SYMMETRIC TENSOR EIGENVALUES AND INEQUALITIES *

ANDREW J. KIM†, DEANNA EASLEY‡, AND TYRUS BERRY§

Abstract. The largest eigenvalue in absolute value of a symmetric matrix is greater than or equal to the absolute value of every entry of the matrix. We introduce a conjecture that the largest eigenvalue in absolute value of any symmetric tensor is greater than or equal to the absolute value of every entry of the tensor. We provide partial analytical results including multiple rigorous bounds on certain tensor entries, and multiple sharpness results including 3-tensors. Moreover, we introduce a framework for computer assisted proofs of these inequalities. Using this framework, we show that the inequality holds for 3-tensors up to two significant figures and for certain entries of higher order tensors up to numerical precision.

Key words. tensor eigenvalues, eigenvalue inequalities, CANDECOMP/PARAFAC (CP), tensor decomposition

AMS subject classifications. 15A69,15A18,15A42

1. Introduction. For a symmetric α-tensor T with largest eigenvalue in absolute value λ_{maxabs}, we seek inequalities of the form

$$\lambda_{\text{maxabs}} \geq c|T_{i_1,\ldots,i_\alpha}|. \tag{1.1}$$

In particular, we are interested in finding the largest c for which this inequality holds for all α-tensors.

In the case of symmetric matrices (i.e. 2-tensors), the inequality (1.1) holds for $c = 1$ since if $T \in \mathbb{R}^{d^2}$ is symmetric, it has an orthogonal eigendecomposition, $T = U^\top \Lambda U$, so by the Cauchy-Schwarz inequality,

$$|T_{ij}| = |\langle u_i, \lambda_j u_j\rangle| \leq ||u_i||\,||\lambda_j u_j|| = |\lambda_j| \leq \lambda_{\text{maxabs}} \tag{1.2}$$

where u_j is an eigenvector of T and λ_j is the associated eigenvalue for $1 \leq j \leq d$, and the identity matrix shows that $c = 1$ is the best possible constant for matrices.

Naturally, this method of proof cannot be generalized to arbitrary tensors due to the lack of a similar rank-1 eigendecomposition [5]. In section 3, we provide an alternative derivation of the fact that $c = 1$ for matrices which can be generalized to tensors. This brings us to make the following conjecture for all symmetric tensors.

Conjecture 1.1. *For all symmetric α-order tensors T with largest eigenvalue in absolute value λ_{maxabs}, then*

$$\lambda_{\text{maxabs}} \geq |T_{i_1,\ldots,i_\alpha}|. \tag{1.3}$$

This conjecture would be extremely useful to a particular theorem for finding approximate tensor CANDECOMP/PARAFAC (CP) decompositions that appears in [1]. For the definition of CP decomposition, see section 2. Below we rewrite this theorem to rely on this conjecture.

---

*

Funding: This work was funded by NSF DMS-2006808

†Class of 2022, Thomas Jefferson High School for Science and Technology, Alexandria, VA (kim.andrew.207@gmail.com).

‡Department of Mathematical Sciences, George Mason University, Fairfax, VA (deasley2@gmu.edu).

§Department of Mathematical Sciences, George Mason University, Fairfax, VA (tberry@gmu.edu)

Theorem 1.2. *Let T be a α-order symmetric tensor with size d, i.e. $T \in \mathbb{R}^{d^\alpha}$. Consider the process of finding an approximate CP decomposition of T by starting from $T_0 = T$ and setting $T_{\ell+1} = T_\ell - \lambda_\ell v_\ell^{\otimes\alpha}$ where λ_ℓ is the largest eigenvalue in absolute value of T_ℓ and v_ℓ is the associated eigenvector. Assuming the conjecture $\lambda_\ell \geq |(T_\ell)_{i_1\ldots i_\alpha}|$, we have $\|T_\ell\|_F \to 0$ and for $r = \sqrt{1 - \frac{1}{d^\alpha}} \in (0,1)$*

$$\frac{\|T_{\ell+1}\|_F}{\|T_\ell\|_F} \leq r \quad \text{and} \quad T = \sum_{\ell=1}^{p} \lambda_\ell v_\ell^{\otimes\alpha} + \mathcal{O}(r^L)$$

for all $L \in \mathbb{N}$.

For completeness, the proof of Theorem 1.2 can be found in Appendix A. A weaker version of Theorem 1.2 was first shown for symmetric 3-tensors and 4-tensors using the weaker inequalities

$$\lambda_{\text{maxabs}} \geq \frac{2}{3+4\sqrt{2}+\sqrt{3}}|T_{ijk}| \qquad \text{and} \qquad \lambda_{\text{maxabs}} \geq \frac{6}{323}|T_{ijk\ell}|,$$

respectively in [1]. If Conjecture 1.1 holds, the theoretical convergence rate for finding the approximate CP decomposition of a symmetric tensor would be significantly improved as shown in Fig. 1.

In this paper we develop a computer assisted technique of developing rigorous bounds of the form (1.1). Through extensive computations we obtain (1.1) with $c = 1.02$ for 3-tensors, and we obtain c even closer to 1 for many entries of higher order tensors. We also introduce some analytical results that show that $c = 1$ is the best possible coefficient for many tensors.

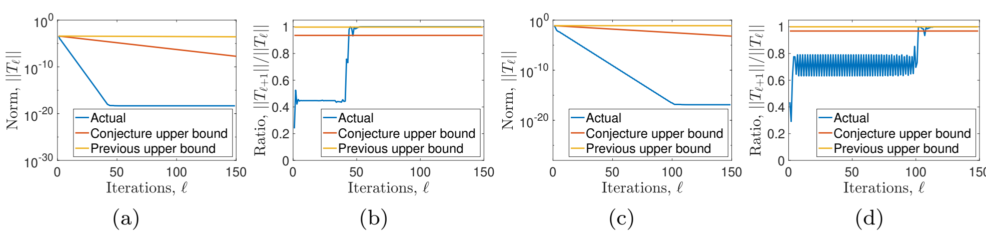


Fig. 1: With $d = 2$ we demonstrate the convergence rate of the approximate CP decomposition of a random symmetric 3-tensor (a,b) and 4-tensor (c,d). The norm of the residual (a,c)(blue) decays to numerical zero linearly and the upper bound assuming Conjecture 1.1 (red) is much stronger than the best previous upper bound (yellow) [1]. The ratio of successive Frobenius norms (b,d)(blue) is always less than the derived upper bounds, r (red and yellow) until the residual norms reach numerical zero.

In Section 2, we review some tensor facts and notation that are necessary for finding tensor eigenvalues and eigenvectors. In Section 3, we present an alternative proof of the inequality (1.2) without using the Cauchy-Schwarz inequality in order to motivate our approach for general tensors. In Section 4, we present a general method of deriving inequalities of the form (1.1) for tensors. Then, in Section 5, we apply the method introduced in Section 4, to validate Conjecture 1.1 for special cases using computer assisted proofs. Finally, In Section 6, we prove the sharpness of our conjecture particularly for 3-tensors and 4-tensors.

2. Tensor Notation. In this section we introduce necessary definitions and notation that will be needed below.

DEFINITION 2.1 (α-order tensor). *For positive integers d and α, a tensor T belonging to $\mathbb{R}^{d^\alpha}$ is called a* **α-order tensor** *or simply a* **α-tensor**.

For example, a $d \times 1$ vector is a 1-tensor and a $d \times d$ matrix is a 2-tensor. For tensors of order $\alpha \geq 3$, we need a different type of multiplication for tensors. For the purposes of this paper, we will only need to know about tensor-vector multiplication, but further discussion on tensor-matrix multiplication can be found in [5].

