Recovery of Third Order Tensors via Convex Optimization

Holger Rauhut
RWTH Aachen University
Lehrstuhl C für Mathematik (Analysis)
Pontdriesch 10,
52056 Aachen, Germany
Email: rauhut@mathc.rwth-aachen.de

Željka Stojanac
RWTH Aachen University
Lehrstuhl C für Mathematik (Analysis)
Pontdriesch 10,
52056 Aachen, Germany
Email: stojanac@mathc.rwth-aachen.de

***Abstract*—We study recovery of low-rank third order tensors from underdetermined linear measurements. This natural extension of low-rank matrix recovery via nuclear norm minimization is challenging since the tensor nuclear norm is in general intractable to compute. To overcome this obstacle we introduce hierarchical closed convex relaxations of the tensor unit nuclear norm ball based on so-called theta bodies – a recent concept from computational algebraic geometry. Our tensor recovery procedure consists in minimization of the resulting new norms subject to the linear constraints. Numerical results on recovery of third order low-rank tensors show the effectiveness of this new approach.**

I. Introduction and Motivation

The recently introduced theory of compressive sensing enables the recovery of sparse vectors from undersampled measurements via efficient algorithms such as ℓ_1-norm minimization. This concept extends to low-rank matrix recovery where the aim is to reconstruct a matrix $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ of rank at most $r \leq \min\{n_1, n_2\}$ from linear measurements $\mathbf{y} = \mathbf{\Phi}(\mathbf{X})$ where $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^m$ with $m \ll n_1 n_2$. However, the natural approach of finding the solution of the optimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2}} \operatorname{rank}(\mathbf{Z}) \quad \text{s.t.} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{y}, \tag{1}$$

is NP-hard. Nevertheless, it has been shown that under suitable conditions on $\mathbf{\Phi}$ the solution of the convex optimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2}} \|\mathbf{Z}\|_* \quad \text{s.t.} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{y}, \tag{2}$$

reconstructs $\mathbf{X}$ exactly. Here, $\|\cdot\|_*$ denotes the nuclear norm of a matrix, i.e., $\|\mathbf{Z}\|_* = \sum_{i=1}^{\min\{n_1,n_2\}} \sigma_i$, where $\{\sigma_i\}_{i=1}^{\min\{n_1,n_2\}}$ is the set of singular values of a matrix $\mathbf{Z}$. The required number of Gaussian measurements scales like $m \geq Cr\max\{n_1, n_2\}$, see [23], [2].

Here, we consider a further extension of compressive sensing to the recovery of low-rank tensors $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ from a small number of linear measurements $\mathbf{y} = \mathbf{\Phi}(\mathbf{X})$, where $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \to \mathbb{R}^m$ with $m \ll n_1 n_2 \cdots n_d$. Again, we are led to consider the rank-minimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}} \operatorname{rank}(\mathbf{Z}) \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Phi}(\mathbf{Z}). \tag{3}$$

978-1-4673-7353-1/15/$31.00

Different notions of the tensor rank corresponding to different decompositions are available [12]. The CP-rank of an arbitrary tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$, where a rank one tensor is of the form $\mathbf{A} = \mathbf{u}^1 \otimes \mathbf{u}^2 \otimes \cdots \otimes \mathbf{u}^d$ or element-wise $A_{i_1 i_2 \ldots i_d} = u^1_{i_1} u^2_{i_2} \cdots u^d_{i_d}$. Expectedly, the problem (3) is NP hard [14]. Although it is possible to define an analog of the nuclear norm $\|\cdot\|_*$ for tensors and consider the minimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}} \|\mathbf{Z}\|_* \quad \text{s.t.} \quad \mathbf{y} = \mathbf{\Phi}(\mathbf{Z}),$$

the computation of $\|\cdot\|_*$, and thereby this problem, is NP hard [14] as well for tensors of order $d \geq 3$.

