Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Maximum relative distance between symmetric rank-two and rank-one tensors

(revised version: November 2021)

by

Henrik Eisenmann and André Uschmajew

Preprint no.: 27 2021

Maximum relative distance between symmetric rank-two and rank-one tensors

Henrik Eisenmann* André Uschmajew*

Abstract

It is shown that the relative distance in Frobenius norm of a real symmetric order-d tensor of rank two to its best rank-one approximation is upper bounded by $\sqrt{1-(1-1/d)^{d-1}}$. This is achieved by determining the minimal possible ratio between spectral and Frobenius norm for symmetric tensors of border rank two, which equals $(1-1/d)^{(d-1)/2}$. These bounds are also verified for nonsymmetric tensors of order $d=3$.

1 Introduction

It is a well-known fact that the minimal possible ratio between spectral and Frobenius norm of a real $n \times n$ matrix is $1/\sqrt{n}$, and is achieved for any matrix with identical singular values, that is, for multiples of orthogonal matrices. Since the spectral norm of a matrix measures the length of its best rank-one approximation, this statement has the geometric meaning that orthogonal matrices achieve the largest possible relative distance to rank-one matrices. More generally, using singular value decomposition, one can show that the minimal ratio between spectral and Frobenius norm of a rank-k matrix is $1/\sqrt{k}$ and is achieved when all nonzero singular values are equal.

There has been considerable interest in determining the minimal possible ratio between spectral norm $\|A\|_\sigma$ and Frobenius norm $\|A\|_F$ of an $n_1 \times \cdots \times n_d$ tensor A; see, e.g., [14, 10, 8, 12, 13, 1]. As in the matrix case, this ratio measures the distance of A to the set of rank-one tensors, and is hence of both theoretical and practical relevance in problems of low-rank approximation and entanglement. The precise relation between the spectral norm of A and its distance to rank-one tensors is as follows:

$$\min_{\operatorname{rank} B \leq 1} \frac{\|A-B\|_F}{\|A\|_F} = \sqrt{1-\frac{\|A\|_\sigma^2}{\|A\|_F^2}}. \tag{1}$$

Therefore, the minimal possible ratio $\|A\|_\sigma/\|A\|_F$ that can be achieved is also called the best rank-one approximation ratio of the given tensor space [14]. By (1), it expresses the maximum relative distance of a tensor to the set of rank-one tensors.

Despite some recent progress achieved in the aforementioned references and others, determining the best rank-one approximation ratio for tensors remains a difficult problem in general and is largely open. One reason is the lack of a suitable analog to the singular value decomposition. Moreover, the best rank-one approximation ratio of tensors usually differs over the real and complex field, as well as for nonsymmetric and symmetric tensors of the same size.

The available results in the literature focus on the best rank-one approximation ratio in the full tensor space. As for matrices, it would however also be useful to estimate its value in dependence of the tensor rank. In this work we take a first step in this direction. We determine the minimal

*Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany

ratio between spectral and Frobenius norm of real symmetric rank-two tensors of any order, as well as for general real rank-two tensors of order three. For matrices, this value equals $1/\sqrt{2}$. For tensors one should also take into account that the set of tensors of rank at most two is not closed. Our main result is on symmetric tensors and reads as follows.

Theorem 1.1. *Let A be a real symmetric tensor of order $d \geq 3$ and rank at most two. Then*

$$\|A\|_\sigma > \left(1 - \frac{1}{d}\right)^{\frac{d-1}{2}} \|A\|_F \tag{2}$$

and this bound is sharp. In particular,

$$\min_{\substack{A \neq 0 \\ \operatorname{brank} A \leq 2}} \frac{\|A\|_\sigma}{\|A\|_F} = \left(1 - \frac{1}{d}\right)^{\frac{d-1}{2}},$$

where brank *denotes border rank, and the minimum is taken over real symmetric tensors. Up to orthogonal transformation and scaling the minimum is achieved only for the tensor*

$$W_d = \lim_{t \to 0} \frac{1}{t} \left[(e_1 + te_2)^d - e_1^d\right] = de_1^{d-1} e_2.$$

Note that while for symmetric tensors the notions of rank and symmetric rank are not the same in general [15], they coincide for rank-two tensors, see, e.g., [17]. The proof of Theorem 1.1 constitutes the main part of this work and will be given in section 2.

Due to the relation (1) the theorem above is equivalent to the following statement on the maximum relative distance of a real symmetric rank-two tensor to the set of rank-one tensors.

Theorem 1.2. *Let A be a real symmetric tensor of order $d \geq 3$ and rank at most two. Then*

$$\min_{\operatorname{rank} B \leq 1} \frac{\|A - B\|_F}{\|A\|_F} < \sqrt{1 - \left(1 - \frac{1}{d}\right)^{d-1}}$$

and this bound is sharp. Equality is achieved for the tensor W_d as above.

It is interesting to note that for $d \to \infty$ our results imply

$$\min_{\substack{A \neq 0 \\ \operatorname{brank}(A) \leq 2}} \frac{\|A\|_\sigma}{\|A\|_F} \searrow \frac{1}{\sqrt{e}} \approx 0.6065$$

and

$$\max_{\substack{A \neq 0 \\ \operatorname{brank}(A) \leq 2}} \min_{\operatorname{rank} B \leq 1} \frac{\|A - B\|_F}{\|A\|_F} \nearrow \sqrt{1 - \frac{1}{e}} \approx 0.7951.$$

In particular, both quantities are bounded independently of d.

In section 3 we raise the question whether the same bounds hold for general nonsymmetric tensors of rank two. We show that the answer is affirmative for third-order tensors. This result is stated in Theorem 3.2.

