$  tensors the maximum two-orthogonal rank is 4 while the maximum rank is 3.

The notions of rank and two-orthogonal rank could potentially coincide up to the maximal rank. For example, in  $(\mathbb{R}^n)^{\otimes 3}$ , the maximal rank is at most  $\frac{1}{2}n(n+2)-1$  [33, Theorem 3.4], while the maximal potential two-orthogonal rank is  $n^2$ . Proposition 2.17 shows that this is the case for small enough ranks. The *border rank* of a tensor  $\mathcal{T}$  is the smallest  $r$  such that  $\mathcal{T} = \lim_{\epsilon \rightarrow 0} \mathcal{T}_\epsilon$  where each  $\mathcal{T}_\epsilon$  has rank  $r$ , and it may be smaller than the rank of  $\mathcal{T}$  (see Example 2.21).

**PROPOSITION 2.17.** *In  $(\mathbb{R}^n)^{\otimes d}$ ,  $W_{r-1} \subsetneq W_r$  if  $r \leq n$ . Moreover, a general tensor in  $W_r \setminus W_{r-1}$  with  $r \leq n$  has border rank  $r$ , so it has rank  $r$ .*

*Proof.* After an orthogonal change of basis, an odec tensor of two-orthogonal rank  $r$  is of the form  $\mathcal{T} = \sum_{i=1}^r \lambda_i e_i^{\otimes d}$  for  $\lambda_i \in \mathbb{R}$ . This tensor has border rank  $r$ , e.g., by looking at the flattenings [22]. Therefore,  $W_{r-1} \neq W_r$ . A property holds for a general tensor if it holds on a dense open set, and odec tensors of rank  $r$  lie in the intersection of all the irreducible components of  $\overline{W_r \setminus W_{r-1}}$ .  $\square$

This result can be improved by considering the following class of two-orthogonal tensors. A decomposition  $\mathcal{T} = \sum_{i=1}^r x_i^{(1)} \otimes \dots \otimes x_i^{(d)} \in \mathbb{R}^{n_1} \otimes \dots \otimes \mathbb{R}^{n_d}$  is called *strong two-orthogonal* ([35, Definition 3.5]) if the orthogonalities between summands always occur in the same partition of the factors: there exists a nonempty set  $\mathcal{J} \subsetneq [d]$  such that for all  $i \neq j$  we have  $x_i^{(k_1)} \perp x_j^{(k_1)}$  and  $x_i^{(k_2)} \perp x_j^{(k_2)}$  for some indices  $k_1 \in \mathcal{J}$  and  $k_2 \in [d] \setminus \mathcal{J}$ . Put differently, the decomposition is still two-orthogonal when viewed as a matrix decomposition in  $(\bigotimes_{j \in \mathcal{J}} \mathbb{R}^{n_j}) \otimes (\bigotimes_{j \notin \mathcal{J}} \mathbb{R}^{n_j})$ .

**PROPOSITION 2.18.** *Let  $V = \mathbb{R}^{n_1} \otimes \dots \otimes \mathbb{R}^{n_d}$  and let*

$$m = \max_{\mathcal{J} \subsetneq [d]} \min \left\{ \prod_{j \in \mathcal{J}} n_j, \prod_{j \notin \mathcal{J}} n_j \right\}.$$

*Then, for all  $r \leq m$ , a sufficiently general tensor of  $W_r$  has border rank  $r$ . In particular, we have that  $W_{m-1} \subsetneq W_m$ .*

*Proof.* By sufficiently general we mean that it holds for some components of  $W_r \setminus W_{r-1}$ . Consider a nonempty  $\mathcal{J} \subsetneq [d]$ . Picking an orthogonal basis for  $\bigotimes_{j \in \mathcal{J}} \mathbb{R}^{n_j}$  and  $\bigotimes_{j \notin \mathcal{J}} \mathbb{R}^{n_j}$  we can construct a strong two-orthogonal decomposition of length  $\min\{\prod_{j \in \mathcal{J}} n_j, \prod_{j \notin \mathcal{J}} n_j\}$ . By flattening the tensor into a matrix, the singular value decomposition implies that strong two-orthogonal decompositions of length  $r$  lead to tensors of border rank  $r$ ; see [35, Theorem 3.6].  $\square$

*Example 2.19.* When  $V = (\mathbb{R}^n)^{\otimes d}$ , Proposition 2.18 implies that  $W_{r-1} \subsetneq W_r$  if  $r \leq n^{\lfloor \frac{d}{2} \rfloor}$ .

**CONJECTURE 2.20.** Let  $r_g$  be the generic rank in  $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ . For every  $r \leq r_g$ , a sufficiently general tensor  $\mathcal{T} \in W_r \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  has rank  $r$ .

The following examples provide evidence for this conjecture. One can compute the rank of the tensors by applying the technique used in [36], which is illustrated in Example 2.24.

*Example 2.21.* The two-orthogonal tensor

$$\mathcal{T} = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 \in W_3 \subset (\mathbb{R}^2)^{\otimes 3}$$

has rank 3, which is the maximal rank in  $(\mathbb{R}^2)^{\otimes 3}$ . This example can also be embedded in  $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$  and  $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3$ , where the generic rank is 3 [22]. However, note that  $\mathcal{T} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} ((e_1 + \epsilon e_2)^{\otimes 3} - e_1^{\otimes 3})$ , so its border rank is two.

*Example 2.22.* The two-orthogonal tensor

$$\begin{aligned} \mathcal{T} = & e_1 \otimes e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 \otimes e_1 \\ & + e_2 \otimes e_2 \otimes e_2 \otimes e_2 \in (\mathbb{R}^2)^{\otimes 4} \end{aligned}$$

has rank 4, which is the generic rank in  $(\mathbb{C}^2)^{\otimes 4}$  [4]. Looking at the flattenings, we see that the border rank of  $\mathcal{T}$  is also 4. Actually, this decomposition is strong two-orthogonal.

*Example 2.23.* The two-orthogonal tensor

$$\mathcal{T} = e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_3 \otimes e_3 + e_2 \otimes e_1 \otimes e_2 + e_3 \otimes e_1 \otimes e_3 \in (\mathbb{R}^3)^{\otimes 3}$$

has rank 5, which is the generic rank in  $(\mathbb{C}^3)^{\otimes 3}$  [32]. The maximal real rank of  $3 \times 3 \times 3$  tensors is also 5 [3].

*Example 2.24.* The two-orthogonal tensor

$$\begin{aligned} \mathcal{T} = & e_1 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_1 \otimes e_3 \otimes e_3 + e_1 \otimes e_4 \otimes e_4 \\ & + e_2 \otimes e_1 \otimes e_2 + e_3 \otimes e_1 \otimes e_3 + e_4 \otimes e_1 \otimes e_4 \in (\mathbb{R}^4)^{\otimes 4} \end{aligned}$$

has rank 7, which is the generic rank in  $(\mathbb{C}^4)^{\otimes 3}$  [24], as follows.

Let  $\mathcal{L} = \text{span}\{\mathcal{T}(e_1, \cdot, \cdot), \mathcal{T}(e_2, \cdot, \cdot), \mathcal{T}(e_3, \cdot, \cdot), \mathcal{T}(e_4, \cdot, \cdot)\}$  be the linear space spanned by the slices of  $\mathcal{T}$  obtained by fixing the first factor. The rank of  $\mathcal{T}$  is the minimal number of rank-one matrices whose linear span contains  $\mathcal{L}$  (e.g., see [36, Proposition 3.3]). The two-orthogonal decomposition above has seven summands, so  $\text{rank}(\mathcal{T}) \leq 7$ . Suppose that  $\text{rank}(\mathcal{T}) \leq 6$ , meaning that there exist six rank-one matrices spanning a space  $\mathcal{K}$  that contains  $\mathcal{L}$ . Then,  $\mathcal{K}$  is spanned by  $\mathcal{L}$  along with two rank-one matrices, so every element of  $\mathcal{K}$  is of the form

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & a_1 \end{pmatrix} + a_5 \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{pmatrix} \otimes \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \end{pmatrix} + a_6 \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{pmatrix} \otimes \begin{pmatrix} y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \end{pmatrix}$$

for fixed  $\{x_{ij}\}, \{y_{ij}\}$  and variable coefficients  $\{a_i\}$ . Computing the  $2 \times 2$  minors of these matrices we get that all the rank-one matrices in  $\mathcal{K}$  have  $a_1 = 0$ , and such matrices do not span  $\mathcal{L}$ . Hence,  $\text{rank}(\mathcal{T}) = 7$ .

Computing the tensor rank is NP-hard [15]. This means that, although a generic tensor has generic rank, it is computationally infeasible to verify whether or not the rank of an arbitrary tensor is smaller, equal, or bigger than the generic rank. We hope that this may be more feasible for two-orthogonal tensors. Perhaps two-orthogonal tensors may be good candidates for tensors with provably high rank, even above the generic rank. Note that the expected generic rank of a tensor of  $(\mathbb{C}^n)^{\otimes d}$  is  $\lceil \frac{n^d}{d(n-1)+1} \rceil$ , which is smaller than the maximal two-orthogonal rank in  $(\mathbb{R}^n)^{\otimes d}$ .

**2.3. Optimal truncations.** When the two-orthogonal rank and the usual rank of a tensor coincide, the two-orthogonal decomposition has good truncation properties. The  $k$ th secant variety of  $X$ , denoted  $\sigma_k(X)$ , is the closure of the set of rank-at-most- $k$  tensors:

$$\sigma_k(X) = \overline{\bigcup_{x_1, \dots, x_k \in X} \text{span}\{x_1, \dots, x_k\}}.$$

Consider a rank- $r$  tensor  $\mathcal{T}$  and let  $k < r$ . A *critical rank- $k$  approximation* of  $\mathcal{T}$  is a tensor  $\mathcal{S} = x_1 + \dots + x_k \in \sigma_k(X)$  smooth such that  $\mathcal{T} - \mathcal{S} \perp T_{\mathcal{S}}\sigma_k(X)$ .

**PROPOSITION 2.25.** *Consider a two-orthogonal decomposition  $\mathcal{T} = \sum_{i=1}^r x_i \in W_r$ . Suppose that  $\text{rank}(\mathcal{T}) = r$  and that the truncation  $\mathcal{S} = \sum_{i=1}^k x_i$ , where  $k < r$ , is sufficiently general as a point in  $\sigma_k(X)$ . Then,  $\mathcal{S}$  is a critical rank- $k$  approximation of  $\mathcal{T}$ .*

*Proof.* Suppose that  $\mathcal{S}$  is sufficiently general in the sense that it satisfies Terracini's lemma, i.e., we have  $T_{\mathcal{S}}\sigma_k(X) = \text{span}\{T_{x_1}X, \dots, T_{x_k}X\}$ . Recall that the tangent space to  $X$  at  $x$  is  $T_x X = \sum_{l=1}^d x^{(1)} \otimes \dots \otimes x^{(l-1)} \otimes \mathbb{R}^{n_l} \otimes x^{(l+1)} \otimes \dots \otimes x^{(d)}$ . Two-orthogonality implies  $x_j \perp T_{x_i}X$  for all  $j > k$  and  $i \leq k$ . Hence,  $\mathcal{S}$  is a critical rank- $k$  approximation of  $\mathcal{T}$ .  $\square$

The previous result suggests that if we order the summands such that  $\|x_1\| \geq \dots \geq \|x_r\|$ , then the truncation  $\mathcal{S} = \sum_{i=1}^k x_i$  is a good candidate for a best rank- $k$  approximation of  $\mathcal{T} = \sum_{i=1}^r x_i$ . Following [35], a *Schmidt–Eckart–Young (SEY) decomposition* is a tensor decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$  such that for all  $k < r$ , retaining the first  $k$  summands gives an optimal solution in the sense that  $\sum_{i=1}^k x_i$  is a minimizer of  $\|\mathcal{T} - \mathcal{S}\|$  over all  $\mathcal{S}$  of rank at most  $k$ .