To motivate tensor-vector multiplication, recall that for a $d \times d$ matrix A and a $d \times 1$ vector v, the matrix-vector multiplication Av is given by $(Av)_i = \sum_{j=1}^{d} A_{ij} v_j$. For a $d \times d \times d$ symmetric 3-tensor T we can similarly define,

$$(T \times_1 v)_{ik} = \sum_{j=1}^{d} T_{jik} v_j$$

which is a matrix. Then naturally,

$$((T \times_1 v) \times_1 v)_k = \sum_{i=1}^{d} \sum_{j=1}^{d} T_{jik} v_j v_i$$

is a vector. This leads to the general definition.

DEFINITION 2.2 (n-mode product of a tensor). *The* **n-mode product** *of a α-order tensor $T \in \mathbb{R}^{d^\alpha}$ with a vector $v \in \mathbb{R}^d$, denoted by $T \times_n v$, is defined elementwise as*

$$(T \times_n v)_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_\alpha} = \sum_{j=1}^{d} T_{i_1,\ldots,i_{n-1},j,i_{n+1},\ldots,i_\alpha} v_j.$$

Note that $T \times_n v \in \mathbb{R}^{d^{\alpha-1}}$, so the order of the resulting tensor is decreased by one. Normally, the choice of index in the tensor-vector multiplication is important, but the results of this paper are restricted to symmetric tensors.

DEFINITION 2.3 (Symmetric Tensor). *A tensor $T \in \mathbb{R}^{d^\alpha}$ is* **symmetric***, if the tensor is invariant to permutations of the indices, i.e.*

$$T_{i_1 \cdots i_\alpha} = T_{p(i_1 \cdots i_\alpha)}$$

for any permutation p.

When a tensor is symmetric, the n-mode product is independent of the mode, i.e. $T \times_n v = T \times_m v$ for any $1 \leq n, m \leq k$. Using this fact, we can have a definition of symmetric tensor eigenvectors and eigenvalues in which we only need one n-mode product. In this case, we choose $n = 1$.

DEFINITION 2.4 (Tensor Eigenvectors and Eigenvalues). *Let $T \in \mathbb{R}^{d^\alpha}$ be a tensor then $v \in \mathbb{R}^d$ is an* **eigenvector** *and $\lambda \in \mathbb{R}$ is the corresponding* **eigenvalue** *of T if*

$$(((T \times_2 v) \times_3 v) \cdots \times_\alpha v) = \lambda v.$$

For a symmetric tensor, the eigenvector-eigenvalue equation is equivalent to

$$\underbrace{(((T \times_1 v) \times_1 v) \cdots \times_1 v)}_{\alpha-1 \text{ times}} = \lambda v$$

since the choice of n-mode product does not affect the definition of a tensor eigenvector. Notice that the tensor eigenvector equation is a degree $\alpha - 1$ polynomial in the components of the eigenvector, namely,

$$\sum_{i_1,\dots,i_{\alpha-1}=1}^{d} T_{i_1,\dots,i_{\alpha-1},i} v_{i_1} \cdots v_{i_{\alpha-1}} = \lambda v_i$$

so finding eigenvectors and eigenvalues analytically is challenging.

3. Matrix Eigenvalue Inequality. Recall the result we showed in the introduction that proved Conjecture 1.1 for symmetric matrices in inequality (1.2). However, notice that (1.2) relied on the Cauchy-Schwarz inequality, which is unique to matrix-vector products. In addition, the proof in (1.2) relies on symmetric matrices having eigendecompositions, and there is no corresponding decomposition of higher order symmetric tensors using tensor eigenvectors [5]. Matrices have a close connection to inner products that higher order tensors do not have, so we need another way of proving this inequality that will generalize to higher order tensors. Thus in this section, we re-prove this basic result for matrices to motivate our approach to tensors.

For a symmetric $d \times d$ matrix A the greatest and least eigenvalue can be defined by the following optimization problems over all possible unit vectors $v \in \mathbb{R}^d$

$$\lambda_{\max} = \max_{||v||=1} v^\top A v = \max_{||v||=1} \sum_{i,j=1}^{d} A_{ij} v_i v_j \qquad \text{and} \qquad \lambda_{\min} = \min_{||v||=1} \sum_{i,j=1}^{d} A_{ij} v_i v_j.$$

Let us fix $s = 1, \dots d$ and let $(e_s)_i = \delta_{is}$, where $\delta_{is} = \begin{cases} 0 & \text{if } \; i \neq s \\ 1 & \text{if } \; i = s \end{cases}$ is the Kronecker delta. Then $\|v_s\| = 1$ so

$$\sum_{i,j} A_{ij} (e_s)_i (e_s)_j = \sum_{i,j} A_{ij} \delta_{is} \delta_{js} = A_{ss}.$$

Hence

$$\lambda_{\min} \leq A_{ss} \leq \lambda_{\max}$$

and thus

$$|\lambda_{\min}| \geq -\lambda_{\min} \geq -A_{ss}.$$

Therefore the largest eigenvalue in absolute value, λ_{maxabs}, is greater than A_{ss} and $-A_{ss}$ so

$$\lambda_{\text{maxabs}} = \max_i |\lambda_i| = \max\{|\lambda_{\max}|, |\lambda_{\min}|\} \geq |A_{ss}|.$$

Next, fix distinct $s, t \in \{1, \dots, d\}$ and let $(w_{s,t})_i = (\delta_{is} + \delta_{it})/\sqrt{2}$. Then $\|(w_{s,t})\| = 1$ and so

$$\sum A_{ij} v_i v_j = \frac{1}{2} \sum A_{ij} (\delta_{is} + \delta_{it})(\delta_{js} + \delta_{jt}) = \frac{1}{2} \left(A_{ss} + 2A_{st} + A_{tt}.\right)$$

Thus

$$\lambda_{\max} \geq \frac{A_{ss} + 2A_{st} + A_{tt}}{2} \tag{3.1}$$

and similarly

$$\lambda_{\min} \leq \frac{A_{ss} + 2A_{st} + A_{tt}}{2} \tag{3.2}$$

therefore

$$\lambda_{\max} \geq \frac{A_{ss} + 2A_{st} + A_{tt}}{2} \geq \lambda_{\min}$$

Finally, using $(w_{s,t})_i = (\delta_{is} - \delta_{it})/\sqrt{2}$ we have

$$\lambda_{\max} \geq \frac{A_{ss} - 2A_{st} + A_{tt}}{2} \tag{3.3}$$

Subtracting (3.2) from (3.1) results in the inequality

$$\lambda_{\max} - \lambda_{\min} \geq 2A_{st}$$

and since

$$2\lambda_{\text{maxabs}} = \lambda_{\text{maxabs}} + \lambda_{\text{maxabs}} \geq \lambda_{\max} - \lambda_{\min}$$

we get $\lambda_{\text{maxabs}} \geq A_{st}$. Likewise, subtracting (3.2) from (3.3) gives us

$$\lambda_{\max} - \lambda_{\min} \geq -2A_{st}$$

hence $\lambda_{\text{maxabs}} \geq -A_{st}$ so together that gives

$$\lambda_{\text{maxabs}} \geq |A_{st}|.$$

If we consider the identity matrix, its only eigenvalue is 1, and its only nonzero entry is 1, meaning this inequality is sharp. Now we have seen a way to prove this inequality that does not rely on the Cauchy-Schwarz inequality. It just relies on this maximum property of eigenvalues. We can now generalize this derivation to tensors of order greater than 2.