To overcome this difficulty, previous approaches to low-rank tensor recovery and tensor completion via convex optimization [6] and [16] focus on the Tucker decomposition and the corresponding norm defined as sum of nuclear norms of the unfoldings (see below for the notion of unfolding). However, it has been shown in [18] that in this scenario the necessary number of measurements for recovery of rank $\mathbf{r}$-tensors via Gaussian measurement ensembles scales exponentially in the dimension, i.e., $m \geq Crn^{d-1}$, where $\mathbf{r} = (r, r, \ldots, r) \in \mathbb{R}^d$. Other approaches, not necessarily based on convex optimization, include iterative hard thresholding algorithms [20], [21], recovery by Riemannian optimization [13], and by the ALS method [10]. However, all these approaches consider the Tucker, the tensor train [19] or in general the hierarchical Tucker decomposition [8]. A recent paper [24] considers tensor completion via tensor nuclear norm minimization and gives theoretical analysis with a significant improvement on the necessary number of measurements for recovery of rank r tensors. However, solving this optimization problem, as already mentioned before, is NP-hard.

As an alternative approach, we build new tensor norms, θ_k-norms, which rely on so-called theta bodies [17], [7]. We treat each entry of a tensor as a polynomial variable. The idea is to define a polynomial ideal J which vanishes on the set $\nu_{\mathbb{R}}(J)$ of all rank-one norm-one third order tensors. This is achieved by taking all minors of order two of every unfolding (to satisfy the rank-one condition) and a polynomial $\sum_{i,j,k} x_{ijk}^2 - 1$ (to satisfy the unit norm condition) as its basis. In this scenario, the convex hull of the set $\nu_{\mathbb{R}}(J)$ will be the unit tensor nuclear norm ball. The unit θ_k-norm balls form a set of hierarchical relaxations of the set $\mathrm{conv}(\nu_{\mathbb{R}}(J))$, that is, of the tensor unit nuclear norm ball. We focus on third order tensors and the largest unit norm ball in this set, the unit θ_1-norm ball. We provide semidefinite programs (SDPs) for computing the θ_1-norm of a given third order tensor and for recovery of low-rank third order tensors via θ_1-norm minimization. The performance of the latter SDP is illustrated with numerical results.

II. Notation

We denote vectors with small bold letters, matrices and tensors with capital bold letters and sets with capital calligraphic letters. The cardinality of a set $\mathcal{S}$ will be denoted by $|\mathcal{S}|$ and with $[m]$ we denote the set $\{1,2,\ldots,m\}$. With $\mathrm{conv}(\mathcal{S})$ we denote the convex hull of the set $\mathcal{S}$.

III. Tensors

We are interested in recovery of low-rank third order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ from underdetermined linear measurements.

We define the Frobenius norm of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ as

$$\|\mathbf{X}\|_F = \sqrt{\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\sum_{i_3=1}^{n_3} X_{i_1 i_2 i_3}^2}.$$

A third order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a rank one tensor if there exist vectors $\mathbf{u}^1 \in \mathbb{R}^{n_1}, \mathbf{u}^2 \in \mathbb{R}^{n_2}, \mathbf{u}^3 \in \mathbb{R}^{n_3}$ such that $\mathbf{X} = \mathbf{u}^1 \otimes \mathbf{u}^2 \otimes \mathbf{u}^3$ or element-wise

$$X_{i_1 i_2 i_3} = u_{i_1}^1 u_{i_2}^2 u_{i_3}^3.$$

The CP-rank (or canonical rank and in the following just rank) of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$. Then the analog of the matrix nuclear norm for tensors is

$$\|\mathbf{X}\|_* = \min\Bigg\{\sum_{k=1}^{r}|c_k| : \mathbf{X} = \sum_{k=1}^{r} c_k \mathbf{u}^{1,k} \otimes \mathbf{u}^{2,k} \otimes \mathbf{u}^{3,k}, \; r \in \mathbb{N}, \left\|\mathbf{u}^{i,k}\right\|_{\ell_2} = 1, \text{ for all } i \in [3], k \in [r]\Bigg\}.$$

However, in the tensor case, computing the canonical rank of a tensor, as well as computing the nuclear norm of a tensor is in general NP-hard, see [9], [15].