Two-orthogonality is necessary for having an SEY decomposition [35, Theorem 3.3], but it is insufficient because a tensor  $\mathcal{T}$  cannot have an SEY decomposition if its border rank and rank disagree (Example 2.21). Taking flattenings, the singular value decomposition shows that strong two-orthogonality is sufficient to be an SEY decomposition [35, Theorem 3.6]. The authors in [35] state the following: “It remains an open question whether strong two-orthogonality is also necessary.” The following counterexample shows that it is not necessary.

*Example 2.26.* Consider the two-orthogonal decomposition

$$(2.1) \quad \mathcal{T} = 3e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes \left( \frac{1}{2}e_1 + 2e_2 \right) + e_1 \otimes e_2 \otimes e_3 \in \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^3.$$

We claim that this is an SEY decomposition, despite not being strong two-orthogonal (there is no factor for which the three summands are orthogonal). The border rank of  $\mathcal{T}$  is three, e.g., by looking at its flattenings. Critical rank-one approximations correspond to singular vector tuples, by Proposition 2.5. If  $x$  is a singular vector tuple of  $\mathcal{T}$  with singular value  $\lambda$ , then  $\|\mathcal{T} - \lambda x\|^2 = \|\mathcal{T}\|^2 - \lambda^2$ . The tensor  $\mathcal{T}$  has

eight singular vector tuples, which is the generic number for a tensor of this format [12]. The singular vector tuple  $e_1 \otimes e_1 \otimes e_1$  has singular value 3, and one can check that all the other singular vector tuples have singular values smaller than 3. Hence,  $3e_1 \otimes e_1 \otimes e_1$  is the best rank-one approximation of  $\mathcal{T}$ .

The tensor  $\mathcal{T}$  has three critical rank-two approximations (which can be computed, for example, using the notion of critical ideal defined in [8]). One of them comes from the first two summands in our decomposition:  $\mathcal{S} = 3e_1 \otimes e_1 \otimes e_1 + e_2 \otimes e_2 \otimes (1/2e_1 + 2e_2)$ , which leads to a residual  $\|\mathcal{T} - \mathcal{S}\|^2 = 1$ . The other two are the tensors

$$\left( \begin{array}{cc|cc|cc} \frac{120-63\alpha}{38} & 0 & \frac{12-12\alpha}{19} & 0 & 0 & 1 \\ 0 & \frac{72-53\alpha}{76} & 0 & \alpha & 0 & 0 \end{array} \right) \in \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^3,$$

where  $505\alpha^2 - 1010\alpha + 144 = 0$ . The rank of these tensors is greater than two, but their border rank is two. One can check that both solutions lead to a residual greater than one. Hence,  $\mathcal{S}$  is the best rank-two approximation of  $\mathcal{T}$ , so (2.1) is an SEY decomposition.

**2.4. Two-orthogonality for partially symmetric tensors.** We study two-orthogonal decompositions for partially symmetric tensors. A tensor  $\mathcal{T} \in (\mathbb{R}^n)^{\otimes d}$  is symmetric if  $\mathcal{T}_{i_1, \dots, i_d} = \mathcal{T}_{\pi(i_1), \dots, \pi(i_d)}$  for any permutation  $\pi \in S_d$ . For  $d=2$ , this says  $M = M^\top$ . Let  $S^d(\mathbb{R}^n)$  denote the space of symmetric tensors of format  $n \times \dots \times n$  ( $d$  times), which is naturally identified with the space of homogeneous polynomials in  $n$  variables of degree  $d$  [5]. Using this identification, we express the rank-one tensors in  $S^d(\mathbb{R}^n)$  as  $\ell^d$  for  $\ell \in (\mathbb{R}^n)^*$ .

Let  $V = S^{d_1}\mathbb{R}^{n_1} \otimes \dots \otimes S^{d_p}\mathbb{R}^{n_p}$  and let  $X = \{\ell_1^{d_1} \otimes \dots \otimes \ell_p^{d_p} \mid \ell_i \in (\mathbb{R}^{n_i})^*\} \subset V$  be the cone over the Segre–Veronese variety, the set of rank-one tensors in  $V$ . An inner product  $\langle \cdot, \cdot \rangle$  in  $(\mathbb{R}^n)^*$  gives a unique inner product on  $S^d\mathbb{R}^n$  defined on two rank-one tensors as  $\langle f^d, g^d \rangle = \langle f, g \rangle^d$  and extended to all  $S^d\mathbb{R}^n$  by linearity. This implies that  $\langle f_1 \cdots f_d, g_1 \cdots g_d \rangle = \frac{1}{d!} \sum_{\pi \in S_d} \prod_{k=1}^d \langle f_k, g_{\pi(k)} \rangle$ . As before, a set of inner products in  $(\mathbb{R}^{n_1})^*, \dots, (\mathbb{R}^{n_p})^*$  defines an inner product in  $V$ .

**DEFINITION 2.27.** A partially symmetric tensor  $\mathcal{T} \in S^{d_1}\mathbb{R}^{n_1} \otimes \dots \otimes S^{d_p}\mathbb{R}^{n_p}$  is called two-orthogonal if it admits a decomposition  $\mathcal{T} = \sum_{i=1}^r x_i$ , where  $x_i = \ell_{1,i}^{d_1} \otimes \dots \otimes \ell_{p,i}^{d_p} \neq 0$  and all pairs of summands  $x_i$  and  $x_j$  with  $i \neq j$  are orthogonal in at least two factors, counting multiplicities  $d_k$ . We define  $W_r$  as the closure of the set of length-at-most- $r$  two-orthogonal partially symmetric tensors, and  $W = \overline{\bigcup_r W_r}$  as the two-orthogonal variety.

**Example 2.28.** The tensor  $f_1^2 \otimes g_1 + f_2^2 \otimes g_2 \in S^2\mathbb{R}^m \otimes \mathbb{R}^n$  with  $f_1 \perp f_2$  is two-orthogonal.

**PROPOSITION 2.29.** Let  $x = \ell_1^{d_1} \otimes \dots \otimes \ell_p^{d_p} \in X \setminus \{0\} \subseteq V$ .

1. The tangent space to  $X$  at  $x$  is

$$T_x X = \left\{ \sum_{k=1}^d \ell_1^{d_1} \otimes \dots \otimes \ell_k^{d_k-1} \ell \otimes \dots \otimes \ell_p^{d_p} \mid \ell \in (\mathbb{R}^{n_k})^* \right\}.$$

2. If  $d_k \geq 2$  for all  $k$ , then

$$N_x X \cap X = \left\{ y_1^{d_1} \otimes \dots \otimes y_p^{d_p} \mid \langle \ell_k, y_k \rangle = 0, \text{ for some } k \right\}.$$

*Proof.* Part (1) follows by applying the Leibniz rule to the parametrization of  $X$ . For part (2), suppose that  $y = y_1^{d_1} \otimes \cdots \otimes y_p^{d_p} \in N_x X$ , then

$$\langle y, \ell_1^{d_1} \otimes \cdots \otimes \ell_k^{d_k-1} \ell \otimes \cdots \otimes \ell_p^{d_p} \rangle = \langle y_k, \ell_k \rangle^{d-1} \langle y_k, \ell \rangle \prod_{j \neq k} \langle y_j, \ell_j \rangle^{d_j} = 0$$

for all  $k \in [p]$  and all  $\ell \in (\mathbb{R}^{n_k})^*$ . So there exists  $k \in [p]$  such that  $\langle y_k, \ell_k \rangle = 0$ .  $\square$

Proposition 2.29 implies that Theorem 1.2 also holds for partially symmetric tensors. Theorem 2.30 is an extension of Theorem 1.3 to partially symmetric tensors. The proof is similar; we include it for completeness.

**THEOREM 2.30.** *Let  $V = S^{d_1} \mathbb{R}^{m_1} \otimes \cdots \otimes S^{d_p} \mathbb{R}^{m_p} \otimes \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_q}$ ,  $n_1 \leq \cdots \leq n_q$ , and  $d_k \geq 2$  for all  $k \in [p]$ . Then, the maximal length of a two-orthogonal decomposition is  $N = m_1 \cdots m_p n_1 \cdots n_{q-1}$ . In particular,  $W = W_N$ .*

*Proof.* Let  $\mathcal{T} = \sum_{i=1}^r \ell_{1,i}^{d_1} \otimes \cdots \otimes \ell_{p,i}^{d_p} \otimes u_{1,i} \otimes \cdots \otimes u_{q,i} \in V$  be two-orthogonal. Consider the tensor  $\mathcal{T}' = \sum_{i=1}^r \ell_{1,i} \otimes \cdots \otimes \ell_{p,i} \otimes u_{1,i} \otimes \cdots \otimes u_{q-1,i}$ . Each pair of summands of  $\mathcal{T}'$  are orthogonal in at least one factor, so they are linearly independent. Therefore,  $r \leq m_1 \cdots m_p n_1 \cdots n_{q-1}$ . It remains to show that the bound can be achieved: let  $\{e_{k,1}, \dots, e_{k,m_k}\}$  be an orthogonal basis of  $(\mathbb{R}^{m_k})^*$ , let  $\{e'_{k,1}, \dots, e'_{k,n_k}\}$  be an orthogonal basis of  $\mathbb{R}^{n_k}$ , and let  $\mathcal{J} \subseteq [n_1] \times \cdots \times [n_q]$  be a maximal set of indices given by a Latin hypercube, as in Lemma 2.13. Then, the tensor

$$\mathcal{T} = \sum_{\mathbf{j} \in \mathcal{J}, i_1, \dots, i_p} e_{1,i_1}^{d_1} \otimes \cdots \otimes e_{p,i_p}^{d_p} \otimes e'_{1,j_1} \otimes \cdots \otimes e'_{d,j_d}$$

is two-orthogonal and consists of  $m_1 \cdots m_p n_1 \cdots n_{q-1}$  terms.  $\square$

**3. Dimension of the two-orthogonal variety.** In this section, we prove Theorem 1.4, which lower bounds the dimension of the two-orthogonal variety. To do this, we construct a dimension-preserving map that parametrizes a set of basis-aligned two-orthogonal tensors. Recall that the *Hamming distance* on  $[n]^d$  is  $d_H(\mathbf{i}, \mathbf{j}) = |\{k | i_k \neq j_k\}|$ .