4. α-Tensor Eigenvalue Inequality. In this section, we will explore the various cases of Conjecture 1.1. First, we introduce a new notation for entries of the tensor $T \in \mathbb{R}^{d^\alpha}$ with repeated indices. We will denote the entry where we have only one index that is repeated, i.e. $T_{s\ldots s}$, by T_{s^α} for some fixed $s \in \{1,\ldots,d\}$. Then in the case we have two distinct indices with different numbers of repeats we fix distinct $s,t \in \{1,\ldots,d\}$ and denote the entry as $T_{s^i t^j}$ with i terms that are s, j terms that are t, and $i+j=\alpha$. For example, we would denote the entry $T_{s\ldots st}$ by $T_{s^{\alpha-1}t}$ or the entry $T_{s\ldots stt}$ by $T_{s^{\alpha-2}t^2}$. When we have three distinct indices we fix distinct $s,t,u \in \{1,\ldots,d\}$ and denote the entry as $T_{s^i t^j u^k}$ with s being repeated i times, t repeated j times, u repeated k times, and $i+j+k=\alpha$. More generally, we could have the entry of T with ℓ distinct indices be $T_{s_1^{i_1} s_2^{i_2} \ldots s_\ell^{i_\ell}}$ where $s_1, s_2, \ldots, s_\ell \in \{1,\ldots,d\}$ are all distinct and $i_1 + i_2 + \cdots + i_\ell = \alpha$. Note that this is well-defined since T is symmetric and independent of the order of the indices.

4.1. Identical indices. We begin by discussing the case where all indices of T are the same. The maximum eigenvalue, $\lambda_{\max}$, and minimum eigenvalue, $\lambda_{\min}$, are defined by the equations

$$\lambda_{\max} = \max_{||v||=1} \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} v_{i_1} v_{i_2} \cdots v_{i_\alpha} \tag{4.1}$$

$$\lambda_{\min} = \min_{||v||=1} \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} v_{i_1} v_{i_2} \cdots v_{i_\alpha} \tag{4.2}$$

respectively where $v \in \mathbb{R}^d$ is any unit vector. Take $(e_s)_i = \delta_{is}$. Then $\|e_s\| = 1$ so

$$\max_{\|v\|=1} \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} v_{i_1} v_{i_2} \cdots v_{i_\alpha} \geq \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} (e_s)_{i_1} (e_s)_{i_2} \cdots (e_s)_{i_\alpha} = T_{s^\alpha}$$
$$\min_{\|v\|=1} \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} v_{i_1} v_{i_2} \cdots v_{i_\alpha} \leq \sum_{i_1,\dots,i_\alpha} T_{i_1 i_2 \dots i_\alpha} (e_s)_{i_1} (e_s)_{i_2} \cdots (e_s)_{i_\alpha} = T_{s^\alpha}.$$

Thus

$$\lambda_{\max} \geq T_{s^\alpha} \geq \lambda_{\min}.$$

The latter inequality tells us that

$$|\lambda_{\min}| \geq -\lambda_{\min} \geq -T_{s^\alpha}.$$

Therefore, if we call λ_{maxabs} the greatest absolute value over all eigenvalues,

$$\lambda_{\text{maxabs}} \geq |T_{s^\alpha}|.$$

Thus the inequality (1.1) holds with $c = 1$ for all diagonal entries of a tensor. In the next two sections we derive inequalities for certain types of non-diagonal entries.

4.2. Two distinct indices. Next, we consider the inequality for entries of the form, $T_{s^{\alpha-1}t}$, meaning that instead of all the indices being identical (diagonal entries), there are two distinct indices. We can derive an inequality by considering any unit vector of the form

$$(w_{s,t})_i = \frac{a\delta_{is} + b\delta_{it}}{\sqrt{a^2+b^2}}.$$

From the definition of the largest tensor eigenvalue, $\lambda_{\max}$, we have

$$\begin{aligned}
\lambda_{\max} &\geq \sum_{i_1,\dots,i_\alpha} T_{i_1 \dots i_\alpha} (w_{s,t})_{i_1} \dots (w_{s,t})_{i_\alpha} \\
&= \left(\frac{1}{\sqrt{a^2+b^2}}\right)^\alpha \left(a^0 b^\alpha \binom{\alpha}{0} T_{s..s} + a^1 b^{\alpha-1} \binom{\alpha}{1} T_{s\dots st} + a^2 b^{\alpha-2} \binom{\alpha}{2} T_{s\dots stt} + \cdots \right. \\
&\qquad \left. \cdots + a^{\alpha-1} b^1 \binom{\alpha}{\alpha-1} T_{st\dots t} + a^\alpha b^0 \binom{\alpha}{\alpha} T_{t\dots t} \right) \\
&= \left(\frac{1}{\sqrt{a^2+b^2}}\right)^\alpha \sum_{i=0}^{\alpha} a^i b^{\alpha-i} \binom{\alpha}{i} T_{s^{\alpha-i}t^i}
\end{aligned}$$

and similarly we also have,

$$\lambda_{\min} \leq \left(\frac{1}{\sqrt{a^2+b^2}}\right)^\alpha \sum_{i=0}^{\alpha} a^i b^{\alpha-i} \binom{\alpha}{i} T_{s^{\alpha-i}t^i}$$

where $\lambda_{\min}$ is the smallest (most negative) eigenvalue. Now for any positive constant $c > 0$ we have

$$c\lambda_{\max} \geq c\left(\frac{1}{\sqrt{a^2+b^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a^i b^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}$$

and for any negative constant $c < 0$ we have

$$c\lambda_{\min} \geq c\left(\frac{1}{\sqrt{a^2+b^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a^i b^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}.$$

Thus, defining $\lambda_{\text{maxabs}} = \max\{|\lambda_{\max}|, |\lambda_{\min}|\}$ we have $\lambda_{\text{maxabs}} \geq |\lambda_{\max}|$ and $\lambda_{\text{maxabs}} \geq |\lambda_{\min}|$, so for any c (positive or negative) we have,

$$|c|\lambda_{\text{maxabs}} \geq |c||\lambda_{\max}| \geq c\lambda_{\max}$$
$$|c|\lambda_{\text{maxabs}} \geq |c||\lambda_{\min}| \geq c\lambda_{\min}$$

which implies that

$$|c|\lambda_{\text{maxabs}} \geq c\left(\frac{1}{\sqrt{a^2+b^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a^i b^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}$$

so we no longer have to restrict to positive c.

We now consider $\alpha+1$ distinct vectors,

$$(w^0_{s,t})_i = \frac{a_0\delta_{is}+b_0\delta_{it}}{\sqrt{a_0^2+b_0^2}}, \quad (w^1_{s,t})_i = \frac{a_1\delta_{is}+b_1\delta_{it}}{\sqrt{a_1^2+b_1^2}}, \quad \cdots, \quad (w^{\alpha}_{s,t})_i = \frac{a_{\alpha}\delta_{is}+b_{\alpha}\delta_{it}}{\sqrt{a_0^2+b_0^2}},$$

which yields the following system of inequalities,

$$\begin{aligned}
|c_0|\lambda_{\text{maxabs}} &\geq c_0\left(\frac{1}{\sqrt{a_0^2+b_0^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a_0^i b_0^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}\\
|c_1|\lambda_{\text{maxabs}} &\geq c_1\left(\frac{1}{\sqrt{a_1^2+b_1^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a_1^i b_1^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}\\
&\vdots\\
|c_{\alpha}|\lambda_{\text{maxabs}} &\geq c_{\alpha}\left(\frac{1}{\sqrt{a_{\alpha}^2+b_{\alpha}^2}}\right)^{\alpha}\sum_{i=0}^{\alpha} a_{\alpha}^i b_{\alpha}^{\alpha-i}\binom{\alpha}{i}T_{s^{\alpha-i}t^i}
\end{aligned}$$