Notation

We consider the subspace $\mathrm{Sym}_d(\mathbb{R}^n)$ of real symmetric $n \times \cdots \times n$ tensors $A = [a_{i_1,\dots,i_d}]$ of order d. It inherits the Euclidean inner product $\langle A, B\rangle_F = \sum_{i_1,\dots,i_d} a_{i_1\dots i_d} b_{i_1\dots i_d}$ from the ambient space, which induces the Frobenius norm via

$$\|A\|_F^2 = \langle A, A\rangle_F.$$

It will be convenient to introduce the notation

$$\pm u^d = \pm u \otimes \cdots \otimes u$$

for symmetric rank-one tensors, and similarly

$$u_1 u_2 \dots u_d = \frac{1}{d!} \sum_{\sigma \in \mathfrak{S}_d} u_{\sigma 1} \otimes u_{\sigma 2} \otimes \cdots \otimes u_{\sigma_d}$$

for the symmetrization of a nonsymmetric rank-one tensor $u_1 \otimes u_2 \otimes \cdots \otimes u_d$. It equals the orthogonal projection of $u_1 \otimes u_2 \otimes \cdots \otimes u_d$ onto $\mathrm{Sym}_d(\mathbb{R}^n)$. Specifically, the notation $u^k v^\ell$ denotes the symmetrization of the rank-one tensor $u^{\otimes k} \otimes v^{\otimes \ell}$. For symmetric rank-one tensors u^d and v^d it holds that $\langle u^d, v^d\rangle_F = \langle u, v\rangle^d$ and, therefore, $\|u^d\|_F = \|u\|^d$.

To any symmetric tensor A one associates a homogeneous polynomial

$$p_A(u) = \sum_{i_1,\dots,i_d} a_{i_1\dots i_d} u_{i_1} \dots u_{i_d} = \langle A, u^d\rangle_F.$$

The spectral norm of A is then defined as

$$\|A\|_\sigma = \max_{u \neq 0} \frac{1}{\|u\|^d} |\langle A, u^d\rangle_F| = \max_{u \neq 0} \frac{1}{\|u\|^d} |p_A(u)|.$$

Due to a result of Banach [2], this definition of spectral norm for symmetric tensors is consistent with the general one, which is given in (14) further below. If w is a normalized maximizer of $\frac{1}{\|w\|^d}|p_A(w)|$, then λw^d with $\lambda = p_A(w) = \langle A, w^d\rangle_F$ is a best symmetric rank-one approximation of A in Frobenius norm, that is, it satisfies

$$\|A - \lambda w^d\|_F = \min_{u \in \mathbb{R}^n, \mu \in \mathbb{R}} \|A - \mu u^d\|_F,$$

and vice versa.

A symmetric tensor of rank at most two takes the form

$$A = \alpha u^d - \beta v^d$$

for vectors u, v and scalars $\alpha, \beta \neq 0$, and the rank is equal to two if and only if u and v are linearly independent. Note that the difference notation will turn out to be convenient later. Technically, this defines tensors of *symmetric rank* at most two. But since for rank two both notions of rank coincide [17], we can just use the word rank throughout. It is well known that the set of tensors of rank at most two is not closed [7]. This is also true when restricting to symmetric tensors. The tensors in the closure are said to have border rank at most two, denoted as $\operatorname{brank} A \leq 2$.

2 Proof of the main result

For proving Theorem 1.1 we will determine the infimum value of the optimization problem

$$\inf_{\substack{\alpha,\beta\in\mathbb{R}\\ \|u\|=\|v\|=1}} F(\alpha,\beta,u,v) = \frac{\|\alpha u^d - \beta v^d\|_\sigma^2}{\|\alpha u^d - \beta v^d\|_F^2}. \tag{3}$$

Here we can always additionally assume that $\langle u,v\rangle \geq 0$ and $\alpha > 0$. We will proceed in several steps. First, in section 2.1, we validate that the tensor W_d, which has symmetric border rank two, achieves equality in (2). Hence the infimum in (3) cannot be larger than $(1-\frac{1}{d})^{d-1}$. We next consider in section 2.2 the first-order necessary optimality condition for (3) and show that it cannot be fulfilled for rank-two tensors admitting a unique symmetric best rank-one approximation (Proposition 2.1). In other words, the potential candidates for achieving the infimum in (3) are rank-two tensors with more than one symmetric best rank-one approximation. In section 2.3 we therefore derive a criterion for a symmetric rank-two tensor to have a unique symmetric best rank-one approximation (Proposition 2.3), and validate by hand in sections 2.4 and 2.5 that for tensors which do not satisfy this criterion the value of F is strictly larger than $(1-\frac{1}{d})^{d-1}$. It then remains to show in section 2.6, that among the tensors of border rank two, and up to orthogonal transformation, only tensor W_d achieves the infimum. Taken together, these steps provide a complete proof of Theorem 1.1.

In our proofs we will frequently assume that $\alpha u^d - \beta v^d \in \mathrm{Sym}_d(\mathbb{R}^2)$ since we can always restrict to $\mathrm{Sym}_d(\mathrm{span}\{u,v\})$.

2.1 The ratio for tensor W_d

Recall that $W_d = e_1^{d-1}e_2 = \frac{d}{dt}(e_1+te_2)^d|_{t=0}$. We have $\|W_d\|_F^2 = d$. The spectral norm is given by following optimization problem:

$$\max\ dx^{d-1}y \quad \text{s.t.} \quad x^2+y^2=1.$$

The KKT conditions for this problem lead to the relation

$$(d-1)x^{d-2}y^2 - x^d = 0,$$

that is, either $x=0$, or $x^2=(d-1)y^2$. We find that $x=\sqrt{\frac{d-1}{d}}$ and $y=\frac{1}{\sqrt{d}}$ is a maximizer with the value $\|W_d\|_\sigma = d\left(\frac{d-1}{d}\right)^{(d-1)/2}\frac{1}{\sqrt{d}}$, and therefore

$$\frac{\|W_d\|_\sigma^2}{\|W_d\|_F^2} = \left(1-\frac{1}{d}\right)^{d-1}.$$

2.2 Optimality condition for symmetric rank-two tensors

The target function in (3) can be written as a composition

$$F(\alpha,\beta,u,v) = G(\varphi(\alpha,\beta,u,v))$$

where

$$G\colon \mathrm{Sym}_d(\mathbb{R}^n)\to\mathbb{R}, \quad G(A) = \frac{\|A\|_\sigma^2}{\|A\|_F^2},$$

and

$$\varphi\colon \mathbb{R}\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n \to \operatorname{Sym}_d(\mathbb{R}^n), \quad \varphi(\alpha,\beta,u,v) = \alpha u^d - \beta v^d.$$