**LEMMA 3.1.** *Let  $n \geq 2$ ,  $d \geq 3$ , and let  $L$  be the Latin hypercube on  $[n]^{d-1}$  cells given by  $L(i_1, \dots, i_{d-1}) = i_{d-1} + \sum_{k=1}^{d-2} (-1)^{k+1} i_k \pmod{n}$ . Let  $\mathcal{I} = \{(i_1, \dots, i_{d-1}, L(i_1, \dots, i_{d-1}))\}$ . Then there exists a subset  $\mathcal{J} \subset \mathcal{I}$  with  $|\mathcal{J}| = n - 1$  and  $d_H(\mathbf{i}, \mathbf{j}) \geq 3$  for all distinct  $\mathbf{i}, \mathbf{j} \in \mathcal{J}$ .*

*Proof.* If  $d$  is even, then  $L(i, \dots, i) = i$ , so we can let  $\mathcal{J} = \{(i, i, \dots, i) | i \in [n-1]\}$ . If  $d$  is odd, then  $L(i, \dots, i) = 2i \pmod{n}$ , so we consider two cases. If  $n$  is odd, let  $\mathcal{J} = \{(i, \dots, i, 2i \pmod{n}) | i \in [n-1]\}$ . If  $n$  is even, let

$$\mathcal{J} = \left\{ (i, \dots, i, i, 2i) \mid 1 \leq i \leq \frac{n}{2} \right\} \cup \left\{ (i, \dots, i, i+1, 2i+1-n) \mid \frac{n}{2} < i \leq n-1 \right\}. \quad \square$$

**THEOREM 1.4.** *The dimension of the two-orthogonal variety in  $(\mathbb{R}^n)^{\otimes d}$  is at least*

$$n^{d-1} + d \binom{n}{2}.$$

*Proof.* We parametrize a set of two-orthogonal tensors via a dimension-preserving map. This lower bounds the dimension. Let  $L$  be the Latin hypercube in Lemma 3.1, and let  $\mathcal{I} = \{(i_1, \dots, i_{d-1}, L(i_1, \dots, i_{d-1})) | i_j \in [n]\}$ . Given a set of scalars  $\{\lambda_{\mathbf{i}} | \mathbf{i} \in \mathcal{I}\} \subset \mathbb{R}$  and an orthonormal basis  $\{u_1^{(k)}, \dots, u_n^{(k)}\} \subset \mathbb{R}^n$  for each  $k \in [d]$ , we construct

$$\mathcal{T} = \sum_{\mathbf{i} \in \mathcal{I}} \lambda_{\mathbf{i}} u_{i_1}^{(1)} \otimes \cdots \otimes u_{i_d}^{(d)} \in W_{n^{d-1}}.$$

There are  $n^{d-1}$  degrees of freedom to choose the  $\lambda_i$ 's and, for each factor  $k \in [d]$ , we have  $\binom{n}{2}$  degrees of freedom to choose  $\{u_1^{(k)}, \dots, u_n^{(k)}\}$ . Hence, the expected dimension of the variety of tensors with such a decomposition is  $n^{d-1} + d\binom{n}{2}$ .

We can rescale each  $u_i^{(k)}$  and scale the  $\lambda_i$ 's accordingly. Hence, in the following parametrization we fix an entry of the vectors to be one, instead of imposing that they have norm one. Let  $\theta = (\theta_{ij} \mid 1 \leq j < i \leq n) \in \mathbb{R}^{\binom{n}{2}}$  and consider the matrix

$$U(\theta) = \begin{pmatrix} & & \\ \vdots & & \vdots \\ u_1(\theta) & \cdots & u_n(\theta) \\ & & \vdots \end{pmatrix} = \begin{pmatrix} 1 & p_{1,2}(\theta) & p_{1,3}(\theta) & \cdots & p_{1,n}(\theta) \\ \theta_{2,1} & 1 & p_{2,3}(\theta) & \cdots & p_{2,n}(\theta) \\ \theta_{3,1} & \theta_{3,2} & 1 & \cdots & p_{3,n}(\theta) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{n,1} & \theta_{n,2} & \cdots & \cdots & 1 \end{pmatrix},$$

where  $p_{ij}(\theta) \in \mathbb{R}(\theta)$  are chosen such that the columns of  $U$  are orthogonal (for simplicity in notation, we consider the Euclidean inner product throughout this proof). We parametrize a family of basis-aligned tensors in  $W_{n^{d-1}}$  by

$$\begin{aligned} \phi: \quad \mathbb{R}^{n^{d-1}} \times \left(\mathbb{R}^{\binom{n}{2}}\right)^d &\rightarrow (\mathbb{R}^n)^{\otimes d}, \\ (\lambda_i) \quad , \quad \theta^{(1)}, \dots, \theta^{(d)} &\mapsto \sum_{i \in \mathcal{I}} \lambda_i u_{i1}(\theta^{(1)}) \otimes \cdots \otimes u_{id}(\theta^{(d)}). \end{aligned}$$

We show that the Jacobian of  $\phi$  is generically full-rank. In other words, given a generic  $\mathcal{S} \in \text{Im}(\phi) \subseteq W_{n^{d-1}}$ , we show that  $J\phi: \mathbb{R}^{n^{d-1}} \times \left(\mathbb{R}^{\binom{n}{2}}\right)^d \rightarrow T_{\mathcal{S}}W_{n^{d-1}}$  is injective. We study  $J\phi|_{\theta^{(1)} = \dots = \theta^{(d)} = 0}$ . Orthogonality implies that  $p_{ij}(0) = 0$  for all  $i, j$ . Let  $i < j$ , then

$$0 = \langle u_i(\theta), u_j(\theta) \rangle = \sum_{k < i} p_{ki}(\theta)p_{kj}(\theta) + p_{ij}(\theta) + \sum_{i < k < j} \theta_{ki}p_{kj}(\theta) + \theta_{ji} + \sum_{k > j} \theta_{ki}\theta_{kj}.$$

Taking derivatives and evaluating at  $\theta = 0$ , we get

$$\frac{\partial p_{ij}(\theta)}{\partial \theta_{kl}} \Big|_{\theta=0} = \begin{cases} -1 & \text{if } (i, j) = (l, k), \\ 0 & \text{otherwise.} \end{cases}$$

In what follows, we write  $\Theta = 0$  to denote  $\theta^{(1)} = \dots = \theta^{(d)} = 0$ . Note that  $d_H(\mathbf{i}, \mathbf{j}) \geq 2$  for all distinct  $\mathbf{i}, \mathbf{j} \in \mathcal{I}$ . Hence

$$\begin{aligned} \frac{\partial \phi_{\mathbf{i}}}{\partial \lambda_{\mathbf{j}}} \Big|_{\Theta=0} &= \begin{cases} \delta_{ij} & \text{if } \mathbf{j} \in \mathcal{I}, \\ 0 & \text{otherwise,} \end{cases} \\ \frac{\partial \phi_{\mathbf{i}}}{\partial \theta_{\ell_1 \ell_2}^{(k)}} \Big|_{\Theta=0} &= \begin{cases} \pm \lambda_j & \text{if } \mathbf{j} \in \mathcal{I}, d_H(\mathbf{i}, \mathbf{j}) = 1, i_k \neq j_k, \text{ and } \{\ell_1, \ell_2\} = \{i_k, j_k\}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The sign in the last expression depends on  $L$  and whether  $i_k < j_k$ . Note that  $\frac{\partial \phi_{\mathbf{i}}}{\partial \theta_{\ell_1 \ell_2}^{(k)}}|_{\Theta=0} = 0$  if  $\mathbf{i} \in \mathcal{I}$  and  $\frac{\partial \phi_{\mathbf{i}}}{\partial \lambda_{\mathbf{j}}}|_{\Theta=0} = 0$  if  $\mathbf{i} \notin \mathcal{I}$ . Therefore, after reordering rows,  $J\phi|_{\Theta=0}$  is a block matrix. The block of partial derivatives with respect to  $\{\lambda_i\}$  is a full-rank  $n^{d-1} \times n^{d-1}$  matrix (a permutation matrix). The other block, corresponding to partial derivatives with respect to  $\{\theta_{\ell_1 \ell_2}^{(k)}\}$ , is an  $n^{d-1}(n-1) \times d\binom{n}{2}$  matrix. Entries outside these two blocks are zero (see Example 3.2). We show that  $J\phi|_{\Theta=0}$  is full-rank

for generic  $\lambda_i$ 's. Let  $\mathcal{J}$  be as in Lemma 3.1 and consider the following submatrix of  $J\phi|_{\Theta=0}$  where we set  $\lambda_i$  to zero if  $i \notin \mathcal{J}$ :

$$M = \left( \begin{array}{c|c} \frac{\partial \phi_i}{\partial \theta_{\ell_1 \ell_2}^{(k)}} & \\ \hline \Theta=0, \lambda_i=0 & \text{if } i \notin \mathcal{J} \end{array} \right)_{i,(k,\ell_1,\ell_2)} \in \mathbb{R}^{n^d \times d \binom{n}{2}}.$$

Since  $|\mathcal{J}| = (n-1)$  for every  $k \in [d]$  and every  $\ell_2 < \ell_1 \in [n]$ , there is some  $\mathbf{j} \in \mathcal{J}$  such that  $j_k = \ell_1$  or  $j_k = \ell_2$ , so every column of  $M$  is nonzero. Moreover, we have  $d_H(\mathbf{j}_1, \mathbf{j}_2) \geq 3$  for all distinct  $\mathbf{j}_1, \mathbf{j}_2 \in \mathcal{J}$ , so there is no  $\mathbf{i} \in [n]^d$  with  $d_H(\mathbf{i}, \mathbf{j}_1) = d_H(\mathbf{i}, \mathbf{j}_2) = 1$ . Hence, every row of  $M$  has at most one nonzero entry. Therefore,  $M$  is full-rank as long as  $\lambda_i \neq 0$  for  $i \in \mathcal{J}$ . By continuity of the determinant,  $J\phi|_{\Theta=0}$  is full-rank for generic  $\lambda_i$ 's.  $\square$

*Example 3.2.* The Jacobian of  $\phi$  in the proof of Theorem 1.4 is given below for  $(\mathbb{R}^2)^{\otimes 3}$ :

$$J\phi|_{\Theta=0} = \begin{array}{ccccccccc} & \lambda_{112} & \lambda_{121} & \lambda_{211} & \lambda_{222} & \theta_{21}^{(1)} & \theta_{21}^{(2)} & \theta_{21}^{(3)} \\ \begin{array}{c} 112 \\ 121 \\ 211 \\ 222 \\ 111 \\ 122 \\ 212 \\ 221 \end{array} & \left( \begin{array}{cccccc} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & -\lambda_{211} & -\lambda_{121} & -\lambda_{112} \\ & & & & & -\lambda_{222} & \lambda_{112} & \lambda_{121} \\ & & & & & \lambda_{112} & -\lambda_{222} & \lambda_{211} \\ & & & & & \lambda_{121} & \lambda_{211} & -\lambda_{222} \end{array} \right) \end{array}.$$

This matrix is full-rank for generic  $\lambda_i \in \mathbb{R}$ . Hence, the two-orthogonal decompositions

$$\lambda_{112} u_1^{(1)} \otimes u_1^{(2)} \otimes u_2^{(3)} + \lambda_{121} u_1^{(1)} \otimes u_2^{(2)} \otimes u_1^{(3)} + \lambda_{211} u_2^{(1)} \otimes u_1^{(2)} \otimes u_1^{(3)} + \lambda_{222} u_2^{(1)} \otimes u_2^{(2)} \otimes u_2^{(3)},$$

where  $u_1^{(k)} = (1, \theta_{21}^{(k)}) \perp (-\theta_{21}^{(k)}, 1) = u_2^{(k)}$ , parametrize a variety of dimension  $4 + 3 \binom{2}{2} = 7$ .