The ℓ-th unfolding $\mathbf{X}^{\{\ell\}} \in \mathbb{R}^{n_\ell \times \prod_{k \in [3]\setminus\{\ell\}} n_k}$ of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a matrix defined element-wise as

$$\mathbf{X}^{\{\ell\}}_{i_\ell (i_1 \ldots i_{\ell-1} i_{\ell+1} \ldots i_3)} = \mathbf{X}_{i_1 i_2 i_3}.$$

We often use MATLAB notation. Specifically, for a third order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we denote the second order subtensor in $\mathbb{R}^{n_1 \times n_2}$ obtained by fixing the last index i_3 to k by $\mathbf{X}(:,:,k)$. Vectorization of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ converts a tensor into a column vector $\mathrm{vec}(\mathbf{X}) \in \mathbb{R}^{n_1 n_2 n_3}$. The ordering of the elements in $\mathrm{vec}(\mathbf{X})$ is not important as long as it is consistent.

IV. Theta bodies

Since the tensor nuclear norm is in general NP-hard to compute [14], we propose an alternative approach. We introduce new tensor norms via closed convex relaxations (theta bodies) of the tensor unit nuclear norm. The computation of these norms requires the following definitions and tools from algebraic geometry.

For a non-zero polynomial $f = \sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ in $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, x_2, \ldots, x_n]$ and a monomial order $\succeq$, we define

a) the multidegree of f by $\mathrm{multideg}(f) = \max\left(\boldsymbol{\alpha} \in \mathbb{Z}^n_{\geq 0} : a_{\boldsymbol{\alpha}} \neq 0\right)$,
b) the leading coefficient of f by $\mathrm{LC}(f) = a_{\mathrm{multideg}(f)} \in \mathbb{R}$,
c) the leading monomial of f by $\mathrm{LM}(f) = \mathbf{x}^{\mathrm{multideg}(f)}$,
d) the leading term of f by $\mathrm{LT}(f) = \mathrm{LC}(f)\,\mathrm{LM}(f)$.

In this paper we use the *graded reverse lexicographic ordering* (*grevlex*) ordering, see [4].

Let J be a polynomial ideal in $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, x_2, \ldots, x_n]$. The real algebraic variety of the ideal $\nu_{\mathbb{R}}(J)$ is the set of all points $\mathbf{x} \in \mathbb{R}^n$ where the ideal vanishes, i.e.,

$$\nu_{\mathbb{R}}(J) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0 \text{ for all } f \in J\}.$$

By Hilbert's basis theorem we can assume that the polynomial ideal J is generated by the finite set $\mathcal{F} = \{f_1, f_2, \ldots, f_k\}$ of polynomials in $\mathbb{R}[\mathbf{x}]$. We write

$$J = \langle f_1, f_2, \ldots, f_k \rangle = \left\langle \{f_i\}_{i \in [k]} \right\rangle \quad \text{or simply} \quad J = \langle \mathcal{F} \rangle.$$

Then its real algebraic variety is the set

$$\nu_{\mathbb{R}}(J) = \{\mathbf{x} \in \mathbb{R}^n : f_i(\mathbf{x}) = 0, \text{ for all } i \in [k]\}.$$

Groebner bases are crucial for computations with polynomial ideals.

Definition 1 (Groebner basis): Fix a monomial order. A basis $\mathcal{G} = \{g_1, g_2, \ldots, g_s\}$ of a polynomial ideal $J \subset \mathbb{R}[\mathbf{x}]$ is a *Groebner basis* (or standard basis) if for all $f \in \mathbb{R}[\mathbf{x}]$ there exist **unique** $r \in \mathbb{R}[\mathbf{x}]$ and $g \in J$ such that

$$f = g + r$$

and no monomial of r is divisible by any of the leading monomials in $\mathcal{G}$, i.e., by any of the $\mathrm{LM}(g_1), \mathrm{LM}(g_2), \ldots, \mathrm{LM}(g_s)$. A Groebner basis is not unique, but the reduced version (defined below) is.