While φ is smooth, the map G is not differentiable in all points. However, it is the quotient of the smooth function $A \mapsto \|A\|_F^2$ and the convex function $A \mapsto \|A\|_\sigma^2$. Therefore, the rules for generalized gradients of regular functions are applicable; see [5, Section 2.3]. It follows that the subdifferential of G in a point A can be computed using a quotient rule, which yields

$$\partial G(A) = \frac{2\|A\|_\sigma}{\|A\|_F^4}[\partial(\|A\|_\sigma)\|A\|_F^2 - A\|A\|_\sigma].$$

Here $\partial(\|A\|_\sigma)$ denotes the subdifferential of the spectral norm in A. The derivative of φ equals

$$\varphi'(\alpha,\beta,u,v)[\delta\alpha,\delta\beta,\delta u,\delta v] = u^{d-1}(\alpha d\cdot\delta u + \delta\alpha\cdot u) - v^{d-1}(d\beta\cdot\delta v + \delta\beta\cdot v),$$

which leads to

$$\begin{aligned}&\partial F(\alpha,\beta,u,v)[\delta\alpha,\delta\beta,\delta u,\delta v]\\ &= \frac{2\|A\|_\sigma}{\|A\|_F^4}\langle\partial(\|A\|_\sigma)\|A\|_F^2 - A\|A\|_\sigma, u^{d-1}(\alpha d\cdot\delta u + \delta\alpha\cdot u) - v^{d-1}(d\beta\cdot\delta v + \delta\beta\cdot v)\rangle_F.\end{aligned} \tag{4}$$

The subdifferential of the spectral norm can be characterized as

$$\partial(\|A\|_\sigma) = \operatorname{conv}\operatorname{arg\,max}\{\langle A, X\rangle_F\colon X\in\operatorname{Sym}_d(\mathbb{R}^n),\ \operatorname{rank}X = 1,\ \|X\|_F = 1\}, \tag{5}$$

see [4, Theorem 2.1] in general, and [1, Section 2.3] in particular. In words, $\partial(\|A\|_\sigma)$ equals the convex hull of the normalized symmetric best rank-one approximations of A.

From (4) and (5) one concludes that the first-order optimality condition $0 \in \partial F(\alpha,\beta,u,v)$ (see, e.g., [5, Proposition 2.3.2]) for problem (3) implies that

$$\lambda(\alpha u^d - \beta v^d) \in P_{u,v}\operatorname{conv}\operatorname{arg\,max}\{\langle\alpha u^d - \beta v^d, X\rangle_F\colon X\in\operatorname{Sym}_d(\mathbb{R}^n),\ \operatorname{rank}X = 1,\ \|X\|_F = 1\}, \tag{6}$$

where $P_{u,v}$ is the orthogonal projection onto the linear subspace $\{u^{d-1}\delta u + v^{d-1}\delta v\colon \delta u,\delta v\in\mathbb{R}^n\}$ of $\operatorname{Sym}_d(\mathbb{R}^n)$, and $\lambda\in\mathbb{R}$ is a Lagrange multiplier.

We now show that the optimality condition (6) cannot hold for tensors $\alpha u^d - \beta v^d$ admitting a unique best symmetric rank-one approximation. This is an interesting analogy to the fact that matrices achieving a minimal ratio of spectral and Frobenius norm have equal singular values.

Proposition 2.1. *Let $A = \alpha u^d - \beta v^d$ have rank two. If A has a unique best symmetric rank-one approximation, then A is not a critical point of the optimization problem* (3).

We use the following lemma that shows $P_{u,v}w^d = au^{d-1}w + bv^{d-1}w$ for any $w\in\mathbb{R}^n$ with some $a,b\in\mathbb{R}$.

Lemma 2.2. *Let $\|u\| = \|v\| = 1$. The projection $P_{u,v}w^d$ is given by*

$$\frac{1}{1-\langle u,v\rangle^{2d-2}}\big[(\langle u,w\rangle^{d-1} - \langle u,v\rangle^{d-1}\langle v,w\rangle^{d-1})u^{d-1} + (\langle v,w\rangle^{d-1} - \langle u,v\rangle^{d-1}\langle u,w\rangle^{d-1})v^{d-1}\big]w.$$

Proof. This follows from the definition of orthogonal projection by a direct calculation. □

Proof of Proposition 2.1. Let one of $\pm w^d$ be the normalized best symmetric rank-one approximation of A. Since it is unique, the optimality condition becomes

$$\lambda(\alpha u^d + \beta v^d) \in P_{u,v} w^d. \tag{7}$$

From $p_A(w) = \langle A, w^d \rangle_F \neq 0$ and $A = \alpha u^d + \beta v^d \in \{u^{d-1}\delta u + v^{d-1}\delta v \colon \delta u, \delta v \in \mathbb{R}^n\}$ we have $P_{u,v} w^d \neq 0$, which excludes $\lambda = 0$. By Lemma 2.2, $P_{u,v} w^d = a u^{d-1} w + b v^{d-1} w$ for some $a, b \in \mathbb{R}$. However, since u and v are linearly independent, we have the decomposition

$$\{u^{d-1}\delta u + v^{d-1}\delta v \colon \delta u, \delta v \in \mathbb{R}^n\} = \{u^{d-1}\delta u \colon \delta u \in \mathbb{R}^n\} \oplus \{v^{d-1}\delta v \colon \delta v \in \mathbb{R}^n\}$$

into two complementary subspaces. Therefore, (7) would only be possible if w is both a multiple of u and v, which contradicts the linear independence of u and v. $\square$

2.3 A condition for unique symmetric best rank-one approximation

We now present a class of symmetric rank-two tensors admitting unique best symmetric rank-one approximations. By the result of Proposition 2.1 these can then be excluded from the further discussion on the minimal norm ratio.

Proposition 2.3. *Let*

$$A = \alpha u^d - \beta v^d$$

with $u \neq v$, $\|u\| = \|v\| = 1$, $\langle u, v \rangle \geq 0$ and $\alpha > \beta > 0$. Then A has exactly one best symmetric rank-one approximation.

For the proof we require auxiliary results. One is the following fact about polynomials.