We have seen how some components defined by maximal basis-aligned tensors are nondefective, i.e., the dimension is the expected one. However, these components could be part of a higher-dimensional component. Indeed, there exist two-orthogonal decompositions of maximal length that are not basis-aligned, which implies having more degrees of freedom in the parametrization. The following example shows that the lower bound of Theorem 1.4 is not tight. Calculating the dimension of the two-orthogonal variety remains an open question.

*Example 3.3.* Consider the following family of two-orthogonal tensors in  $(\mathbb{R}^4)^{\otimes 3}$ ,

$$\begin{aligned} & \lambda_{111} e_1 \otimes e_1 \otimes e_1 + \lambda_{122} e_1 \otimes e_2 \otimes e_2 + \lambda_{133} e'_1 \otimes e_3 \otimes e_3 + \lambda_{144} e'_1 \otimes e_4 \otimes e_4 \\ & + \lambda_{212} e_2 \otimes e_1 \otimes e_2 + \lambda_{221} e_2 \otimes e_2 \otimes e_1 + \lambda_{234} e'_2 \otimes e_3 \otimes e_4 + \lambda_{243} e'_2 \otimes e_4 \otimes e_3 \\ & + \lambda_{313} e_3 \otimes e'_1 \otimes e'_3 + \lambda_{324} e_3 \otimes e'_2 \otimes e'_4 + \lambda_{331} e'_3 \otimes e'_3 \otimes e'_1 + \lambda_{342} e'_3 \otimes e'_4 \otimes e'_2 \\ & + \lambda_{414} e_4 \otimes e'_1 \otimes e'_4 + \lambda_{423} e_4 \otimes e'_2 \otimes e'_3 + \lambda_{432} e'_4 \otimes e'_3 \otimes e'_2 + \lambda_{441} e'_4 \otimes e'_4 \otimes e'_1, \end{aligned}$$

where  $e'_1, e'_2 \in \text{span}\{e_1, e_2\}$ ,  $e'_3, e'_4 \in \text{span}\{e_3, e_4\}$ ,  $\langle e'_1, e'_2 \rangle = \langle e'_3, e'_4 \rangle = 0$ , and  $\|e'_i\| = 1$  for all  $i$ . We have  $\binom{4}{2} = 6$  degrees of freedom to choose  $\{e_1, \dots, e_4\}$ , one degree of

freedom to choose  $\{e'_1, e'_2\}$ , and one degree of freedom to choose  $\{e'_3, e'_4\}$ . This adds up to eight degrees of freedom for each factor, so a total of 24. We can arbitrarily scale each summand, which adds 16 degrees of freedom. Using the computer algebra software *Macaulay2* [13], we computed the Jacobian of the corresponding parametrization and checked that it is generically full-rank. Hence, the dimension of the variety parametrized by these decompositions is 40, while the lower bound given by Theorem 1.4 is 34.

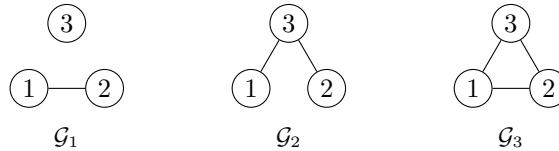
**4. Parametric description of the two-orthogonal variety.** We give a combinatorial parametrization of the two-orthogonal variety using a graph-theoretical perspective. For binary tensors, the resulting graphs are bipartite; we study this case in more depth.

**4.1. Graphical descriptions.** A *graph* is a pair  $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ , where  $V(\mathcal{G})$  is a set of vertices and  $E(\mathcal{G})$  is a set of unordered pairs of vertices, called edges. We describe two-orthogonal decompositions via graphs, as follows.

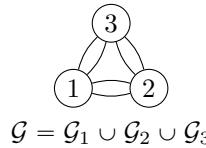
**DEFINITION 4.1.** Let  $V = \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ . The graphical description of a two-orthogonal decomposition  $\mathcal{T} = \sum_{i=1}^r x_i^{(1)} \otimes \cdots \otimes x_i^{(d)} \in V$  is a tuple of  $d$  graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  such that  $V(\mathcal{G}_k) = [r]$  and  $\{i, j\} \in E(\mathcal{G}_k)$  if and only if  $x_i^{(k)} \perp x_j^{(k)}$ .

*Example 4.2.* The graphical description of an odeco decomposition of rank  $r$  is  $(K_r, \dots, K_r)$ , where  $K_r$  is the complete graph on  $r$  vertices.

*Example 4.3.* Consider the tensor  $e_1 \otimes e_1 \otimes e_1 + e_2 \otimes (e_1 + e_2) \otimes e_2 + (e_1 - e_2) \otimes e_3 \otimes e_3$ . Of the three vectors  $x_i^{(1)}$  appearing in the first factor, the first and second vectors are orthogonal. In the second factor, the first and third vectors are orthogonal and the second and third vectors are orthogonal. In the third factor, all three vectors are orthogonal. Hence the graphical description of the decomposition is the following:

 $\mathcal{G}_1$  $\mathcal{G}_2$  $\mathcal{G}_3$ 

A *multigraph* is  $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ , where  $V(\mathcal{G})$  is the set of vertices and  $E(\mathcal{G})$  is a multiset of unordered pairs of vertices. Given  $\mu, n \in \mathbb{N}$ , the complete multigraph  $\mu K_n$  is the complete graph  $K_n$  in which every edge appears  $\mu$  times. We define the union of two graphs  $\mathcal{G}_1, \mathcal{G}_2$  as the multigraph  $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ , where  $V(\mathcal{G}) = V(\mathcal{G}_1) \cup V(\mathcal{G}_2)$  and  $E(\mathcal{G}) = E(\mathcal{G}_1) \sqcup E(\mathcal{G}_2)$ . Given a pair of (multi)graphs  $\mathcal{G}_1, \mathcal{G}_2$ , we say that  $\mathcal{G}_1 \subseteq \mathcal{G}_2$  if  $V(\mathcal{G}_1) \subseteq V(\mathcal{G}_2)$  and  $E(\mathcal{G}_1) \subseteq E(\mathcal{G}_2)$ . For example, the multigraph obtained as the union of the three graphs from Example 4.3 is

 $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$ 

**DEFINITION 4.4** (see [16]). Given a (multi)graph  $\mathcal{G}$  and a positive integer  $n$ , an orthogonal vector  $n$ -coloring of  $\mathcal{G}$  is an assignment of vectors of  $\mathbb{R}^n \setminus \{0\}$  to  $V(\mathcal{G})$  such that if  $\{v, w\} \in E(\mathcal{G})$  then  $v$  and  $w$  receive orthogonal vectors. The vector chromatic number  $\chi_v(\mathcal{G})$  is the least  $n$  such that  $\mathcal{G}$  has an orthogonal vector  $n$ -coloring.

**PROPOSITION 4.5.** Let  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  be the graphical description of a two-orthogonal decomposition  $\mathcal{T} = \sum_{i=1}^r x_i^{(1)} \otimes \cdots \otimes x_i^{(d)} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ . Then  $\chi_v(\mathcal{G}_j) \leq n_j$  for all  $j \in [d]$ , and  $2K_r \subseteq \bigcup_j \mathcal{G}_j$ .

*Proof.* For each  $j \in [d]$ , assigning  $x_i^{(j)}$  to vertex  $i \in [r]$  gives an orthogonal vector  $n_j$ -coloring of  $\mathcal{G}_j$ . The fact that  $2K_r \subseteq \bigcup_j \mathcal{G}_j$  follows from two-orthogonality.  $\square$

In view of Proposition 4.5, we define the notion of a valid graphical description for the two-orthogonal variety  $W \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  to mean a tuple of  $d$  graphs that encode the orthogonality relations of some two-orthogonal tensor in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ .

**DEFINITION 4.6.** A valid graphical description for  $W \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  is a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  such that  $V(\mathcal{G}_1) = \cdots = V(\mathcal{G}_d) = [r]$ ,  $\chi_v(\mathcal{G}_j) \leq n_j$  for all  $j$ , and  $2K_r \subseteq \bigcup_j \mathcal{G}_j$ .

**Remark 4.7.** Lemma 2.10 can be proved using the combinatorial tools in this section. Given two graphs  $\mathcal{G}, \mathcal{G}'$  on the same vertex set, we have  $\chi_v(\mathcal{G} \cup \mathcal{G}') \leq \chi_v(\mathcal{G})\chi_v(\mathcal{G}')$  [16, Lemma 1.14]. Fix  $n_1 \leq \cdots \leq n_d$  and let  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  with  $V(\mathcal{G}_j) = [r]$  for all  $j$  be a valid graphical description for  $W \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ . Since  $2K_r \subseteq \bigcup_{j=1}^d \mathcal{G}_j$  we have  $K_r \subseteq \bigcup_{j=1}^{d-1} \mathcal{G}_j$ , so

$$r = \chi_v(K_r) = \chi_v \left( \bigcup_{j=1}^{d-1} \mathcal{G}_j \right) \leq \chi_v(\mathcal{G}_1) \cdots \chi_v(\mathcal{G}_{d-1}) \leq n_1 \cdots n_{d-1}.$$

Given a valid graphical description for  $W \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$ , we are interested in obtaining the two-orthogonal decompositions that have that graphical description.

**DEFINITION 4.8.** The orthogonal vector  $n$ -coloring ideal of a graph  $\mathcal{G}$  is

$$I_{\mathcal{G},n} = (\langle u_i, u_j \rangle \mid \{i, j\} \in E(\mathcal{G})) + (\langle u_i, u_i \rangle - 1 \mid i \in V(\mathcal{G})) \subset \mathbb{R}[u_{i,1}, \dots, u_{i,n} \mid i \in V(\mathcal{G})].$$

The zero locus  $Z(I_{\mathcal{G},n}) \subseteq (\mathbb{R}^n)^{\times |V(\mathcal{G})|}$  consists of all orthogonal vector  $n$ -colorings of  $\mathcal{G}$  with unit vectors. If  $\chi_v(\mathcal{G}) \leq n$ , then  $Z(I_{\mathcal{G},n}) \neq \emptyset$  by definition. This implies the following.