Adding all the above inequalities note that the left hand sides add to

$$(|c_0|+|c_1|+\cdots+|c_{\alpha}|)\,\lambda_{\text{maxabs}} = \lambda_{\text{maxabs}}||\vec{c}||_1$$

where $\vec{c} = \begin{pmatrix} c_0 & c_1 & \cdots & c_{\alpha}\end{pmatrix}^{\top}$, and the 1-norm is defined as $||\vec{c}||_1 = \sum_{i=0}^{\alpha}|c_i|$. Similarly, adding the right-hand-sides of all the above inequalities results in the matrix-vector product

$$\begin{pmatrix} c_0\\ c_1\\ \vdots\\ c_{\alpha}\end{pmatrix}^{\top}
\begin{pmatrix}
\binom{\alpha}{0}\frac{b_0^{\alpha}}{(a_0^2+b_0^2)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_0 b_0^{\alpha-1}}{(a_0^2+b_0^2)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_0^{\alpha}}{(a_0^2+b_0^2)^{\alpha/2}}\\
\binom{\alpha}{0}\frac{b_1^{\alpha}}{(a_1^2+b_1^2)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_1 b_1^{\alpha-1}}{(a_1^2+b_1^2)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_1^{\alpha}}{(a_1^2+b_1^2)^{\alpha/2}}\\
\vdots & \vdots & \ddots & \vdots\\
\binom{\alpha}{0}\frac{b_{\alpha}^{\alpha}}{(a_{\alpha}^2+b_{\alpha}^2)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_{\alpha} b_{\alpha}^{\alpha-1}}{(a_{\alpha}^2+b_{\alpha}^2)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_{\alpha}^{\alpha}}{(a_{\alpha}^2+b_{\alpha}^2)^{\alpha/2}}
\end{pmatrix}
\begin{pmatrix} T_{s^{\alpha}}\\ T_{s^{\alpha-1}t}\\ \vdots\\ T_{t^{\alpha}}\end{pmatrix} \tag{4.3}$$

For simplicity we set $b_i = 1$ for $i = 0, ..., \alpha$ so that we have the matrix-vector inequality,

$$||\vec{c}||_1 \lambda_{\text{maxabs}} \geq \vec{c}^\top A\vec{T}$$

where $\vec{T}_i = T_{s^{\alpha-i+1}t^{i-1}}$ for $j = 1, ..., \alpha+1$ and

$$A = \begin{pmatrix} \binom{\alpha}{0}\frac{1}{(a_0^2+1)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_0}{(a_0^2+1)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_0^\alpha}{(a_0^2+1)^{\alpha/2}} \\ \binom{\alpha}{0}\frac{a_1^\alpha}{(a_1^2+1)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_1}{(a_1^2+1)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_1^\alpha}{(a_1^2+1)^{\alpha/2}} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{\alpha}{0}\frac{1}{(a_\alpha^2+1)^{\alpha/2}} & \binom{\alpha}{1}\frac{a_\alpha}{(a_\alpha^2+1)^{\alpha/2}} & \cdots & \binom{\alpha}{\alpha}\frac{a_\alpha^\alpha}{(a_\alpha^2+1)^{\alpha/2}} \end{pmatrix}. \tag{4.4}$$

We note that the matrix A is closely connected to the Vandermonde matrix and we have the following immediate result.

THEOREM 4.1. *Assuming that $a_0, ..., a_\alpha$ are distinct, the matrix A as defined in (4.4) is invertible and its inverse matrix is described by*

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}(a_j^2+1)^{\alpha/2} \sum\limits_{0 \leq i_1 < \cdots < i_{\alpha+1-i} \leq \alpha+1} a_{i_1}a_{i_2}\cdots a_{i_{j-1}}a_{i_{j+1}}\cdots a_{\alpha+1-i}}{\binom{\alpha}{i}\prod\limits_{\ell<j}^{\alpha+1}(a_j - a_\ell)}.$$

We include the proof of Theorem 4.1 in Appendix B. Now that we know the matrix A is invertible we have the following eigenvalue inequality which is of the form 1.1 as desired.

THEOREM 4.2. *Let $a_0, a_1, ..., a_\alpha$ be distinct and let A be as defined in (4.4), then for any $i \in \{0, ..., \alpha\}$ we have*

$$\lambda_{\text{maxabs}} \geq \frac{|T_{s^{\alpha-i}t^i}|}{||\vec{e}_{i+1}^\top A^{-1}||_1} \tag{4.5}$$

Proof. By (4.3), we have $\|\vec{c}\|_1\lambda_{\text{maxabs}} \geq \vec{c}^\top A\vec{T}$ and since A is invertible we can choose $\vec{c} = A^{-\top}\vec{e}_{i+1}$ (where $\vec{e}_{i+1}$ is the standard unit vector) so that

$$\vec{c}^\top A = \vec{e}_{i+1}^\top.$$

and (4.3) becomes,

$$||\vec{e}_{i+1}^\top A^{-1}||_1\lambda_{\text{maxabs}} \geq \vec{e}_{i+1}^\top\vec{T} = \vec{T}_{i+1} = T_{s^{\alpha-i}t^i}$$

and similarly for $\vec{c} = -A^{-\top}\vec{e}_{i+1}$ we have,

$$||\vec{e}_{i+1}^\top A^{-1}||_1\lambda_{\text{maxabs}} \geq -\vec{e}_{i+1}^\top\vec{T} = -\vec{T}_{i+1} = -T_{s^{\alpha-i}t^i}$$

and together these inequalities yield the result. ∎

Theorem 4.2 gives us a method of finding bounds on the best possible coefficient c in equation (1.1). Define

$$c_{\text{opt}} = \min_T \left\{ \frac{\lambda_{\text{maxabs}}(T)}{|T_{i_1,...,i_\alpha}|} \right\}$$

where $\lambda_{\text{maxabs}}(T)$ is the largest eigenvalue of T in absolute value. Then for any choice of $a_0, ..., a_\alpha$ we have

$$c_{\text{opt}} \geq \frac{1}{||e_{i+1}^\top A^{-1}||_1}.$$

Conjecture 1.1 claims that equation (1.1) holds with $c = 1$. In section 5 we use numerical optimization methods to find the values of $a_0, ..., a_\alpha$ that minimize $||e_{i+1}^\top A^{-1}||_1$ in order to verify Conjecture 1.1 on several examples.

4.3. Three distinct indices. Next we consider the case of entries of T that have three different indices, $T_{s^i t^j u^{\alpha-i-j}}$. To obtain similar bounds we consider test vectors that have up to three nonzero entries and can be written in the form,

$$(w_{s,t,u})_i = \frac{a\delta_{is} + b\delta_{it} + d\delta_{iu}}{\sqrt{a^2 + b^2 + d^2}}$$

and substituting this vector in (4.1) we have

$$\begin{aligned}
\lambda_{\max} &\geq \sum_{i_1,...,i_\alpha} T_{i_1...i_\alpha}(w_{s,t,u})_{i_1}...(w_{s,t,u})_{i_\alpha} \\
&= \left(\frac{1}{\sqrt{a^2+b^2+d^2}}\right)^\alpha \sum_{i=0}^{\alpha}\sum_{j=0}^{\alpha-i} a^i b^j d^{\alpha-i-j}\binom{\alpha}{i, j, \alpha-i-j} T_{s^i t^j u^{\alpha-i-j}} \\
&\geq \lambda_{\min}
\end{aligned}$$

We can see that the total number of possible permutations of the numbers $i, j, \alpha-i-j$ is given by

$$m = (\alpha+1) + \alpha + \cdots + 2 + 1 = \frac{(\alpha+1)(\alpha+2)}{2}.$$

Now, just as we did in the two distinct indices case, we can construct an $m \times m$ matrix. Combining the above inequalities with multipliers $c_0, ..., c_m$, we have the system of equations,

$$\begin{aligned}
|c_0|\lambda_{\text{maxabs}} &\geq c_0 \left(\frac{1}{\sqrt{a_0^2+b_0^2+d_0^2}}\right)^\alpha \sum_{i=0}^{\alpha}\sum_{j=0}^{\alpha-i} a_0^i b_0^j d_0^{\alpha-i-j}\binom{\alpha}{i, j, \alpha-i-j} T_{s^i t^j u^{\alpha-i-j}} \\
|c_1|\lambda_{\text{maxabs}} &\geq c_1 \left(\frac{1}{\sqrt{a_1^2+b_1^2+d_1^2}}\right)^\alpha \sum_{i=0}^{\alpha}\sum_{j=0}^{\alpha-i} a_1^i b_1^j d_1^{\alpha-i-j}\binom{\alpha}{i, j, \alpha-i-j} T_{s^i t^j u^{\alpha-i-j}} \\
&\vdots \\
|c_m|\lambda_{\text{maxabs}} &\geq c_m \left(\frac{1}{\sqrt{a_m^2+b_m^2+d_m^2}}\right)^\alpha \sum_{i=0}^{\alpha}\sum_{j=0}^{\alpha-i} a_m^i b_m^j d_m^{\alpha-i-j}\binom{\alpha}{i, j, \alpha-i-j} T_{s^i t^j u^{\alpha-i-j}}
\end{aligned}$$