Lemma 2.4. *Let $a, \gamma > 0$ and $b \geq 0$ and $d \geq 2$. The equation $x = \gamma(x-a)(x+b)^{d-1}$ has two real solutions if d is even, and three real solutions if d is odd.*

Proof. Let $p(x) = \gamma(x-a)(x+b)^{d-1} - x$. Then by the intermediate value theorem, p must have at least two real zeros, namely one in the interval $[-b, 0]$ and another one in the interval (a, ∞). On the other hand,

$$p'(x) = \gamma d (x+b)^{d-2} \left(x - \frac{(d-1)a - b}{d} \right) - 1,$$

has at most two sign changes, one at a value larger than $\frac{(d-1)a-b}{d}$ and another at one at a value smaller than $-b$ if d is odd. Therefore, p has at most three real zeros. The statement follows from the fact that the number of real zeros of a polynomial with real coefficients has the same parity as its degree. $\square$

The second lemma narrows the possible locations of maximizers of the homogeneous form $|p_A|$.

Lemma 2.5. *Under the assumptions of Proposition 2.3, let w be a maximizer of $|p_A(w)| = \left|\langle \alpha u^d - \beta v^d, w^d \rangle_F\right|$ subject to $\|w\| = c > 0$. Then $|\langle u, w \rangle| \geq |\langle v, w \rangle|$.*

Proof. Assume to the opposite that $|\langle u, w \rangle| < |\langle v, w \rangle|$ and without loss of generality $\langle v, w \rangle > 0$. Let Q be the symmetric orthogonal matrix mapping u to v and v to u (i.e. $Q = I - z^T z$ with $z = (u+v)/\|u+v\|$), and let $\bar{w} = Qw$. Then $\langle u, w \rangle = \langle v, \bar{w} \rangle$ and $\langle v, w \rangle = \langle u, \bar{w} \rangle$. By assumption, we then have

$$\left|\langle \alpha u^d - \beta v^d, \bar{w}^d \rangle_F\right| = \langle \alpha u^d - \beta v^d, \bar{w}^d \rangle_F.$$

If $\left|\langle \alpha u^d - \beta v^d, w^d \rangle_F\right| = \langle \alpha u^d - \beta v^d, w^d \rangle_F$ this yields $\left|\langle \alpha u^d - \beta v^d, \bar{w}^d \rangle_F\right| > \left|\langle \alpha u^d - \beta v^d, w^d \rangle_F\right|$ (by using $(\alpha+\beta)\langle v, w \rangle^d > (\alpha+\beta)\langle u, w \rangle^d$) which contradicts the optimality of w. In the other case, $\left|\langle \alpha u^d - \beta v^d, w^d \rangle_F\right| = -\langle \alpha u^d - \beta v^d, w^d \rangle_F$, optimality implies $\beta(\langle u, w \rangle^d + \langle v, w \rangle^d) > \alpha(\langle u, w \rangle^d + \langle v, w \rangle^d)$ which contradicts $\alpha > \beta$. $\square$

We are now in the position to prove Proposition 2.3.

Proof of Proposition 2.3. We can assume that $A \in \mathrm{Sym}_d(\mathbb{R}^2)$, so that $u, v \in \mathbb{R}^2$. Without loss of generality, since we can change coordinates, we can consider $\alpha = 1$, $u = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\sqrt[d]{\beta} v = \begin{pmatrix} a \\ b \end{pmatrix}$ with $a > 0$, $b \geq 0$ (since $\langle u, v \rangle \geq 0$), and $a^2 + b^2 < 1$ (since $\beta < \alpha = 1$). Writing $w = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ for points on the unit circle, where $\lambda > 0$ is a normalization constant, we then have

$$p_A(w) = \lambda^d [y^d - (ax + by)^d]. \tag{8}$$

Critical points on the circle are characterized by $\langle w, \nabla p_A(w) \rangle = 0$, which means

$$y^{d-1} x - (bx - ay)(ax + by)^{d-1} = 0$$

independent of λ. Note that here $y = 0$ is not possible since both a and b are nonzero. Recall that a symmetric best rank-one approximation of A is given as $p_A(w) w^d$, where w maximizes $|p_A(w)|$ on the circle. Since $p_A(-w) = (-1)^d p_A(w)$, in order to prove the assertion it suffices to show that $|p_A(w)|$ has exactly one maximizer w with $y = 1$. The optimality condition at such a w reduces to

$$x = (bx - a)(ax + b)^{d-1}. \tag{9}$$

Hence, we only need to show that there is exactly one solution x of this equation corresponding to a global maximum of $|p_A|$ on the unit circle.

If $y = 1$, then p_A in (8) has a zero at $x_0 = \frac{1-b}{a}$. Then

$$x_0 = \frac{1-b}{a} > \frac{b - b^2 - a^2}{a} = (bx_0 - a)(ax_0 + b)^{d-1}.$$

This shows that (9) has at least one solution $x^* > x_0$. We consider such a solution x^* such that the corresponding unit vector $w = \lambda \begin{pmatrix} x^* \\ 1 \end{pmatrix}$ is a local maximum of $|p_A|$ on the unit circle. We have

$$|\langle u, w \rangle| = \lambda < \frac{\lambda}{\sqrt[d]{\beta}} = \frac{\lambda}{\sqrt[d]{\beta}}(ax_0 - b) < \lambda \frac{1}{\sqrt[d]{\beta}}(ax^* - b) = |\langle v, w \rangle|.$$

By Lemma 2.5, w is not a global maximum of $|p_A|$. If d is even, then by Lemma 2.4 equation (9) has exactly two solutions and therefore only one corresponds to a global maximum. If d is odd, then by the same lemma (9) has three solutions. Taking into account that p_A in (8) has only one zero for $y = 1$, one of these solutions corresponds to a local minimizer of $|p_A|$. Hence, there is only one global maximizer. $\square$

2.4 The case $\alpha > 0 \geq \beta$

We show that $\frac{\|\alpha u^d - \beta v^d\|_\sigma^2}{\|\alpha u^d - \beta v^d\|_F^2} \geq \frac{1}{2}$ if $\langle u^d, v^d \rangle_F \geq 0$ and $\alpha > 0 \geq \beta$. This shows that for $d > 2$ such tensors do not attain the infimum in (3) since $\frac{1}{2} > \left(1 - \frac{1}{d}\right)^{d-1}$. We formulate this statement without α and β by removing the restriction $\|u\| = \|v\| = 1$.