**PROPOSITION 4.9.** The two-orthogonal variety  $W \subset \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  is equal to the union of  $\overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})}$  over all valid graphical descriptions  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  for  $W$ , where

$$\begin{aligned} \psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)} : \quad & Z(I_{\mathcal{G}_1, n_1}) \times \cdots \times Z(I_{\mathcal{G}_d, n_d}) \times \mathbb{R}^r \rightarrow \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}, \\ & ((x_1^{(1)}, \dots, x_r^{(1)}), \dots, (x_1^{(d)}, \dots, x_r^{(d)}), (\lambda_1, \dots, \lambda_r)) \mapsto \sum_{i=1}^r \lambda_i x_i^{(1)} \otimes \cdots \otimes x_i^{(d)}. \end{aligned}$$

**Remark 4.10.** We cannot guarantee that  $\overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})}$  is irreducible, since  $I_{\mathcal{G},n}$  is not prime in general. For example, for  $\mathcal{G} = ([4], \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\})$ ,  $I_{\mathcal{G},3}$  is generated by

$$\begin{aligned} u_{1,1}u_{2,1} + u_{1,2}u_{2,2} + u_{1,3}u_{2,3}, \quad & u_{2,1}u_{3,1} + u_{2,2}u_{3,2} + u_{2,3}u_{3,3}, \quad u_{2,1}^2 + u_{2,2}^2 + u_{2,3}^2 - 1, \\ u_{1,1}u_{3,1} + u_{1,2}u_{3,2} + u_{1,3}u_{3,3}, \quad & u_{2,1}u_{4,1} + u_{2,2}u_{4,2} + u_{2,3}u_{4,3}, \quad u_{3,1}^2 + u_{3,2}^2 + u_{3,3}^2 - 1, \\ u_{1,1}u_{4,1} + u_{1,2}u_{4,2} + u_{1,3}u_{4,3}, \quad & u_{1,1}^2 + u_{1,2}^2 + u_{1,3}^2 - 1, \quad u_{4,1}^2 + u_{4,2}^2 + u_{4,3}^2 - 1, \end{aligned}$$

and its primary decomposition consists of four components.

**Remark 4.11.** Graphical descriptions give inclusions between components of  $W$ . We define a partial order on the valid graphical descriptions for  $W$  as follows:

$$(\mathcal{G}_1, \dots, \mathcal{G}_d) \preceq (\mathcal{G}'_1, \dots, \mathcal{G}'_d) \text{ if } V(\mathcal{G}_j) = V(\mathcal{G}'_j) \text{ and } E(\mathcal{G}'_j) \subseteq E(\mathcal{G}_j) \text{ for all } j \in [d].$$

If  $(\mathcal{G}_1, \dots, \mathcal{G}_d) \preceq (\mathcal{G}'_1, \dots, \mathcal{G}'_d)$ , then  $\overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})} \subseteq \overline{\text{Im}(\psi_{(\mathcal{G}'_1, \dots, \mathcal{G}'_d)})}$ . The graphical description depends on the order of the summands. To avoid this dependence, the

partial order can be defined up to graph isomorphism. This leads to a notion of maximal valid graphical descriptions. Counting and characterizing the maximal valid graphical descriptions remains an open question.

**4.2. Binary tensors.** In this section, we focus on two-orthogonal tensors in  $V = (\mathbb{R}^2)^{\otimes d}$ . Let  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  be the graphical description of a two-orthogonal decomposition in  $V$ . Then, each  $\mathcal{G}_j$  is a disjoint union of complete bipartite graphs, as follows. If  $u, v, w \in \mathbb{R}^2$ ,  $u \perp v$ , and  $u \perp w$ , then  $v$  and  $w$  are collinear. Therefore, without loss of generality, in this section we will only consider valid graphical descriptions such that each graph is a disjoint union of complete bipartite graphs.

Remark 4.10 shows that  $I_{\mathcal{G},n}$  is not prime in general, so a valid graphical description containing  $\mathcal{G}$  may not parametrize an irreducible variety. The following result shows that these varieties are irreducible for binary tensors.

**PROPOSITION 4.12.** *Fix a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  on vertex set  $[r]$ . Assume that it is a valid graphical description for  $W \subset (\mathbb{R}^2)^{\otimes d}$ , so each  $\mathcal{G}_j$  is a disjoint union of  $c_j$  complete bipartite graphs. Then  $\overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})}$  is irreducible and its expected dimension is  $r + \sum_{j=1}^d c_j$ .*

*Proof.* First, suppose that  $\mathcal{G}$  is a complete bipartite graph with  $V(\mathcal{G}) = [r]$ . Let  $(e_{i_1}, \dots, e_{i_r})$  with  $i_j \in \{1, 2\}$  be an orthogonal vector two-coloring of  $\mathcal{G}$ . Acting with  $\text{SO}(2)$ , we get all orthogonal vector two-colorings of  $\mathcal{G}$  with unit vectors, up to sign. This is an irreducible subvariety of  $Z(I_{\mathcal{G},2})$  of dimension one, let us call it  $Y_{\mathcal{G}}$ . Now suppose that  $\mathcal{G}$  is a disjoint union of complete bipartite graphs, i.e.,  $\mathcal{G} = \mathcal{H}_1 \sqcup \dots \sqcup \mathcal{H}_c$ . Then,  $Y_{\mathcal{G}} = Y_{\mathcal{H}_1} \times \dots \times Y_{\mathcal{H}_c}$  is a product of irreducible varieties, hence irreducible, and its dimension is  $c$ .

Using the previous reasoning for each  $\mathcal{G}_j$ , we get that  $Y_{\mathcal{G}_1} \times \dots \times Y_{\mathcal{G}_d} \times \mathbb{R}^r$  is a product of irreducible varieties, hence irreducible, and its dimension is  $r + \sum_{j=1}^d c_j$ . In the tensor decomposition given by  $\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)}$ , the signs of the vectors can be absorbed by  $\lambda \in \mathbb{R}^r$ , so

$$\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)}(Z(I_{\mathcal{G}_1}) \times \dots \times Z(I_{\mathcal{G}_d}) \times \mathbb{R}^r) = \psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)}(Y_{\mathcal{G}_1} \times \dots \times Y_{\mathcal{G}_d} \times \mathbb{R}^r).$$

Finally, since  $\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)}$  is a polynomial map, the variety  $\overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})}$  is also irreducible, and its expected dimension is  $\dim(Y_{\mathcal{G}_1} \times \dots \times Y_{\mathcal{G}_d} \times \mathbb{R}^r) = r + \sum_{j=1}^d c_j$ .  $\square$

Theorem 1.4 says that  $\dim(W) \geq 2^{d-1} + d$ . We conjecture that for binary tensors we have equality, and the top-dimensional component is  $W_{2^{d-1}} \setminus W_{2^{d-1}-1}$ .

**CONJECTURE 4.13.** *Fix a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  on vertex set  $[r]$ . Assume that it is a valid graphical description for  $W \subset (\mathbb{R}^2)^{\otimes d}$ , so each  $\mathcal{G}_j$  is a disjoint union of  $c_j$  complete bipartite graphs. Then*

$$r + \sum_{j=1}^d c_j \leq 2^{d-1} + d$$

*with equality only if  $r = 2^{d-1}$ .*

Table 1 shows that Conjecture 4.13 is true for  $d = 3$ ; we compute the graphical descriptions of all two-orthogonal decompositions for  $2 \times 2 \times 2$  tensors. Using an analogous approach, one can check that this conjecture is also true for  $d = 4$ . In what follows, we provide further evidence for Conjecture 4.13 and study its consequences.

TABLE 1

Graphical descriptions of the two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$ . Vertices are summands in a two-orthogonal decomposition. An edge in  $\mathcal{G}_k$  represents an orthogonality in the  $k$ th factor. The table shows that  $W_1 \subseteq \text{odeco} \subseteq W_2$  and  $W_3 \setminus W_2 \subset W_4 \setminus W_3$ . See section 5 for more details.

	$\mathcal{G}_1$	$\mathcal{G}_2$	$\mathcal{G}_3$
$W_1$	(1)	(1)	(1)
odeco	(1) — (2)	(1) — (2)	(1) — (2)
$W_2^{\{2,3\}}$	(1) (2)	(1) — (2)	(1) — (2)
$W_2^{\{1,3\}}$	(1) — (2)	(1) (2)	(1) — (2)
$W_2^{\{1,2\}}$	(1) — (2)	(1) — (2)	(1) (2)
$W_3 \setminus W_2$	(1) — (2) 3	(1) — (2) 3	(1) — (2) 3
$W_4 \setminus W_3$	(4) — (3) 1 — 2	(4) — (3) 1 — 2	(4) — (3) 1 — 2

PROPOSITION 4.14. *The complete multigraph  $2K_r$  can be expressed as the union of  $d$  bipartite graphs if and only if  $r \leq 2^{d-1}$ .*

*Proof.* This is equivalent to Theorem 1.3 for binary tensors.  $\square$

PROPOSITION 4.15. *Fix a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  on vertex set  $[2^{d-1}]$ . Assume that it is a valid graphical description for  $W \subset (\mathbb{R}^2)^{\otimes d}$ . Then each  $\mathcal{G}_j$  is a complete bipartite graph  $K_{2^{d-2}, 2^{d-2}}$ . This graphical description satisfies the bound from Conjecture 4.13 with equality.*

*Proof.* First, we show that  $\mathcal{G}_d$  is connected, and the same argument works for any  $\mathcal{G}_j$ . Since each  $\mathcal{G}_j$  is bipartite, they are two-colorable. We label each vertex with a binary string of length  $d-1$  such that if two vertices are connected in  $\mathcal{G}_j$ , their labels differ in the  $j$ th position. All vertices have distinct labels, as follows. Two vertices with the same label are not connected in any of  $\mathcal{G}_1, \dots, \mathcal{G}_{d-1}$ . So even if they were connected in  $\mathcal{G}_d$ , there is at most one edge between them. Hence each of the  $2^{d-1}$  possible labels is used exactly once. Moreover, if two labels differ in only one position, then they are only connected in one of  $\mathcal{G}_1, \dots, \mathcal{G}_{d-1}$ , hence they are connected in  $\mathcal{G}_d$ .

We construct a path between any two vertices  $u$  and  $v$  in  $\mathcal{G}_d$ , as follows. Let  $u$  and  $v$  be labeled by binary strings  $\ell_u$  and  $\ell_v$  in  $\{0, 1\}^{d-1}$ . We can convert  $\ell_u$  into  $\ell_v$  via a sequence of binary strings, each differing from the previous one in only one position. Since every binary string is a vertex in our graph, this is a sequence of vertices in the graph. All vertices with labels that differ in only one position are connected in  $\mathcal{G}_d$ . Hence it is a path in  $\mathcal{G}_d$ .

We have shown that each  $\mathcal{G}_j$  is connected. Let  $A, B$  be two independent sets of  $\mathcal{G}_j$ , so that every edge in  $\mathcal{G}_j$  connects a vertex from  $A$  with a vertex from  $B$ .