Adding the above inequalities, the left hand sides sums to

$$(|c_0| + |c_1| + \cdots + |c_m|)\,\lambda_{\text{maxabs}} = ||\vec{c}||_1 \lambda_{\text{maxabs}}.$$

Expressing the sum of the right hand sides as a matrix yields,

(4.6)

$$\vec{c}^{\top}\begin{pmatrix} \binom{\alpha}{\alpha,0,0}\frac{a_0^{\alpha}}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_0^{\alpha-1}b_0}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_0^{\alpha}}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} \\ \binom{\alpha}{\alpha,0,0}\frac{a_1^{\alpha}}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_1^{\alpha-1}b_1}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_1^{\alpha}}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{\alpha}{\alpha,0,0}\frac{a_m^{\alpha}}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_m^{\alpha-1}b_m}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_m^{\alpha}}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} \end{pmatrix}\begin{pmatrix} T_{s^{\alpha}} \\ T_{s^{\alpha-1}t} \\ \vdots \\ T_{st^{\alpha-1}} \\ T_{t^{\alpha}} \\ T_{s^{\alpha-1}u} \\ \vdots \\ T_{u^{\alpha}} \end{pmatrix}$$

To clarify the ordering, consider the entry $T_{s^{\alpha-i-j}t^iu^j}$, when $i=j=0$ we get the first entry $T_{s^{\alpha}}$ and then keeping $i=0$ we continue with $j=0,1,2,...,\alpha$ for a total of $\alpha+1$ entries with $i=0$. Next we move to $i=1$ and j goes from 0 to $\alpha-1$ for a total of α entries with $i=1$. Then when $i=2$, j goes from 0 to $\alpha-2$ for a total of $\alpha-1$ entries with $i=2$. So each block with a fixed i value has $\alpha-i+1$ entries. Thus in order to get to the block of entries that have subscript t^i we must pass $(\alpha+1)+(\alpha)+(\alpha-1)+\cdots+(\alpha-i+2)$ entries corresponding to all entries having subscript t^{ℓ} where $\ell<i$. Once we reach the block of entries having t^i, in order to reach the entry $T_{s^{\alpha-i-j}t^iu^j}$ we simply go an additional j entries further down the list. Thus the index of $T_{s^{\alpha-i-j}t^iu^j}$ is given by,

$$\begin{aligned} I_{\alpha}(i,j) &= (\alpha+1)+(\alpha)+(\alpha-1)+\cdots+(\alpha-i+2)+j \\ &= \sum_{\ell=0}^{i-1}\alpha+1-\ell \\ &= i(\alpha+1)-\frac{(i-1)i}{2} \end{aligned}$$

and we define a vector $\tilde{T}_{I_{\alpha(i,j)}} = T_{s^{\alpha-i-j}t^iu^j}$ and we define the matrix,

(4.7)

$$\tilde{A} = \begin{pmatrix} \binom{\alpha}{\alpha,0,0}\frac{a_0^{\alpha}}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_0^{\alpha-1}b_0}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_0^{\alpha}}{(a_0^2+b_0^2+d_0^2)^{\alpha/2}} \\ \binom{\alpha}{\alpha,0,0}\frac{a_1^{\alpha}}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_1^{\alpha-1}b_1}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_1^{\alpha}}{(a_1^2+b_1^2+d_1^2)^{\alpha/2}} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{\alpha}{\alpha,0,0}\frac{a_m^{\alpha}}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} & \binom{\alpha}{\alpha-1,1,0}\frac{a_m^{\alpha-1}b_m}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} & \cdots & \binom{\alpha}{0,0,\alpha}\frac{d_m^{\alpha}}{(a_m^2+b_m^2+d_m^2)^{\alpha/2}} \end{pmatrix}$$

which leads to the following result.

THEOREM 4.3. *For any a_i, b_i and d_i for any $i\in\{0,...,\alpha\}$, let $\tilde{A}$ be as stated in (4.7), then assuming $\tilde{A}$ is invertible we have*

$$\lambda_{\text{maxabs}} \geq \frac{|T_{s^{\alpha-i-j}t^iu^j}|}{||e_{I_{\alpha}}^{\top}\tilde{A}^{-1}||_1} \tag{4.8}$$

The proof of Theorem 4.3 is identical to that of Theorem 4.2.

Theorem 4.3 yields a method of finding lower bounds for the coefficient c in equation (1.1). Similarly to the previous section, any choice of $a_0,...,a_{\alpha},b_0,...,b_{\alpha},d_0,...,d_{\alpha}$ will yield a rigorous lower bound,

$$c \geq \frac{1}{||e_{I_{\alpha}}^{\top}\tilde{A}^{-1}||_1}.$$

5. Computer Assisted Proofs. Now we will validate our conjecture for various special cases with computer assisted proofs using the inequalities we proved in section 4 [Theorem 4.2]. Let T be a α-order symmetric tensor with size d. We will consider the case where we have two distinct indices where we denote an entry of T as $T_{s^i t^j}$ with $s, t \in \{1, \dots, d\}$, i is the number of s terms, j is the number of t terms, and $i + j = \alpha$.

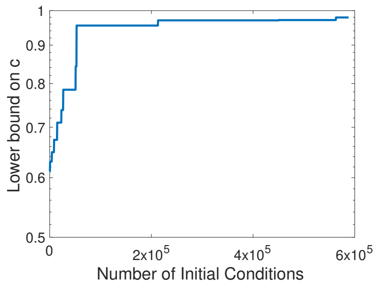
	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
$\alpha=3$	1	1					
$\alpha=4$	1	1	1				
$\alpha=5$	1	1	1				
$\alpha=6$	1	1	1	0.9794			
$\alpha=7$	1	1	1	0.9891			
$\alpha=8$	0.9980	0.9960	0.9891	0.7252	0.6414		
$\alpha=9$	0.9625	0.9970	0.9285	0.5896	0.7704		
$\alpha=10$	0.8460	0.8547	0.9579	0.6219	0.5851	0.3095	
$\alpha=11$	0.8058	0.8628	0.7386	0.9625	0.3979	0.1771	
$\alpha=12$	0.7547	0.4103	0.7474	0.3754	0.2181	0.2981	0.1016

(a)

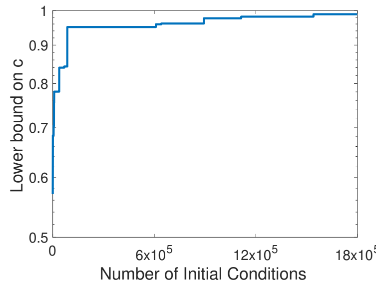
	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
$\alpha=3$	1	1					
$\alpha=4$	1	1	1				
$\alpha=5$	1	1	1				
$\alpha=6$	1	1	1	1			
$\alpha=7$	1	1	1	1			
$\alpha=8$	1	1	1	0.9881	0.9950		
$\alpha=9$	0.9980	1	1	0.9911	0.9862		
$\alpha=10$	0.9930	0.9980 0.9930	0.9747	0.9662	0.7924		
$\alpha=11$	0.9980	0.9775	0.9921	0.6579	0.9009	0.7067	
$\alpha=12$	0.9911	0.9434	0.8993	0.7305	0.6859	0.5456	0.4337

(b)

Fig. 2: (a) num=50, maxiter=1000 (b) num=500, maxiter=2000



(a)



(b)

Fig. 3: (a) By running 6×10^5 initial conditions we were able to show that c_{opt} is at least 0.9793 for the 3-tensor ($\alpha = 3$) in the T_{u^3} entry. (b) Similarly running 18×10^5 initial conditions obtains a lower bound of 0.9891 for the 3-tensor in the T_{u^2t} entry.