Proposition 2.6. *Let $u \neq v$ and $\langle u, v \rangle \geq 0$. Then $\frac{\|u^d + v^d\|_\sigma^2}{\|u^d + v^d\|_F^2} \geq \frac{1}{2}$.*

Proof. We can assume $\|u\| \geq \|v\|$. Using that $\|u^d + v^d\|_\sigma \geq \langle u^d + v^d, \frac{u^d}{\|u\|^d}\rangle_F$, we have

$$\frac{\|u^d + v^d\|_\sigma^2}{\|u^d + v^d\|_F^2} \geq \frac{\|u\|^{2d} + 2\langle u, v\rangle^d + \left(\frac{\langle u,v\rangle}{\|u\|}\right)^{2d}}{\|u\|^{2d} + 2\langle u, v\rangle^d + \|v\|^{2d}} = 1 - \frac{\|v\|^{2d} - \left(\frac{\langle u,v\rangle}{\|u\|}\right)^{2d}}{\|u\|^{2d} + 2\langle u, v\rangle^d + \|v\|^{2d}} \geq 1 - \frac{\|v\|^{2d}}{2\|v\|^{2d}} = \frac{1}{2},$$

as asserted. □

2.5 The case $\alpha = \beta > 0$

In this section we verify by a direct calculation that the infimum in (3) is not attained for the difference of two rank-one tensors with the same norm, i.e. when $\alpha = \beta$ in (3).

Proposition 2.7. *Let $u \neq v$, $\|u\| = \|v\| \neq 0$, $\langle u, v\rangle \geq 0$ and $d \geq 3$. Then*

$$\frac{\|u^d - v^d\|_\sigma^2}{\|u^d - v^d\|_F^2} > \left(1 - \frac{1}{d}\right)^{d-1}.$$

We require the following version of Jensen's inequality.

Lemma 2.8. *Let $f : [a, b] \to \mathbb{R}$ be convex and continuously differentiable. If $a + b = a' + b'$ and $a < a' < b' < b$, then*

$$\frac{1}{b-a}\int_a^b f(x)\,dx \geq \frac{1}{b'-a'}\int_{a'}^{b'} f(x)\,dx \geq f\left(\frac{a+b}{2}\right).$$

The inequalities are strict if f is strictly convex.

Proof. Without loss of generality let $a = -b$ and $a' = -b'$. Then using substitution we have

$$\begin{aligned}
\frac{1}{b}\int_{-b}^{b} f(x)\,dx &= \frac{1}{b'}\int_{-b'}^{b'} f\left(\frac{b}{b'}x\right) - f(x) + f(x)\,dx \\
&= \frac{1}{b'}\int_{-b'}^{b'} f(x)\,dx + \frac{1}{b'}\int_0^{b'}\int_x^{\frac{bx}{b'}} f'(y) - f'(-y)\,dydx \geq \frac{1}{d}\int_{-b'}^{b'} f(x)\,dx,
\end{aligned}$$

by monotonicity of the derivative of a convex function. This shows the first of the asserted inequalities. The second inequality is just Jensen's inequality, noting that $\frac{a+b}{2} = \frac{a'+b'}{2}$. If f is strictly convex, then f' is strictly monotone and the inequalities are strict. □

Proof of Proposition 2.7. We can assume that $A \in \mathrm{Sym}_d(\mathbb{R}^2)$, so that $u, v \in \mathbb{R}^2$. After rotation and rescaling we have $u = \begin{pmatrix}1\\t\end{pmatrix}$ and $v = \begin{pmatrix}1\\-t\end{pmatrix}$ with $t \in (0, 1]$. Then

$$\|u^d - v^d\|_F^2 = 2(1 + t^2)^d - 2(1 - t^2)^d =: g(t). \tag{10}$$

First, we apply the estimate

$$\|u^d - v^d\|_\sigma \geq \left\langle u^d - v^d, \frac{u^d}{\|u\|^d}\right\rangle_F = \frac{(1 + t^2)^d - (1 - t^2)^d}{\sqrt{1 + t^2}^d},$$

which yields

$$\frac{\|u^d - v^d\|_\sigma^2}{\|u^d - v^d\|_F^2} \geq \frac{(1 + t^2)^d - (1 - t^2)^d}{2(1 + t^2)^d} = \frac{1}{2}\left(1 - \left(\frac{1 - t^2}{1 + t^2}\right)^d\right).$$

The right-hand side is monotonically increasing in the interval $(0, 1]$. For $t = \sqrt{\frac{1}{d-1}}$ it equals

$$\frac{1}{2}\left(1 - \left(\frac{d-2}{d}\right)^d\right) = \frac{d^d - (d-2)^d}{2d^d}.$$

This value is larger than $\left(1 - \frac{1}{d}\right)^{d-1} = \left(\frac{d-1}{d}\right)^{d-1}$ since, using Lemma 2.8 with $f(t) = t^{d-1}$, it holds that $d^d - (d-2)^d > 2d(d-1)^{d-1}$ for $d \geq 3$. This shows that

$$\frac{\|u^d - v^d\|_\sigma^2}{\|u^d - v^d\|_F^2} > \left(1 - \frac{1}{d}\right)^{d-1}$$

for all $t \in \left[\sqrt{\frac{1}{d-1}}, 1\right]$. It hence remains to verify this inequality for all $t \in \left(0, \sqrt{\frac{1}{d-1}}\right)$, which is a little bit more involved. The starting point is another lower bound for the spectral norm, namely

$$\|u^d - v^d\|_\sigma \geq \left\langle u^d - v^d, \begin{pmatrix} \sqrt{(d-1)/d} \\ 1/\sqrt{d} \end{pmatrix}^d \right\rangle_F = \frac{1}{\sqrt{d}^d}\left(\left(\sqrt{d-1}+t\right)^d - \left(\sqrt{d-1}-t\right)^d\right) =: h(t).$$