Then  $|A| = |B| = 2^{d-2}$ , otherwise we would need to connect more than  $2^{d-2}$  vertices, with at least two edges for every pair of vertices, using only  $d-1$  graphs, a contradiction to Proposition 4.14.  $\square$

**COROLLARY 4.16.** *If the difference  $W_{2^{d-1}} \setminus W_{2^{d-1}-1} \subset (\mathbb{R}^2)^{\otimes d}$  is nonempty, then it consists of basis-aligned tensors (up to closure), so  $\dim(\overline{W_{2^{d-1}} \setminus W_{2^{d-1}-1}}) = 2^{d-1} + d$ .*

*Proof.* The first part of the statement follows from Proposition 4.15. The second part follows from Theorem 1.4 and Proposition 4.12.  $\square$

**COROLLARY 4.17.** *The variety  $\overline{W_{2^{d-1}} \setminus W_{2^{d-1}-1}} \subset (\mathbb{R}^2)^{\otimes d}$  is irreducible.*

*Proof.* The statement is vacuously true if  $W_{2^{d-1}} \setminus W_{2^{d-1}-1} = \emptyset$ . Otherwise, Corollary 4.16 implies that  $W_{2^{d-1}} \setminus W_{2^{d-1}-1}$  consists of basis-aligned tensors. In particular, a two-orthogonal decomposition with  $2^{d-1}$  summands comes from a Latin hypercube. Two Latin cubes are isotopic if one can be obtained from the other by permuting its rows, columns, and symbols. Similarly, two Latin hypercubes are called isotopic if one can be obtained from the other by permuting its mode- $k$  fibers (e.g., rows, columns, tubes, etc.) and symbols. Two-orthogonal decompositions coming from two isotopic Latin squares lie in the same irreducible component since permuting mode- $k$  fibers and symbols can be seen as acting with an element of  $\text{SO}(n)$ . The number of isotopy classes of a Latin  $d$ -cube on  $n=2$  symbols is 1, by induction. The case  $d=2$  (i.e., a Latin square) is well known. A 3-cube can be thought of as a 2-cube on symbols (1, 2) and (2, 1). The same reasoning extends to higher-dimensional cubes.  $\square$

The following two results provide more evidence for Conjecture 4.13.

**PROPOSITION 4.18.** *Fix a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  on vertex set  $[r]$ . Assume that it is a valid graphical description for  $W \subset (\mathbb{R}^2)^{\otimes d}$ . Let  $\mathcal{G}_d$  be a disjoint union of  $c_d$  complete bipartite graphs. Then  $c_d + r \leq 2^{d-1} + 1$  with equality only if  $r = 2^{d-1}$ .*

*Proof.* If  $r = 2^{d-1}$ , then  $c_d = 1$ , by Proposition 4.15. Suppose that  $r < 2^{d-1}$ . Each vertex in  $\mathcal{G}_d$  can be labeled with a binary string of length  $d-1$ , given by two-colorings of  $\mathcal{G}_1, \dots, \mathcal{G}_{d-1}$ . All labels are distinct and labels at Hamming distance one are connected in  $\mathcal{G}_d$ . We show that for every connected component in  $\mathcal{G}_d$  there is a binary string that cannot be a label for a vertex in  $\mathcal{G}_d$ . Consider a Hamiltonian cycle on the hypercube  $\{0, 1\}^{d-1}$ , where vertices of Hamming distance one are connected. This induces a linear ordering  $\ell_1 \prec \dots \prec \ell_{2^{d-1}}$ . Since  $r < 2^{d-1}$  we can assume that  $\ell_1$  does not label any vertex in  $\mathcal{G}_d$ . Given a connected component  $\mathcal{H}_i$  of  $\mathcal{G}_d$ , let  $\ell_{m_i}$  be the minimal label appearing in  $\mathcal{H}_i$ . Then  $\ell_{m_i-1}$  cannot label any vertex in  $\mathcal{G}_d$ . Of course, for different connected components, the corresponding minimal labels are different. Therefore,  $c_d + r \leq 2^{d-1} < 2^{d-1} + 1$ .  $\square$

**COROLLARY 4.19.** *Fix a tuple of graphs  $(\mathcal{G}_1, \dots, \mathcal{G}_d)$  on vertex set  $[r]$ . Assume that it is a valid graphical description for  $W \subset (\mathbb{R}^2)^{\otimes d}$ . Suppose that  $\mathcal{G}_1, \dots, \mathcal{G}_{d-1}$  are connected. If  $\mathcal{G}_d$  is not connected, then*

$$\dim \overline{\text{Im}(\psi_{(\mathcal{G}_1, \dots, \mathcal{G}_d)})} < 2^{d-1} + d.$$

*Proof.* It follows from Proposition 4.12 and Proposition 4.18.  $\square$

The previous results combine to give the following.

**COROLLARY 4.20.** *If Conjecture 4.13 is true, then  $W_{2^{d-1}-1} \neq W_{2^{d-1}}$  and  $\overline{W_{2^{d-1}} \setminus W_{2^{d-1}-1}}$  is an irreducible variety. Moreover,  $\dim \overline{W_{2^{d-1}} \setminus W_{2^{d-1}-1}} = 2^{d-1} + d$  and  $\dim W_{2^{d-1}-1} < 2^{d-1} + d$ .*

## 5. $2 \times 2 \times 2$ tensors.

**5.1. Stratification and equations of the two-orthogonal variety.** In this section we study the two-orthogonal variety  $W \subseteq (\mathbb{R}^2)^{\otimes 3}$ . We describe the varieties in the chain

$$W_1 \subseteq W_2 \subseteq W_3 \subseteq W_4 = W.$$

In Table 1 we saw a combinatorial description of the two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$ . Here we describe this variety algebraically. Consider the Euclidean inner product in  $\mathbb{R}^2$  and let  $\{e_0, e_1\}$  be an orthonormal basis of  $\mathbb{R}^2$ . Using these coordinates, we express a tensor as  $\mathcal{T} = \sum_{i,j,k \in \{0,1\}} t_{ijk} e_i \otimes e_j \otimes e_k$ . The equations shown below were obtained with Macaulay2.

Recall that each of the  $W_r$ 's is invariant under the action of  $G = \mathrm{SO}(2)^{\otimes 3}$  on  $(\mathbb{R}^2)^{\otimes 3}$ . This gives a concise way to describe our varieties: we only give a representative of each orbit. Let  $u \in \mathbb{R}^2$  be an arbitrary unit vector, and let  $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$  be arbitrary scalars.

First, we have  $W_1 = X$ , the cone over the Segre variety. The tensors in  $W_1$  are of the form

$$\lambda_1 e_0 \otimes e_0 \otimes e_0 \in W_1,$$

up to the action of  $G$ . The equations defining  $W_1$  are the  $2 \times 2$  minors of all the flattenings [14, Example 2.11].

Tensors in  $W_2 \setminus W_1$  consist of two summands that are orthogonal in at least two factors. Let us denote by  $W_2^{\{i,j\}}$  the tensors in  $W_2$  whose two-orthogonal decompositions are orthogonal in the  $i$ th and  $j$ th factors. We have three families of tensors:

$$\begin{aligned} \lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 u \otimes e_1 \otimes e_1 &\in W_2^{\{2,3\}}, \\ \lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 e_1 \otimes u \otimes e_1 &\in W_2^{\{1,3\}}, \\ \lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 e_1 \otimes e_1 \otimes u &\in W_2^{\{1,2\}}. \end{aligned}$$

The equations defining each of these irreducible components are

$$W_2^{\{2,3\}} : \quad \begin{cases} -t_{000}t_{101} + t_{100}t_{001} - t_{010}t_{111} + t_{110}t_{011} = 0, \\ -t_{000}t_{110} + t_{100}t_{010} - t_{001}t_{111} + t_{101}t_{011} = 0, \end{cases}$$

and we obtain the equations for the other two by swapping indices. Each component is defined by a subset of the equations defining the oedeco tensors, as discussed in Remark 2.9. The intersection of the three components is the oedeco variety.

Given  $\mathcal{T} = \lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 u \otimes e_1 \otimes e_1 \in W_2$ , to add a third summand to  $\mathcal{T}$  satisfying two-orthogonality we need  $u = \pm e_0$ . We can then add  $\lambda_3 e_1 \otimes e_0 \otimes e_1$  or  $\lambda_3 e_1 \otimes e_1 \otimes e_0$ . Both two-orthogonal decompositions have the same orthogonality pattern between summands, up to permutation. If we repeat this analysis for the other two components of  $W_2$  we also get equivalent tensors. Hence  $\overline{W_3 \setminus W_2}$  is irreducible and consists of tensors of the form

$$\lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 e_0 \otimes e_1 \otimes e_1 + \lambda_3 e_1 \otimes e_0 \otimes e_1.$$

The variety  $\overline{W_3 \setminus W_2}$  is defined by two polynomials. One is Cayley's hyperdeterminant:

TABLE 2  
*Description of the two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$ .*

Variety	Codimension	Degree	#components	Generators per component
$W_1 = X$	4	6	1	9 quadrics
odeco	3	8	1	3 quadrics
$\overline{W_2 \setminus W_1}$	2	12	3	2 quadrics
$\overline{W_3 \setminus W_2}$	2	16	1	2 quartics
$\overline{W_4 \setminus W_3}$	1	4	1	1 quartic
$W = \overline{W_4 \setminus W_3} \cup W_2$	1	4	4	1 quartic / 2 quadrics

$$\begin{aligned} \text{Det}(\mathcal{T}) = & t_{000}^2 t_{111}^2 + t_{001}^2 t_{110}^2 + t_{010}^2 t_{101}^2 + t_{100}^2 t_{011}^2 \\ & - 2t_{000} t_{001} t_{110} t_{111} - 2t_{000} t_{010} t_{101} t_{111} - 2t_{000} t_{100} t_{011} t_{111} \\ & - 2t_{001} t_{100} t_{011} t_{110} - 2t_{001} t_{010} t_{101} t_{110} - 2t_{010} t_{100} t_{011} t_{101} \\ & + 4t_{000} t_{011} t_{101} t_{110} + 4t_{001} t_{010} t_{100} t_{111}. \end{aligned}$$

The other polynomial is a quartic with 40 monomials. This quartic is invariant under the action of  $\text{SO}(2)^{\times 3}$ . In particular, it is invariant under flipping indices  $0 \leftrightarrow 1$ . After such relabeling, there are 5 distinct monomials. Consider the homogenization of the polynomial that defines the ellipotope [23]:

$$(5.1) \quad g(z_1, z_2, z_3, z_4) = 2z_1 z_2 z_3 + z_1^2 z_4 + z_2^2 z_4 + z_3^2 z_4 - z_4^3.$$

Then, the other quartic that defines  $\overline{W_3 \setminus W_2}$  is

$$(5.2) \quad f(\mathcal{T}) = \sum_{(i,j,k) \in \{0,1\}^3} (-1)^{i+j+k} t_{i,j,k} g(t_{i+1,j,k}, t_{i,j+1,k}, t_{i,j,k+1}, t_{i+1,j+1,k+1}),$$

where the sum in the indices is taken modulo 2. Hence,  $\overline{W_3 \setminus W_2} = Z(f) \cap Z(\text{Det})$ .