Definition 2: Fix a monomial order. The *reduced Groebner basis* for a polynomial ideal $J \in \mathbb{R}[\mathbf{x}]$ is a Groebner basis $\mathcal{G} = \{g_1, g_2, \ldots, g_s\}$ for J such that

1) $\mathrm{LC}(g_i) = 1$, for all $i \in [s]$.
2) $\mathrm{LM}(g_i)$ does not divide $\mathrm{LM}(g_j)$, for all $i \neq j$.

Throughout the paper, $\mathbb{R}[\mathbf{x}]_k$ denotes the set of polynomials of degree at most k. The following definition is central for the definition of theta bodies [17], [7], which will be used below for defining our new tensor norms.

Definition 3 ([7]): Let J be an ideal in $\mathbb{R}[\mathbf{x}]$. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is *k-sos mod J* if there exists a finite set of polynomials $h_1, h_2, \ldots, h_t \in \mathbb{R}[\mathbf{x}]_k$ such that $f \equiv \sum_{j=1}^{t} h_j^2$ mod J, i.e, if $f - \sum_{j=1}^{t} h_j^2 \in J$.

We recall that a degree one polynomial is also known as *linear* polynomial.

Definition 4 (Theta body, [7]): Let $J \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the *k-th theta body of an ideal J* is

$$\mathrm{TH}_k(J) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every linear } f \text{ that is } k\text{-sos mod } J\}.$$

By definition,

$$\mathrm{TH}_1(J) \supseteq \mathrm{TH}_2(J) \supseteq \cdots \supseteq \mathrm{conv}(\nu_{\mathbb{R}}(J)). \quad (4)$$

The theta bodies are closed convex sets, while $\mathrm{conv}(\nu_{\mathbb{R}}(J))$ may not be closed. We say that an ideal $J \subseteq \mathbb{R}[\mathbf{x}]$ is TH_k-*exact* if $\mathrm{TH}_k(J)$ equals to the closure of $\mathrm{conv}(\nu_{\mathbb{R}}(J))$, i.e., to $\overline{\mathrm{conv}(\nu_{\mathbb{R}}(J))}$. Guarantees on convergence can be found in [7]. However, to our knowledge, none of the existing guarantees apply in our case.

Checking whether a given polynomial is k-sos mod J using this definition requires knowledge of all linear polynomials that are k-sos mod J. To overcome this difficulty, we need an alternative description of $\mathrm{TH}_k(J)$.

As in [1] and [7], we assume that there are no linear polynomials in the ideal J and we consider only the monomial bases $\mathcal{B}$ of $\mathbb{R}[\mathbf{x}]/J$. The degree of an equivalence class $f + J$, denoted by $\deg(f + J)$, is the smallest degree of an element in the class. Each element in the basis $\mathcal{B} = \{f_i + J\}$ of $\mathbb{R}[\mathbf{x}]/J$ is represented by the polynomial f_i such that $\deg(f_i + J) = \deg(f_i)$. We assume that $\mathcal{B} = \{f_i + J\}$ is ordered so that $f_{i+1} \succeq_{grevlex} f_i$. We define the set

$$\mathcal{B}_k := \{f + J \in \mathcal{B} : \deg(f + J) \leq k\}.$$

Definition 5 (θ-basis, [7]): Let $J \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. A basis $\mathcal{B} = \{f_0 + J, f_1 + J, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/J$ is called a θ-basis

1) if $\mathcal{B}_1 = \{1 + J, x_1 + J, \ldots, x_n + J\}$;
2) if $\deg(f_i + J), \deg(f_j + J) \leq k$ then $f_i f_j + J$ is in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$.

For computing a θ-basis of $\mathbb{R}[\mathbf{x}]/J$, we first need to compute the reduced Groebner basis $\mathcal{G} = \{g_1, \ldots g_s\}$ of the ideal J with respect to an ordering which first compares the total degree (for example, the grevlex ordering). Then, a set $\mathcal{B} = \{f_0 + J, f_1 + J, \ldots\}$ will be a θ-basis of $\mathbb{R}[\mathbf{x}]/J$ if it contains all $f_i + J$ such that

1) f_i is a monomial
2) f_i is not divisible by any of the monomials in the set $\{\mathrm{LT}(g_i) : i \in [s]\}$.