Note that $\frac{u^d - v^d}{\|u^d - v^d\|_F} \to \frac{W_d}{\|W_d\|_F}$ for $t \to 0$. This can be seen by taking the limit of $\frac{u^d - v^d}{t}$ and noting that $g(t) = \|u^d - v^d\|_F^2$ is of order t^2 by (10). We therefore have

$$\lim_{t \to 0} \frac{h(t)^2}{g(t)} = \left\langle \frac{W_d}{\|W_d\|_F}, \begin{pmatrix} \sqrt{(d-1)/d} \\ 1/\sqrt{d} \end{pmatrix}^d \right\rangle_F^2 = \frac{\|W_d\|_\sigma^2}{\|W_d\|_F^2} = \left(1 - \frac{1}{d}\right)^{d-1},$$

where the second and third equalities have been shown in section 2.1. We now claim that

$$\frac{d}{dt}\frac{h(t)^2}{g(t)} > 0 \text{ for } t \in \left(0, \sqrt{\tfrac{1}{d-1}}\right)$$

which then proves the assertion. This claim is equivalent to the positivity of

$$\begin{aligned} \frac{\sqrt{d}^d}{4d}\big(2h'(t)g(t) - g'(t)h(t)\big) = &\left[\left(\sqrt{d-1}+t\right)^{d-1} + \left(\sqrt{d-1}-t\right)^{d-1}\right]\left[(1+t^2)^d - (1-t^2)^d\right] \\ &- t\left[\left(\sqrt{d-1}+t\right)^d - \left(\sqrt{d-1}-t\right)^d\right]\left[(1+t^2)^{d-1} + (1-t^2)^{d-1}\right]. \end{aligned}$$

Elementary manipulations give

$$
\begin{aligned}
&\frac{\sqrt{d}^d}{4d}\big(2h'(t)g(t)-g'(t)h(t)\big)\\
&=\left[\left(\sqrt{d-1}+t\right)^{d-1}(1+t^2)^{d-1}-\left(\sqrt{d-1}-t\right)^{d-1}(1-t^2)^{d-1}\right]\left(1-t\sqrt{d-1}\right)\\
&\quad-\left[\left(\sqrt{d-1}+t\right)^{d-1}(1-t^2)^{d-1}-\left(\sqrt{d-1}-t\right)^{d-1}(1+t^2)^{d-1}\right]\left(1+t\sqrt{d-1}\right)\\
&=\left[\Big(\underbrace{\sqrt{d-1}+t+t^2\sqrt{d-1}+t^3}_{=:b}\Big)^{d-1}-\Big(\underbrace{\sqrt{d-1}-t-t^2\sqrt{d-1}+t^3}_{=:a}\Big)^{d-1}\right]\left(1-t\sqrt{d-1}\right)\\
&\quad-\left[\Big(\underbrace{\sqrt{d-1}+t-t^2\sqrt{d-1}-t^3}_{=:b'}\Big)^{d-1}-\Big(\underbrace{\sqrt{d-1}-t+t^2\sqrt{d-1}-t^3}_{=:a'}\Big)^{d-1}\right]\left(1+t\sqrt{d-1}\right).
\end{aligned}
\tag{11}
$$

Note that for $t\in\left(0,\sqrt{\frac{1}{d-1}}\right)$ we have $b>b'>a'>a$ and

$$
b-a=2t\left(1+t\sqrt{d-1}\right),\quad b'-a'=2t\left(1-t\sqrt{d-1}\right).
$$

Therefore with $f(t)=(d-1)t^{d-2}$, we can rewrite (11) as

$$
\frac{1}{4d}\sqrt{d}^d\big(2h'(t)g(t)-g'(t)h(t)\big)=\frac{1}{2t}\left[(b'-a')\int_a^b f(x)\,dx-(b-a)\int_{a'}^{b'}f(x)\,dx\right].
$$

Moreover,

$$
\frac{a+b}{2}=\sqrt{d-1}+2t^3>\sqrt{d-1}-2t^3=\frac{a'+b'}{2},
$$

and therefore $a'':=\frac{a+b-(b'-a')}{2}>a'>a$ and $b>b'':=\frac{a+b+(b'-a')}{2}>b'$. Since $a''+b''=a+b$ and $a''-b''=a'-b'$, Lemma 2.8 yields

$$
(b'-a')\int_a^b f(x)\,dx\geq(b-a)\int_{a''}^{b''}f(x)\,dx>(b-a)\int_{a'}^{b'}f(x)\,dx,
$$

where the second inequality follows from monotonicity of f. This shows that (11) is positive. □

2.6 Tensors of border rank two

We now consider tensors lying on the boundary of the set of symmetric rank-two tensors.

Proposition 2.9. *Let A be a limit of symmetric rank-two tensors and* $\operatorname{rank} A>2$. *Then*

$$
\frac{\|A\|_\sigma^2}{\|A\|_F^2}\geq\left(1-\frac{1}{d}\right)^{d-1}=\frac{\|W_d\|_\sigma^2}{\|W_d\|_F^2}
$$

and equality is attained if and only if $A=u^{d-1}v$ for some orthogonal u and v, that is, for tensors arising from scaling and orthogonal transformations of tensor W_d.

The boundary of rank-two tensors is well studied. We require the following well-known parametrization, see, e.g., [3]. We offer a self-contained proof for completeness.

Lemma 2.10. *Let A be a limit of symmetric rank-two tensors and* $\operatorname{rank} A > 2$. *Then A is of the form*

$$A = au^d + bdu^{d-1}v$$

with $\langle u, v\rangle = 0$ and $\|u\| = \|v\| = 1$.