Finally, the tensors in  $W_4 \setminus W_3$  are of the form

$$\lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 e_0 \otimes e_1 \otimes e_1 + \lambda_3 e_1 \otimes e_0 \otimes e_1 + \lambda_4 e_1 \otimes e_1 \otimes e_0,$$

and  $\overline{W_4 \setminus W_3}$  is the zero locus of  $f$ , from (5.2). The above combines to give the following.

**PROPOSITION 5.1.** *Table 2 describes the stratification of the two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$ . We have the inclusions  $W_1 \subset \text{odeco} \subset W_2$  and  $\overline{W_3 \setminus W_2} \subset \overline{W_4 \setminus W_3}$ . Hence,  $W = \overline{W_4 \setminus W_3} \cup W_2$ .*

Given  $\mathcal{T} \in (\mathbb{R}^2)^{\otimes 3}$ , we seek the closest two-orthogonal tensor to  $\mathcal{T}$ . To do so, we compute the critical points of  $\|\mathcal{T} - \mathcal{S}\|^2$  over  $\mathcal{S} \in W$  and choose the one that minimizes this quantity. The *Euclidean distance degree* (ED degree) of a variety is the number of critical points of the squared Euclidean distance to a general point outside the variety [8]. For example, the ED degree of the variety of rank-one tensors  $X \subset (\mathbb{R}^2)^{\otimes 3}$  is six [12]. Applying the technique in [8], we compute  $\text{EDdegree}(\overline{W_4 \setminus W_3}) = 12$  and  $\text{EDdegree}(\overline{W_2^{i,j}}) = 4$  for all  $i \neq j \in [3]$ . Since the ED degree is additive over the irreducible components, the ED degree of the two-orthogonal variety  $W = \overline{W_4 \setminus W_3} \cup W_2 \subset (\mathbb{R}^2)^{\otimes 3}$  is 24.

**5.2. Generic identifiability of two-orthogonal decompositions.** In Example 2.15 we saw that two-orthogonal decompositions are not always unique. We show that they are unique for sufficiently general  $2 \times 2 \times 2$  tensors.

*Remark 5.2.* There exist symmetric two-orthogonal tensors that do not have a symmetric two-orthogonal decomposition. We will see that the symmetric tensor

$$\mathcal{T} = \lambda e_0 \otimes e_0 \otimes e_0 + e_0 \otimes e_1 \otimes e_1 + e_1 \otimes e_0 \otimes e_1 + e_1 \otimes e_1 \otimes e_0 \in W_4 \cap S^3(\mathbb{R}^2) \subset (\mathbb{R}^2)^{\otimes 3}$$

does not admit a two-orthogonal decomposition with fewer terms for a sufficiently general  $\lambda$ . The real tensors with symmetric two-orthogonal decompositions are odecos. Hence  $\mathcal{T}$  does not admit a symmetric odecos decomposition.

**THEOREM 1.5.** *A generic tensor in  $W \subseteq (\mathbb{R}^2)^{\otimes 3}$  has a unique two-orthogonal decomposition.*

*Proof.* The two-orthogonal variety in  $(\mathbb{R}^2)^{\otimes 3}$  has 4 irreducible components:

$$W = \overline{W_4 \setminus W_3} \cup W_2^{\{2,3\}} \cup W_2^{\{1,3\}} \cup W_2^{\{1,2\}}.$$

We show that a generic tensor from each component has six singular vector tuples and only those appearing in the two-orthogonal decomposition are orthogonal. This implies the uniqueness of the two-orthogonal decomposition. First, let  $\mathcal{T}$  be a generic tensor from  $\overline{W_4 \setminus W_3}$ :

$$\mathcal{T} = \lambda_1 e_0 \otimes e_0 \otimes e_0 + \lambda_2 e_0 \otimes e_1 \otimes e_1 + \lambda_3 e_1 \otimes e_0 \otimes e_1 + \lambda_4 e_1 \otimes e_1 \otimes e_0.$$

The singular vector tuples are  $u \otimes v \otimes w$  such that

$$\det \begin{pmatrix} \lambda_1 v_0 w_0 + \lambda_2 v_1 w_1 & u_0 \\ \lambda_3 v_0 w_1 + \lambda_4 v_1 w_0 & u_1 \end{pmatrix} = \det \begin{pmatrix} \lambda_1 u_0 w_0 + \lambda_3 u_1 w_1 & v_0 \\ \lambda_2 u_0 w_1 + \lambda_4 u_1 w_0 & v_1 \end{pmatrix} = \det \begin{pmatrix} \lambda_1 v_0 u_0 + \lambda_4 v_1 u_1 & w_0 \\ \lambda_2 v_0 u_1 + \lambda_3 v_1 u_0 & w_1 \end{pmatrix} = 0.$$

This gives linear conditions in the monomials  $\{u_i v_j w_k\}$ . Suppose that  $u_0 = 0$  and  $u_1 = 1$ . If  $v_0 = 0$ , then  $w_1 = 0$ , which gives solution  $e_1 \otimes e_1 \otimes e_0$ . If  $v_0 = 1$ , then  $w_0 = v_1 = 0$ , which gives solution  $e_1 \otimes e_0 \otimes e_1$ . The case  $u_1 = 0$  is analogous and we obtain the other two terms in our decomposition.

Now, fix  $u_1 = v_1 = w_1 = 1$  and let  $g$  be the homogenization of the ellipope (5.1). We get

$$\begin{aligned} u_0 - \frac{2\lambda_1\lambda_2\lambda_3 + \lambda_1^2\lambda_4 + \lambda_2^2\lambda_4 + \lambda_3^2\lambda_4 - \lambda_4^3}{\lambda_1^2\lambda_2 - \lambda_2^3 + \lambda_2\lambda_3^2 + 2\lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_4^2} w_0 &= u_0 - \frac{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)}{g(\lambda_1, \lambda_3, \lambda_4, \lambda_2)} w_0 = 0, \\ v_0 - \frac{2\lambda_1\lambda_2\lambda_3 + \lambda_1^2\lambda_4 + \lambda_2^2\lambda_4 + \lambda_3^2\lambda_4 - \lambda_4^3}{\lambda_1^2\lambda_3 + \lambda_2^2\lambda_3 - \lambda_3^3 + 2\lambda_1\lambda_2\lambda_4 + \lambda_3\lambda_4^2} w_0 &= v_0 - \frac{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)}{g(\lambda_1, \lambda_2, \lambda_4, \lambda_3)} w_0 = 0, \\ w_0^2 + \frac{(\lambda_1^2\lambda_3 + \lambda_2^2\lambda_3 - \lambda_3^3 + 2\lambda_1\lambda_2\lambda_4 + \lambda_3\lambda_4^2)(\lambda_1^2\lambda_2 - \lambda_2^3 + \lambda_2\lambda_3^2 + 2\lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_4^2)}{(2\lambda_1\lambda_2\lambda_3 + \lambda_1^2\lambda_4 + \lambda_2^2\lambda_4 + \lambda_3^2\lambda_4 - \lambda_4^3)(\lambda_1^3 - \lambda_1\lambda_2^2 - \lambda_1\lambda_3^2 - 2\lambda_2\lambda_3\lambda_4 - \lambda_1\lambda_4^2)} &= \\ &= w_0^2 - \frac{g(\lambda_1, \lambda_2, \lambda_4, \lambda_3)g(\lambda_1, \lambda_3, \lambda_4, \lambda_2)}{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)g(\lambda_2, \lambda_3, \lambda_4, \lambda_1)} = 0. \end{aligned}$$

So we obtain two solutions:

$$\begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} = \pm \begin{pmatrix} \left( \frac{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)g(\lambda_1, \lambda_2, \lambda_4, \lambda_3)}{g(\lambda_1, \lambda_3, \lambda_4, \lambda_2)g(\lambda_2, \lambda_3, \lambda_4, \lambda_1)} \right)^{\frac{1}{2}} \\ \left( \frac{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)g(\lambda_1, \lambda_3, \lambda_4, \lambda_2)}{g(\lambda_1, \lambda_2, \lambda_4, \lambda_3)g(\lambda_2, \lambda_3, \lambda_4, \lambda_1)} \right)^{\frac{1}{2}} \\ \left( \frac{g(\lambda_1, \lambda_2, \lambda_4, \lambda_3)g(\lambda_1, \lambda_3, \lambda_4, \lambda_2)}{g(\lambda_1, \lambda_2, \lambda_3, \lambda_4)g(\lambda_2, \lambda_3, \lambda_4, \lambda_1)} \right)^{\frac{1}{2}} \end{pmatrix}.$$

This gives six singular vector tuples, the number of singular vector tuples of a generic  $2 \times 2 \times 2$  tensor [12]. For generic  $\lambda_i$ 's we have  $u_0^2, v_0^2, w_0^2 \neq 0, 1$ , which implies that the two new critical two rank-one approximations are not orthogonal to each other or to the summands appearing in the decomposition.

Now let  $\mathcal{T}$  be a generic tensor from  $W_2^{\{1,2\}}$ . The argument works analogously for the other two components of  $W_2$ . After an orthogonal change of basis and rescaling,

$$\mathcal{T} = e_0 \otimes e_0 \otimes e_0 + \lambda e_1 \otimes e_1 \otimes y$$

with  $y = (y_0, 1) \in \mathbb{R}^2$  and  $\lambda \in \mathbb{R}$ . The singular vector tuples are  $u \otimes v \otimes w$  such that

$$\det \begin{pmatrix} v_0 w_0 & u_0 \\ \lambda v_1 \langle y, w \rangle & u_1 \end{pmatrix} = \det \begin{pmatrix} u_0 w_0 & v_0 \\ \lambda u_1 \langle y, w \rangle & v_1 \end{pmatrix} = \det \begin{pmatrix} v_0 u_0 + \lambda y_0 v_1 u_1 & w_0 \\ \lambda u_1 v_1 & w_1 \end{pmatrix} = 0.$$

Suppose that  $u_0 = 1$  and  $u_1 = 0$ . This, implies  $v_1 = w_1 = 0$ , which gives the solution  $e_0 \otimes e_0 \otimes e_0$ . The cases  $v_1 = 0$  and  $w_1 = 0$  are analogous. Now, fix  $u_1 = v_1 = w_1 = 1$ . We get five solutions. One of them is the second summand from the decomposition given before. The other four are

$$\begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} = \begin{pmatrix} \pm \left( \frac{\lambda^2 y_0^2 + \lambda^2 - \lambda y_0}{1 - \lambda y_0} \right)^{\frac{1}{2}} \\ \pm \left( \frac{\lambda^2 y_0^2 + \lambda^2 - \lambda y_0}{1 - \lambda y_0} \right)^{\frac{1}{2}} \\ \frac{\lambda}{1 - \lambda y_0} \end{pmatrix}, \quad \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} = \begin{pmatrix} \mp \left( \frac{\lambda^2 y_0^2 + \lambda^2 + \lambda y_0}{1 + \lambda y_0} \right)^{\frac{1}{2}} \\ \pm \left( \frac{\lambda^2 y_0^2 + \lambda^2 + \lambda y_0}{1 + \lambda y_0} \right)^{\frac{1}{2}} \\ -\frac{\lambda}{1 + \lambda y_0} \end{pmatrix}.$$

For generic  $y_0, \lambda \in \mathbb{R}$ , the only singular vector tuples orthogonal to each other are the ones from the decomposition given before. So the two-orthogonal decomposition is unique.  $\square$

Using an analogous approach, one can verify that a two-orthogonal decomposition with 8 summands in  $(\mathbb{R}^2)^{\otimes 4}$  is unique. It would be interesting to generalize this result to bigger tensors. We present the following conjecture.