6. Sharpness. Now we wish to show the sharpness of our conjecture, i.e. demonstrate that the inequality (1.3) is optimal. We first show that without loss of generality we only need to show sharpness for 2^α tensors.

LEMMA 6.1. *If there is an example of a symmetric tensor $T \in \mathbb{R}^{2^\alpha}$ such that the inequality is shown to be sharp, then the inequality is also sharp for $\mathbb{R}^{n^\alpha}$.*

Proof. Suppose there is a symmetric tensor $T \in \mathbb{R}^{2^\alpha}$ that demonstrates sharpness of the conjecture. We can extend this to the n^α case where entries with indices that consist of ones and twos are the same as the 2^α case and all other entries 0. This yields the following equations:

$$
\begin{aligned}
\lambda u_1 &= (T \times_2 u \times_3 u)_1 = \sum_{j,k=1}^{n} T_{1jk\ell} u_j u_k u_\ell \\
\lambda u_2 &= (T \times_2 u \times_3 u)_2 = \sum_{j,k=1}^{n} T_{2jk\ell} u_j u_k u_\ell \\
\lambda u_3 &= (T \times_2 u \times_3 u)_3 = \sum_{j,k=1}^{n} T_{3jk\ell} u_j u_k u_\ell = 0 \\
\lambda u_4 &= (T \times_2 u \times_3 u)_4 = \sum_{j,k=1}^{n} T_{4jk\ell} u_j u_k u_\ell = 0 \\
&\vdots \\
\lambda u_n &= (T \times_2 u \times_3 u)_3 = \sum_{j,k=1}^{n} T_{njk\ell} u_j u_k u_\ell = 0 \\
1 &= u_1^1 + u_2^2 + \ldots + u_n^2
\end{aligned}
$$

Obviously $u_3 = u_4 = \ldots = u_n = 0$ and thus $u_1^2 + u_2^2 = 1$. Hence we have the same equations as in the 2^α case. Therefore, we have proved the inequality for any symmetric α-tensor. ∎

We first consider the case of 3-tensors. By the eigenvalue equation, for some 3-tensor $T \in \mathbb{R}^{d^3}$ and eigenvector $\vec{u} \in \mathbb{R}^d$ of length 1 with associated eigenvalue λ we have

$$T \times_2 u \times_3 u = \lambda u \tag{6.1}$$

with the component-wise definition

$$(T \times_2 u \times_3 u)_i = \sum_{j,k=1}^{d} T_{ijk} u_j u_k.$$

By Lemma 6.1, we only need to consider tensors that are in $\mathbb{R}^{2^3}$. Consider the $2\times 2\times 2$ symmetric tensor T with entries $T_{111} = 0, T_{211} = 1, T_{221} = 0, T_{222} = -1$ so

$$T = \begin{bmatrix} T_{111} & T_{121} \\ T_{211} & T_{221} \end{bmatrix} \begin{bmatrix} T_{112} & T_{122} \\ T_{212} & T_{222} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and eigenvector of length 1 $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$. By plugging these values into the eigenvalue equation and taking the first component of both sides of (6.1) we have the following equality

$$\lambda u_1 = (T \times_2 u \times_3 u)_1 = \sum_{j,k=1}^{2} T_{1jk} u_j u_k = T_{111} u_1^2 + T_{112} u_1 u_2 + T_{121} u_2 u_1 + T_{122} u_2^2 = 2u_1 u_2$$

Now, taking the second component we have

$$\begin{aligned} \lambda u_2 = (T \times_2 u \times_3 u)_2 = \sum_{j,k=1}^{2} T_{2jk} u_j u_k = T_{211} u_1^2 + T_{212} u_1 u_2 + T_{221} u_2 u_1 + T_{222} u_2^2 \\ = u_1^2 - u_2^2 \end{aligned}$$

Therefore, we have the system of equations

$$\lambda u_1 = 2u_1 u_2 \tag{6.2}$$

$$\lambda u_2 = u_1^2 - u_2^2 \tag{6.3}$$

$$u_1^2 + u_2^2 = 1 \tag{6.4}$$

Rearranging (6.2) gives $\lambda = 2u_2$. Plugging this value of λ into (6.3) then gives $u_1^2 = 3u_2^2$. Plugging this into (6.4) and solving the equation yields $\lambda = \pm 1$.

Therefore, since $\lambda_{\text{maxabs}} = 1$ and is greater than or equal to the absolute value of each entry in the tensor, this proves that in the inequality $\lambda_{\text{maxabs}} \geq c \cdot T_{stu}$, c cannot have a value greater than 1, proving the sharpness of our inequality for the 3-tensor.

We next consider the case of 4-tensors. By Lemma 6.1, we only need to consider tensors that are in $\mathbb{R}^{2^4}$. Note that by the eigenvalue equation, for some $T \in \mathbb{R}^{d^4}$ and eigenvector $\vec{u} \in \mathbb{R}^d$ of length 1 with associated eigenvalue λ we have

$$T \times_2 u \times_3 u \times_4 u = \lambda u \tag{6.5}$$

with the component-wise definition

$$(T \times_2 u \times_3 \times_4 u)_i = \sum_{j,k,\ell=1}^{d} T_{ijk} u_j u_k u_\ell.$$

Now, consider the $2 \times 2 \times 2 \times 2$ symmetric tensor T with entries $T_{1111} = T_{2222} = 3$, $T_{1122} = 1$, and $T_{1112} = T_{1222} = 0$. We then take the first component of both sides of (6.5) to get the following equality

$$\begin{aligned}
\lambda u_1 &= (T \times_2 u \times_3 u \times_4 u)_1 \\
&= \sum_{j,k=1}^{2} T_{1jk\ell} u_j u_k u_\ell \\
&= T_{1111} u_1^3 + (T_{1211} + T_{1121} + T_{1112}) u_1^2 u_2 + (T_{1221} + T_{1212} + T_{1122}) u_1 u_2^2 + T_{1222} u_2^3 \\
&= 3u_1^3 + 3u_1 u_2^2 \\
&= 3(u_1^3 + u_1 u_2^2).
\end{aligned}$$

Now, taking the second component we have

$$\begin{aligned}
\lambda u_2 &= (T \times_2 u \times_3 u \times_4 u)_2 \\
&= \sum_{j,k=1}^{2} T_{2jk\ell} u_j u_k u_\ell \\
&= T_{2111} u_1^3 + (T_{2211} + T_{2112} + T_{2121}) u_1^2 u_2 + (T_{2212} + T_{2221} + T_{2122}) u_1 u_2^2 + T_{2222} u_2^3 \\
&= 3u_1^2 u_2 + 3u_2^3 \\
&= 3(u_1^2 u_2 + u_2^3).
\end{aligned}$$

Therefore, we have the system of equations

$$\lambda u_1 = 3(u_1^3 + u_1 u_2^2) \tag{6.6}$$

$$\lambda u_2 = 3(u_1^2 u_2 + u_2^3) \tag{6.7}$$

$$u_1^2 + u_2^2 = 1 \tag{6.8}$$

Rearranging (6.8) as $u_2^2 = 1 - u_1^2$ and plugging this into (6.6) gives us $\lambda u_1 = 3u_1$ hence $\lambda = 3$. Then we rearrange (6.8) as $u_1^2 = 1 - u_2^2$ and plug this into (6.7) to give us $\lambda u_2 = 3u_2$ thus once again $\lambda = 3$. Therefore $\lambda_{\text{maxabs}} = 3$ and satisfies the conjecture

$$\lambda_{\text{maxabs}} \geq |T_{ijk\ell}|$$

since the largest entry in T is also 3, proving the sharpness of the conjecture for the 4-tensor.