Proof. Let $A_n = u_n^d \pm v_n^d$ with $\lim_{n\to\infty} A_n = A$ or $\lim_{n\to\infty} A_n = -A$. It is not difficult to see that u_n and v_n must be unbounded since otherwise there is a subsequence of A_n converging to a tensor of rank at most two, contradicting $\operatorname{rank} A > 2$. We write $v_n = s_n u_n + t_n w_n$ with $\|w_n\| = 1$ and $\langle u_n, w_n\rangle = 0$. Then

$$A_n = (1 \pm s_n^d)u_n^d \pm \sum_{k=1}^{d} \binom{d}{k} s_n^{d-k} t_n^k u_n^{d-k} w_n^k,$$

and it can be checked that all terms are pairwise orthogonal. Hence, since A_n converges, all terms must be bounded and by passing to a subsequence we can assume that all of them converge. Due to $\|u_n\| \to \infty$ we have $1 \pm s_n^d \to 0$ for the first term, which implies that the sequence s_n is bounded. Therefore, considering the term $k = 1$, the sequence $t_n \|u_n\|^{d-1}$ is bounded which automatically implies $t_n^k \|u_n\|^{d-k} \to 0$ for all $k > 1$. We conclude that

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} (1 \pm s_n^d)u_n^d + \lim_{n\to\infty} ds_n^{d-1} t_n u_n^{d-1} w_n = au^d + bdu^{d-1}v$$

which proves the assertion. □

Proof of Proposition 2.9. Using Lemma 2.10, scaling and orthogonal transformations, we can assume $A = ae_1^d + bde_1^{d-1}e_2 \in \operatorname{Sym}_d(\mathbb{R}^2)$ with $a, b \geq 0$. Then $\|A\|_F^2 = a^2 + b^2 d$ since the tensors e_1^d and $e_1^{d-1}e_2$ are orthogonal and $\|de_1^{d-1}e_2\|_F^2 = d$. We have the following two lower bounds for the spectral norm:

$$\|A\|_\sigma \geq \left\langle ae_1^d + bde_1^{d-1}e_2, \frac{1}{\sqrt{d}^d} \begin{pmatrix} \sqrt{d-1} \\ 1 \end{pmatrix}^d \right\rangle_F = \frac{1}{\sqrt{d}^d} \left(a\sqrt{d-1}^d + bd\sqrt{d-1}^{d-1} \right) \tag{12}$$

and

$$\|A\|_\sigma \geq \left\langle ae_1^d + bde_1^{d-1}e_2, e_1^d \right\rangle_F = a. \tag{13}$$

We can restrict to tensors A with Frobenius norm $\|A\|_F^2 = a^2 + b^2 d = 1$ and need to show that

$$\|A\|_\sigma > \left(1 - \frac{1}{d}\right)^{\frac{d-1}{2}}$$

whenever $a > 0$. The first lower bound (12) implies that this is true whenever $b > \frac{\sqrt{d} - a\sqrt{d-1}}{d}$. Together with $1 = a^2 + b^2 d$ and $a, b \geq 0$ this verifies the claim for $0 < a < \frac{2\sqrt{d(d-1)}}{2d-1}$. If $a \geq \frac{2\sqrt{d(d-1)}}{2d-1}$, then the second lower bound (13) yields the desired estimate

$$\|A\|_\sigma^2 \geq a^2 \geq \left(\frac{2\sqrt{d(d-1)}}{2d-1} \right)^2 > \frac{d-1}{d} > \left(1 - \frac{1}{d}\right)^{d-1}$$

for $d \geq 3$. □

This concludes the proof of Theorem 1.1.

3 Approximation ratio for nonsymmetric rank-two tensors

Recall that the spectral norm for general tensors is defined as

$$\|A\|_\sigma = \max_{\|u_1\|=\cdots=\|u_d\|=1} \langle A, u_1 \otimes \cdots \otimes u_d \rangle_F. \tag{14}$$

The result for symmetric tensors raises the following question.

Question 3.1. Is the inequality

$$\|A\|_\sigma > \left(1 - \frac{1}{d}\right)^{\frac{d-1}{2}} \|A\|_F$$

true for real tensors of order $d \geq 3$ and rank at most two?

With some additional effort the results of Proposition 2.1 and Lemma 2.6 can be carried over to nonsymmetric tensors. However, we could neither confirm nor refute the results of sections 2.3 and 2.5 in the nonsymmetric case. Currently we do not know the answer to the above question in general, but we are able to confirm the inequality for $d = 3$ using different arguments.

Theorem 3.2. *Let $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a real third-order tensor of rank at most two. Then*

$$\|A\|_\sigma > \frac{2}{3}\|A\|_F.$$

The inequality is sharp and equality is attained for the tensor $W_3 \in \mathbb{R}^{2\times 2\times 2}$ (which can be embedded in $\mathbb{R}^{n_1 \times n_2 \times n_3}$ if $n_1, n_2, n_3 \geq 2$). Consequently, the maximum relative distance of a real third-order tensor of border rank two to the set of rank-one tensors is

$$\max_{\substack{A \neq 0 \\ \operatorname{brank} A \leq 2}} \min_{\operatorname{rank} B \leq 1} \frac{\|A - B\|_F}{\|A\|_F} = \frac{\sqrt{5}}{3}.$$

Our proof of Theorem 3.2 makes use of two main tools. The first is Cayley's hyperdeterminant, which is positive if and only if a tensor in $\mathbb{R}^{2\times 2\times 2}$ has real rank two [7, Propositions 5.9 and 5.10]. It is also known as Kruskal's polynomial [16]. The second one is the following characterization of real $2 \times 2 \times 2$ tensors of spectral norm one, which is proven in [9, Proposition 2].

Lemma 3.3. *Let $A \in \mathbb{R}^{2\times 2\times 2}$ be a real tensor with $\|A\|_\sigma = 1$. Then after an orthogonal change of basis*

$$A = \left[\begin{array}{cc|cc} 1 & 0 & 0 & a \\ 0 & b & c & d \end{array}\right]$$

with $|a|, |b|, |c| \leq 1$ and $a^2 + b^2 + c^2 + d^2 + 2abc \leq 1$.