**CONJECTURE 5.3.** *A generic two-orthogonal tensor in  $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$  has a unique two-orthogonal decomposition (up to reordering of the summands). In particular, the chain*

$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_r \subseteq \cdots$$

*stabilizes exactly at  $N = \min_{k \in [d]} \prod_{j \neq k} n_j$ . That is,  $W_{N-1} \neq W_N = W_{N+1} = \cdots$ .*

**Open questions.** In this work, we study tensors that can be decomposed via a sequence of critical rank-one approximations. We focus on the decompositions that are order-independent, which we show to be characterized by the two-orthogonal property. In addition to the conjecture above, there are several open directions for future investigation. We show in Proposition 5.1 that a generic  $2 \times 2 \times 2$  tensor does not have a two-orthogonal decomposition. We believe this is also true for larger higher-order tensors, since it is a closed property on the singular vector tuples, but do not know the dimension of the two-orthogonal variety and if it can fill the space of tensors. We propose to study this problem by introducing the notion of valid graphical descriptions, which also leaves open directions of future work. For binary tensors of order  $d$ , we conjecture that the dimension of the two-orthogonal variety is  $2^{d-1} + d$  (Conjecture 4.13) and connect this to a question about bipartite graphs. Finally, we conjecture that the notions of rank and two-orthogonal rank coincide generically up to the generic rank (Conjecture 2.20) and we provide evidence for small tensors.

**Acknowledgments.** We thank Jan Draisma and Àlex Rodríguez García for helpful discussions.

## REFERENCES

- [1] A. ANANDKUMAR, R. GE, AND M. JANZAMIN, *Guaranteed Non-Orthogonal Tensor Decomposition Via Alternating Rank-1 Updates*, preprint, arXiv:1402.5180, 2014.
- [2] A. BORALEVI, J. DRAISMA, E. HOROBET, AND E. ROBEVA, *Orthogonal and unitary tensor decomposition from an algebraic perspective*, Israel J. Math., 222 (2017), pp. 223–260, <https://doi.org/10.1007/s11856-017-1588-6>.
- [3] M. R. BREMNER AND J. HU, *On Kruskal’s theorem that every  $3 \times 3 \times 3$  array has rank at most, 5*, Linear Algebra Appl., 439 (2013), pp. 401–421, <https://doi.org/10.1016/j.laa.2013.03.021>.
- [4] M. V. CATALISANO, A. V. GERAMTIA, AND A. GIMIGLIANO, *Secant varieties of  $\mathbb{P}^1 \times \cdots \times \mathbb{P}^1$  ( $n$ -times) are not defective for  $n \geq, 5$* , J. Algebraic Geom., 20 (2011), pp. 295–327, <https://doi.org/10.1090/S1056-3911-10-00537-0>.
- [5] P. COMON, G. GOLUB, L.-H. LIM, AND B. MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279, <https://doi.org/10.1137/060661569>.
- [6] L. DE LATHAUWER, P. COMON, B. DE MOOR, AND J. VANDEWALLE, *Higher-order power method—Application in independent component analysis*, in Proceedings of the International Symposium on Nonlinear Theory and its Applications, Vol. 1, Research Society of Nonlinear Theory and its Applications, Las Vegas, UT, 1995, pp. 91–96.
- [7] V. DE SILVA AND L.-H. LIM, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1084–1127, <https://doi.org/10.1137/06066518X>.
- [8] J. DRAISMA, E. HOROBET, G. OTTAVIANI, B. STURMFELS, AND R. R. THOMAS, *The Euclidean distance degree of an algebraic variety*, Found. Comput. Math., 16 (2016), pp. 99–149, <https://doi.org/10.1007/s10208-014-9240-x>.
- [9] J. DRAISMA, G. OTTAVIANI, AND A. TOCINO, *Best rank- $k$  approximations for tensors: Generalizing Eckart–Young*, Res. Math. Sci., 5 (2018), 27, <https://doi.org/10.1007/s40687-018-0145-1>.
- [10] C. ECKART AND G. YOUNG, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218, <https://doi.org/10.1007/BF02288367>.
- [11] A. FALCO AND A. NOUY, *A proper generalized decomposition for the solution of elliptic problems in abstract form by using a functional Eckart–Young approach*, J. Math. Anal. Appl., 376 (2011), pp. 469–480, <https://doi.org/10.1016/j.jmaa.2010.12.003>.
- [12] S. FRIEDLAND AND G. OTTAVIANI, *The number of singular vector tuples and uniqueness of best rank-one approximation of tensors*, Found. Comput. Math., 14 (2014), pp. 1209–1242, <https://doi.org/10.1007/s10208-014-9194-z>.
- [13] D. R. GRAYSON AND M. E. STILLMAN, *Macaulay2, A Software System for Research in Algebraic Geometry*, <http://www2.macaulay2.com>.
- [14] J. HARRIS, *Algebraic Geometry, a First Course*, Grad. Texts in Math., Springer, New York, 1992.
- [15] J. HÅSTAD, *Tensor rank is NP-complete*, J. Algorithms, 11 (1990), pp. 644–654, [https://doi.org/10.1016/0196-6774\(90\)90014-6](https://doi.org/10.1016/0196-6774(90)90014-6).
- [16] G. HAYNES, C. PARK, A. SCHAEFFER, J. WEBSTER, AND L. H. MITCHELL, *Orthogonal vector coloring*, Electron. J. Combin., 17 (2010), r55, <https://doi.org/10.37236/327>.
- [17] C. J. HILLAR AND L.-H. LIM, *Most tensor problems are NP-hard*, J. ACM, 60 (2013), pp. 1–39, <https://doi.org/10.1145/2512329>.
- [18] E. HOROBET AND J. I. RODRIGUEZ, *Data loci in algebraic optimization*, J. Pure Appl. Algebra, 226 (2022), 107144, <https://doi.org/10.1016/j.jpaa.2022.107144>.
- [19] E. HOROBET AND E. T. TURATTI, *When does subtracting a rank-one approximation decrease tensor rank?*, Linear Algebra Appl., 709 (2025), pp. 397–415, <https://doi.org/10.1016/j.laa.2025.01.025>.
- [20] T. G. KOLDA, *Orthogonal tensor decompositions*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 243–255, <https://doi.org/10.1137/S0895479800368354>.
- [21] T. G. KOLDA AND J. R. MAYO, *Shifted power method for computing tensor eigenpairs*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1095–1124, <https://doi.org/10.1137/100801482>.
- [22] J. M. LANDSBERG, *Tensors: Geometry and Applications: Geometry and Applications*, Grad. Stud. Math. 128, American Mathematical Society, Providence, RI, 2011.

- [23] M. LAURENT AND S. POLJAK, *On a positive semidefinite relaxation of the cut polytope*, Linear Algebra Appl., 223–224 (1995), pp. 439–461, [https://doi.org/10.1016/0024-3795\(95\)00271-R](https://doi.org/10.1016/0024-3795(95)00271-R).
- [24] T. LICKTEIG, *Typical tensorial rank*, Linear Algebra Appl., 69 (1985), pp. 95–120, [https://doi.org/10.1016/0024-3795\(85\)90070-9](https://doi.org/10.1016/0024-3795(85)90070-9).
- [25] L.-H. LIM, *Singular values and eigenvalues of tensors: A variational approach*, in 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005, IEEE, Piscataway, NJ, 2005, pp. 129–132.
- [26] B. D. MCKAY AND I. M. WANLESS, *A census of small Latin hypercubes*, SIAM J. Discrete Math., 22 (2008), pp. 719–736, <https://doi.org/10.1137/070693874>.
- [27] L. QI, *The best rank-one approximation ratio of a tensor space*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 430–442, <https://doi.org/10.1137/100795802>.
- [28] E. ROBEVA, *Orthogonal decomposition of symmetric tensors*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 86–102, <https://doi.org/10.1137/140989340>.
- [29] E. ROBEVA AND A. SEIGAL, *Singular vectors of orthogonally decomposable tensors*, Linear Multilinear Algebra, 65 (2017), pp. 2457–2471, <https://doi.org/10.1080/03081087.2016.1277508>.
- [30] L. SORBER, M. VAN BAREL, AND L. DE LATHAUWER, *Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank- $(L_r, L_r, 1)$  terms, and a new generalization*, SIAM J. Optim., 23 (2013), pp. 695–720, <https://doi.org/10.1137/120868323>.
- [31] A. STEGEMAN AND P. COMON, *Subtracting a best rank-1 approximation may increase tensor rank*, Linear Algebra Appl., 433 (2010), pp. 1276–1300, <https://doi.org/10.1016/j.laa.2010.06.027>.
- [32] V. STRASSEN, *Rank and optimal computation of generic tensors*, Linear Algebra Appl., 52–53 (1983), pp. 645–685, [https://doi.org/10.1016/0024-3795\(83\)80041-x](https://doi.org/10.1016/0024-3795(83)80041-x).
- [33] T. SUMI, M. MIYAZAKI, AND T. SAKATA, *About the maximal rank of 3-tensors over the real and the complex number field*, Ann. Inst. Statist. Math., 62 (2010), pp. 807–822, <https://doi.org/10.1007/s10463-010-0294-5>.
- [34] A. USCHMAJEW, *A new convergence proof for the higher-order power method and generalizations*, Pac. J. Optim., 11 (2015), pp. 309–321, <http://www.ybook.co.jp/online-p/PJO/vol11/pjov11n2p309.pdf>.
- [35] N. VANNIEUWENHOVEN, J. NICAISE, R. VANDEBRIL, AND K. MEERBERGEN, *On generic nonexistence of the Schmidt–Eckart–Young decomposition for complex tensors*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 886–903, <https://doi.org/10.1137/130926171>.
- [36] K. WANG AND A. SEIGAL, *Lower bounds on the rank and symmetric rank of real tensors*, J. Symbolic Comput., 118 (2023), pp. 69–92, <https://doi.org/10.1016/j.jsc.2023.01.004>.
- [37] WOLFRAM RESEARCH, INC., *MATHEMATICA*, Version 13.3, Champaign, IL, 2024, <https://www.wolfram.com/mathematica>.
- [38] T. ZHANG AND G. H. GOLUB, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550, <https://doi.org/10.1137/S0895479899352045>.