7. Conclusions and Future Directions. The conjecture we proposed would be invaluable in finding approximate CP decompositions for symmetric tensors as it would provide us an improved convergence rate. We were able to numerically find the optimal c in the inequality (1.1) to be approximately 1 which we demonstrate through rigorous computer assisted proofs. Of course, since our computations were only accurate up to numerical precision, in most cases we did not get $c = 1$ exactly. However, the fact that we were often able to get results that approached 1 was very promising. In every situation we looked at, we were always able to get as close to $c = 1$ as we wanted by increasing run times. In addition, due to our analytical proofs in the previous section, we were able to show that the conjecture is true for all symmetric tensors up to order 4.

Another direction would be to further explore the sharpness of our conjecture for symmetric tensors of order $\alpha \geq 5$. We were able to prove sharpness by solving a system of quadratic equations but once we attempted to construct an appropriate example tensor for $\alpha = 5$, the polynomials become higher order and and more complex. Perhaps symbolic algebra methods could be applied to solve these systems.

Appendix A. Proof of Theorem 1.2.

The notion of the rank-1 matrix can be generalized to tensors as follows.

DEFINITION A.1 (Rank-1 Tensor). *Let $T \in \mathbb{R}^{d^\alpha}$ then T is called a* **rank-1 tensor** *if there exists a $v \in \mathbb{R}^d$ such that*

$$v^{\otimes\alpha} = T.$$

Now for non rank-1 tensors, one may seek to decompose such tensors as a sum of rank-1 tensors.

DEFINITION A.2 (CP Decomposition). *The vectors $v_1, ..., v_p$ form a* **CP decomposition** *of α-tensor T if,*

$$T = \sum_{\ell=1}^{p} v_\ell^{\otimes\alpha}$$

where $v_\ell^{\otimes\alpha}$ for $\ell = 1, \dots, p$ are rank-1 tensors and the minimum value of p for which such a decomposition exists is called the rank *of the tensor T.*

Note that this notion of rank agrees with the classical notion of matrix rank in the case of 2-tensors but many of the properties of matrix rank do not generalize to higher order tensors [5, 2, 3, 7, 4].

We define the Frobenius norm generalized to tensors in the following way.

DEFINITION A.3 (Tensor Frobenius Norm [5]). *The* **Frobenius norm of a tensor** *$T \in \mathbb{R}^{d^\alpha}$ is the square root of the sum of the squares of all its elements:*

$$\|T\|_F = \sqrt{\sum_{i_1=1}^{d} \cdots \sum_{i_\alpha=1}^{d} {T_{i_1,\dots,i_\alpha}}^2}.$$

The following lemma introduces a particularly simple formula involving the tensor Frobenius norm and vectors.

LEMMA A.4. *Let $v \in \mathbb{R}^d$ and α be a positive integer. Then, the tensor Frobenius norm of the αth-order tensor product is the same as the Euclidean norm of v raised to the α, i.e.*

$$\|v^{\otimes\alpha}\|_F = \|v\|^\alpha.$$

Proof. By the definition of the tensor Frobenius norm,

$$\|v^{\otimes\alpha}\|_F^2 = \sum_{i_1=1}^{d} \cdots \sum_{i_\alpha=1}^{d} [(v^{\otimes\alpha})_{i_1,\dots,i_\alpha}]^2$$

and since $(v^{\otimes\alpha})_{i_1\dots i_\alpha} = v_{i_1} v_{i_2} \cdots v_{i_\alpha}$, we have $\|v^{\otimes\alpha}\|_F^2 = \sum_{i_1=1}^{d} \cdots \sum_{i_\alpha=1}^{d} v_{i_1}^2 \cdots v_{i_\alpha}^2$, so

$$\|v^{\otimes\alpha}\|_F^2 = \sum_{i_1=1}^{d} v_{i_1}^2 \sum_{i_2=1}^{d} v_{i_2}^2 \cdots \sum_{i_\alpha=1}^{d} v_{i_\alpha}^2 = \underbrace{\|v\|^2 \|v\|^2 \cdots \|v\|^2}_{\alpha \text{ times}}$$

by definition of $\|v\|$ so
$$\|v^{\otimes\alpha}\|_F = \|v\|^\alpha. \qquad \square$$

Finally, we present the following lemma which demonstrates that an eigenvalue-eigenvector pair provides a rank-1 approximation of a tensor in the Frobenius norm.

LEMMA A.5. *Let T be an α-order symmetric tensor with dimension d, i.e. $T \in \mathbb{R}^{d^\alpha}$, and $v \in \mathbb{R}^d$ be a unit length eigenvector of T with eigenvalue $\lambda \neq 0$. Then*
$$\|T - \lambda v^{\otimes\alpha}\|_F^2 = \|T\|_F^2 - \lambda^2$$
and $\|T\|_F \geq \lambda$.

Proof. We first wish to show that $\|T - \lambda v^{\otimes\alpha}\|_F^2 = \|T\|_F^2 - \lambda^2$.
$$\begin{aligned}\|T - \lambda v^{\otimes\alpha}\|_F^2 &= \sum_{i_1=1}^{d} \cdots \sum_{i_\alpha=1}^{d} [(T - \lambda v^{\otimes\alpha})_{i_1,\ldots,i_\alpha}]^2 \\ &= \sum_{i_1=1}^{d} \cdots \sum_{i_\alpha=1}^{d} (T_{i_1,\ldots,i_\alpha}^2 - 2\lambda T_{i_1,\ldots,i_\alpha} v_{i_1} v_{i_2} \cdots v_{i_\alpha} + \lambda^2 v_{i_1}^2 v_{i_2}^2 \cdots v_{i_\alpha}^2) \\ &= \|T\|_F^2 - 2\lambda \sum_{i=1}^{d} v_i (T \times_2 v \times_3 v \times_4 \cdots \times_k v)_i + \lambda^2 \|v^{\otimes\alpha}\|_F\end{aligned}$$
Since $\|v\| = 1$ and by Lemma A.4, $\|v^{\otimes\alpha}\|_F = 1$, hence
$$\|T - \lambda v^{\otimes\alpha}\|_F^2 = \|T\|_F^2 - 2\lambda\langle v, \lambda v\rangle + \lambda^2 = \|T\|_F^2 - 2\lambda^2\|v\|_2^2 + \lambda^2 = \|T\|_F^2 - \lambda^2$$
Since $\|T - \lambda v^{\otimes\alpha}\|_F \geq 0$, $\|T\|_F^2 - \lambda^2 \geq 0$ so $\|T\|_F^2 \geq \lambda^2$ and taking square roots, $\|T\|_F \geq |\lambda|$. $\square$

Proof. First let λ_{maxabs} be the largest eigenvalue in absolute value of a tensor T and assume $\lambda_{\text{maxabs}} \geq |T_{i_1 \ldots i_\alpha}|$ for all $i_1, \ldots, i_\alpha$. We will show that there exists a constant $c = \frac{1}{d^{\alpha/2}} \in (0, 1]$ such that $\lambda_{\text{maxabs}} \geq c\|T\|_F$. Since $\lambda_{\text{maxabs}} \geq |T_{i_1 \ldots i_\alpha}|$, we have
$$\lambda_{\text{maxabs}}^2 \geq T_{i_1 \ldots i_\alpha}^2$$
which implies that
$$d^\alpha \lambda_{\text{maxabs}}^2 \geq \sum_{i_1,\ldots,i_\alpha} T_{i_1 \ldots i_\alpha}^2$$
so we have $d^{\alpha/2}\lambda_{\text{maxabs}} \geq \sqrt{\sum_{i_1,\ldots,i_\alpha} T_{i_1 \ldots i_\alpha}^2}$ and
$$\lambda_{\text{maxabs}} \geq \frac{1}{d^{\alpha/2}}\|T\|_F, \tag{A.1}$$
where we take $c = \dfrac{1}{d^{\alpha/2}} \in (0, 1)$, since $d \geq 1$. By Lemma A.5 applied to T_ℓ, we have
$$\|T_{\ell+1}\|_F{}^2 = \|T_\ell - \lambda_\ell v_\ell^{\otimes\alpha}\|_F{}^2 = \|T_\ell\|_F{}^2 - \lambda_\ell{}^2.$$
Since λ_ℓ is defined to be the largest eigenvalue of T_ℓ, (A.1) says that $\lambda_\ell \geq c\|T_\ell\|_F$ where $c = \frac{1}{d^{\alpha/2}}$ so
$$\begin{aligned}\|T_{\ell+1}\|_F{}^2 &\leq \|T_\ell\|_F{}^2 - c^2\|T_\ell\|_F{}^2 \\ &\leq (1 - c^2)\|T_\ell\|_F{}^2.\end{aligned}$$