Proof of Theorem 3.2. We can assume that $\|A\|_\sigma = 1$ and that A is given as in Lemma 3.3. Cayley's hyperdeterminant states that $\operatorname{rank} A = 2$ if and only if $d^2 + 4abc > 0$. We therefore consider the optimization problem

$$\max_{a,b,c,d} \|A\|_F^2 = 1 + a^2 + b^2 + c^2 + d^2$$

subject to the constraints

$$a^2 + b^2 + c^2 + d^2 + 2abc \leq 1, \quad a^2, b^2, c^2 \leq 1, \quad d^2 + 4abc \geq 0.$$

The KKT conditions for this problem read

$$\begin{pmatrix} 2a \\ 2b \\ 2c \\ 2d \end{pmatrix} + \lambda \begin{pmatrix} 2a+2bc \\ 2b+2ac \\ 2c+2ab \\ 2d \end{pmatrix} + \begin{pmatrix} \lambda_a 2a \\ \lambda_b 2b \\ \lambda_c 2c \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4bc \\ 4ac \\ 4ab \\ 2d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} \lambda(a^2+b^2+c^2+d^2+2abc-1) \\ \mu(d^2+4abc) \\ \lambda_a(a^2-1) \\ \lambda_b(b^2-1) \\ \lambda_c(c^2-1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\lambda, \lambda_a, \lambda_b, \lambda_c \leq 0, \quad \mu \geq 0.$$

First assume $a^2 = 1$. Then $b^2 + c^2 + d^2 + 2abc \leq 0$ and $d^2 + 4abc \geq 0$. This is only possible if $b = c = d = 0$ and we get $\|A\|_F^2 = 2$. The cases $b^2 = 1$ and $c^2 = 1$ give the same result. For the rest of the proof assume $a^2, b^2, c^2 < 1$ and therefore $\lambda_a = \lambda_b = \lambda_c = 0$. If also $\lambda = 0$, then one easily shows $a = b = c = d = 0$ which yields $\|A\|_F = 1$. We therefore can restrict to the case $\lambda < 0$. Basic calculations lead to either $\lambda = -1$ or $a^2 = b^2 = c^2$. We now need to distinguish between $\mu = 0$ and $\mu > 0$.

If $\mu = 0$ and $\lambda = -1$, then $ab = bc = ac = 0$ and therefore $\|A\|_F^2 = 1 + a^2 + b^2 + c^2 + d^2 = 2$. In the case $\mu = 0$ and $\lambda \neq -1$ we have that $d = 0$ and $a^2 = b^2 = c^2$ satisfies $3a^2 \pm 2a^3 = 1$. The possible solutions are $a^2 = \frac{1}{4}$ or $a^2 = 1$, but since we assumed $a^2 < 1$, only the first one applies and yields $\|A\|_F^2 = \frac{7}{4}$.

Now consider $\mu > 0$ and $\lambda = -1$. It follows that $d = 0$, $abc = 0$ and $\|A\|_F^2 = 1 + a^2 + b^2 + c^2 = 2$. The final case is $\mu > 0$ and $a^2 = b^2 = c^2$. Then $d^2 = -4abc$ and hence we obtain $1 = 3a^2 - 2abc = 3a^2 \pm a^3$. As we have already noted this implies $a^2 = \frac{1}{4}$, hence $d^2 = \frac{1}{2}$ and $\|A\|_F^2 = \frac{9}{4}$.

In conclusion, we have shown $\|A\|_F^2 > \frac{9}{4}$ for tensors of rank two since equality is only achieved at the boundary $d^2 + 4abc = 0$. □

Theorem 3.2 suggests an interesting relation between results in [6] and [11]. The authors in [6] found that the minimal possible ratio of spectral and Frobenius norm among all tensors in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ is $\frac{2}{3}$, while in [11] it is shown that the minimal ratio for tensors in $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ is only $\frac{1}{2}$. However, Theorem 3.2 states that border rank-two tensors in $\mathbb{R}^{2\times 2\times 2}$ have the minimal ratio $\frac{2}{3}$. This might be related to the fact that tensors of real rank two and three both have positive volume in $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$, while almost all tensors in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ have complex rank two.

References

[1] A. Agrachev, K. Kozhasov, and A. Uschmajew. Chebyshev polynomials and best rank-one approximation ratio. *SIAM J. Matrix Anal. Appl.*, 41(1):308–331, 2020.

[2] S. Banach. Über homogene Polynome in (L^2). *Studia Math.*, 7:36–44, 1938.

[3] J. Buczyński and J. M. Landsberg. On the third secant variety. *J. Algebraic Combin.*, 40(2):475–502, 2014.

[4] F. H. Clarke. Generalized gradients and applications. *Trans. Amer. Math. Soc.*, 205:247–262, 1975.

[5] F. H. Clarke. *Optimization and nonsmooth analysis.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[6] F. Cobos, T. Kühn, and J. Peetre. Extreme points of the complex binary trilinear ball. *Studia Math.*, 138(1):81–92, 2000.

[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(3):1084–1127, 2008.

[8] H. Derksen, S. Friedland, L.-H. Lim, and L. Wang. Theoretical and computational aspects of entanglement. *arXiv:1705.07160*, 2017.

[9] R. Grząślewicz and K. John. Extreme elements of the unit ball of bilinear operators on l_2^2. *Arch. Math. (Basel)*, 50(3):264–269, 1988.

[10] X. Kong and D. Meng. The bounds for the best rank-1 approximation ratio of a finite dimensional tensor space. *Pac. J. Optim.*, 11(2):323–337, 2015.

[11] T. Kühn and J. Peetre. Embedding constants of trilinear Schatten-von Neumann classes. *Proc. Estonian Acad. Sci. Phys. Math.*, 55(3):174–181, 2006.

[12] Z. Li, Y. Nakatsukasa, T. Soma, and A. Uschmajew. On orthogonal tensors and best rank-one approximation ratio. *SIAM J. Matrix Anal. Appl.*, 39(1):400–425, 2018.

[13] Z. Li and Y.-B. Zhao. On norm compression inequalities for partitioned block tensors. *Calcolo*, 57(1):Paper No. 11, 27, 2020.

[14] L. Qi. The best rank-one approximation ratio of a tensor space. *SIAM J. Matrix Anal. Appl.*, 32(2):430–442, 2011.

[15] Y. Shitov. A counterexample to Comon's conjecture. *SIAM J. Appl. Algebra Geom.*, 2(3):428–443, 2018.

[16] J. M. F. ten Berge. Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika*, 56(4):631–636, 1991.

[17] X. Zhang, Z.-H. Huang, and L. Qi. Comon's conjecture, rank decomposition, and symmetric rank decomposition of symmetric tensors. *SIAM J. Matrix Anal. Appl.*, 37(4):1719–1728, 2016.