Thus, setting $r = \sqrt{1-c^2} \in (0,1)$ we have $\|T_{\ell+1}\|_F \leq r\|T_\ell\|_F$ and $\|T_{\ell+1}\|_F \leq r^2\|T_{\ell-1}\|_F$ and so forth and proceeding inductively we find,

$$\|T_{\ell+1}\|_F \leq r^{\ell+1}\|T_0\|_F = r^{\ell+1}\|T\|_F.$$

Since $0 < r < 1$, $\lim\limits_{\ell\to\infty} r^{\ell+1} = 0$, so $0 \leq \|T_{\ell+1}\|_F \leq r^{\ell+1}\|T\|_F \to 0$ implies $\|T_{\ell+1}\| \to 0$ as $\ell \to \infty$. Since this limit is 0, an upper bound on the rate of convergence of $\|T_\ell\|_F$ is found by considering

$$\frac{\|T_{\ell+1}\|_F}{\|T_\ell\|_F} \leq r = \sqrt{1 - \frac{1}{d^\alpha}}.$$

□

$$\sum_{1\leq i_1\leq m} a_{i_1} = \sum_{i_1=1}^{m} a_{i_1} = a_1 + a_2 + \cdots + a_m$$

$$\sum_{1\leq i_1<i_2\leq m} a_{i_1}a_{i_2} = a_1a_2 + a_1a_3\cdots + a_1a_m + a_2a_3 + \cdots + a_2a_m\cdots a_{m-1}a_m$$

Appendix B. Proof of Theorem 4.1.

Proof. We will first show that A is invertible. We can rewrite A in terms of the Vandermonde matrix as

$$\begin{pmatrix} \frac{1}{(a_0^2+1)^{\alpha/2}} & & & \\ & \frac{1}{(a_1^2+1)^{\alpha/2}} & & \\ & & \ddots & \\ & & & \frac{1}{(a_\alpha^2+1)^{\alpha/2}} \end{pmatrix} \begin{pmatrix} 1 & a_0 & \cdots & a_0^\alpha \\ 1 & a_1 & \cdots & a_1^\alpha \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_\alpha & \cdots & a_\alpha^\alpha \end{pmatrix} \begin{pmatrix} \binom{\alpha}{0} & & & \\ & \binom{\alpha}{1} & & \\ & & \ddots & \\ & & & \binom{\alpha}{\alpha} \end{pmatrix}$$

so that

$$\begin{aligned} \det(A) &= \begin{vmatrix} \frac{1}{(a_0^2+1)^{\alpha/2}} & & & \\ & \frac{1}{(a_1^2+1)^{\alpha/2}} & & \\ & & \ddots & \\ & & & \frac{1}{(a_\alpha^2+1)^{\alpha/2}} \end{vmatrix} \begin{vmatrix} 1 & a_0 & \cdots & a_0^\alpha \\ 1 & a_1 & \cdots & a_1^\alpha \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_\alpha & \cdots & a_\alpha^\alpha \end{vmatrix} \begin{vmatrix} \binom{\alpha}{0} & & & \\ & \binom{\alpha}{1} & & \\ & & \ddots & \\ & & & \binom{\alpha}{\alpha} \end{vmatrix} \\ &= \prod_{i=0}^{\alpha} \frac{1}{(a_i^2+1)^{\alpha/2}} \prod_{0\leq i\leq j\leq\alpha} (a_j - a_i) \prod_{i=0}^{\alpha} \binom{\alpha}{i} \\ &= \frac{\prod_{i=0}^{\alpha} \binom{\alpha}{i} \prod_{0\leq i\leq j\leq\alpha} (a_j - a_i)}{\prod_{i=0}^{\alpha} (a_i^2+1)^{\alpha/2}}. \end{aligned}$$

As long as $a_0, a_1, \ldots, a_\alpha$ are all distinct, the determinant will always be nonzero and thus A is an invertible matrix.

Now we will show how to find the inverse of A by using the following inverse for the Vandermonde matrix:

$$(V^{-1})_{ij} = \frac{(-1)^{i+j} \sum\limits_{0\leq i_1<\cdots<i_{\alpha+1-i}\leq\alpha+1} a_{i_1}a_{i_2}\cdots a_{i_{j-1}}a_{i_{j+1}}\cdots a_{\alpha+1-i}}{\prod\limits_{\ell<j}^{\alpha+1} (a_j - a_\ell)}.$$

The derivation of this inverse can be found in [6]. Let

$$B = \begin{pmatrix} \frac{1}{(a_0^2+1)^{\alpha/2}} & & \\ & \ddots & \\ & & \frac{1}{(a_\alpha^2+1)^{\alpha/2}} \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} \binom{\alpha}{0} & & \\ & \ddots & \\ & & \binom{\alpha}{\alpha} \end{pmatrix}.$$

Then

$$A^{-1} = (BVC)^{-1} = C^{-1}(BV)^{-1} = C^{-1}V^{-1}B^{-1}$$

Since B and C are invertibile diagonal matrices,

$$B^{-1} = \begin{pmatrix} (a_0^2+1)^{\alpha/2} & & \\ & \ddots & \\ & & (a_\alpha^2+1)^{\alpha/2} \end{pmatrix} \quad \text{and} \quad C^{-1} = \begin{pmatrix} \frac{1}{\binom{\alpha}{0}} & & \\ & \ddots & \\ & & \frac{1}{\binom{\alpha}{\alpha}} \end{pmatrix}.$$

Therefore

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}(a_j^2+1)^{\alpha/2} \sum\limits_{0 \le i_1 < \cdots < i_{\alpha+1-i} \le \alpha+1} a_{i_1} a_{i_2} \cdots a_{i_{j-1}} a_{i_{j+1}} \cdots a_{\alpha+1-i}}{\binom{\alpha}{i} \prod\limits_{\ell<j}^{\alpha+1} (a_j - a_\ell)}.$$

∎

REFERENCES

[1] D. C. Easley and T. Berry, *A higher order unscented transform*, SIAM/ASA Journal on Uncertainty Quantification, 9 (2021), pp. 1094–1131.

[2] J. Haastad, *Tensor rank is np-complete*, in International Colloquium on Automata, Languages, and Programming, Springer, 1989, pp. 451–460.

[3] C. J. Hillar and L.-H. Lim, *Most tensor problems are np-hard*, Journal of the ACM (JACM), 60 (2013), pp. 1–39.

[4] T. G. Kolda, *A counterexample to the possibility of an extension of the Eckart–Young low-rank approximation theorem for the orthogonal rank tensor decomposition*, SIAM Journal on Matrix Analysis and Applications, 24 (2003), pp. 762–767.

[5] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM review, 51 (2009), pp. 455–500.

[6] E. Rawashdeh, *A simple method for finding the inverse matrix of vandermonde matrix*, MATEMATICKI VESNIK, Serbia, (2018).

[7] A. Stegeman and P. Comon, *Subtracting a best rank-1 approximation may increase tensor rank*, Linear Algebra and its Applications, 433 (2010), pp. 1276–1300.