For a θ-basis $\mathcal{B} = \{f_i + J\}$ of $\mathbb{R}[\mathbf{x}]/J$ we define $[\mathbf{x}]_{\mathcal{B}_k}$ to be the column vector formed by all elements of $\mathcal{B}_k$ in order. Then $[\mathbf{x}]_{\mathcal{B}_k} [\mathbf{x}]_{\mathcal{B}_k}^T$ is a square matrix indexed by $\mathcal{B}_k$ and its (i, j)-entry is equal to $f_i f_j + J$. By hypothesis, the entries of $[\mathbf{x}]_{\mathcal{B}_k} [\mathbf{x}]_{\mathcal{B}_k}^T$ lie in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$. Let $\{\lambda_{i,j}^{\ell}\}$ be the unique set of real numbers such that $f_i f_j + J = \sum_{\ell : f_\ell + J \in \mathcal{B}_{2k}} \lambda_{i,j}^{\ell} (f_\ell + J)$.

Definition 6 (k-th combinatorial moment matrix, [7]): Let $J, \mathcal{B}$ and $\{\lambda_{i,j}^{\ell}\}$ be as above. Let $\mathbf{y}$ be a real vector indexed by $\mathcal{B}_{2k}$ with $y_0 = 1$, where y_0 is the first entry of $\mathbf{y}$, indexed by the basis element $1 + J$. The *k-th combinatorial moment matrix* $\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})$ of J is the real matrix indexed by $\mathcal{B}_k$ whose (i, j)-entry is $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{\ell : f_\ell + J \in \mathcal{B}_{2k}} \lambda_{i,j}^{\ell} y_\ell$.

Finally, the following theorem gives us an alternative description of the theta bodies.

Theorem 1 ([7]): The k-th theta body of J, $\mathrm{TH}_k(J)$, is the closure of

$$Q_{\mathcal{B}_k}(J) = \pi_{\mathbb{R}^n} \left\{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : \mathbf{M}_{\mathcal{B}_k}(\mathbf{y}) \succeq 0, y_0 = 1\right\},$$

where $\pi_{\mathbb{R}^n}$ denotes the projection on the variables $y_1, \ldots, y_n$.

V. The tensor θ_k-norm

Let us now provide the announced hierarchical closed convex relaxations of the tensor unit nuclear norm ball. These lead to tensor norms that have not been considered before - at least up to our knowledge.

Remark 1: A similar, but somewhat easier, approach to the one explained in detail for third order tensors below can be used to define closed convex relaxations of the matrix nuclear unit norm ball. In this scenario, all these relaxations coincide with the original matrix unit nuclear norm ball [22]. In the tensor case we cannot expect to obtain equality of all relaxations to the tensor norm, because computing the latter would then not be NP-hard in general. Still, the equality in the matrix case suggests that these relaxations are useful approximations to the nuclear norm in the tensor case.

First, recall that the set of all minors of order two of a matrix $\mathbf{A}$ is

$$\{\det(\mathbf{A}_{\mathcal{I},\mathcal{J}}) : \mathcal{I} \subset [m], \mathcal{J} \subset [n], |\mathcal{I}| = |\mathcal{J}| = 2\},$$

where $\mathbf{A}_{\mathcal{I},\mathcal{J}} \in \mathbb{R}^{2 \times 2}$ denotes the submatrix of $\mathbf{A}$ obtained by deleting all rows $i \in [m] \backslash \mathcal{I}$ and all columns $j \in [n] \backslash \mathcal{J}$.

For notational purposes, we define the following polynomials in $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_{111}, x_{112}, \ldots, x_{n_1 n_2 n_3}]$

$$\begin{aligned}
f_1^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{x}) &= -x_{ijk} x_{\hat{i}\hat{j}\hat{k}} + x_{i\hat{j}\hat{k}} x_{\hat{i}jk}, \quad ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_1 \\
f_2^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{x}) &= -x_{ijk} x_{\hat{i}\hat{j}\hat{k}} + x_{i\hat{j}k} x_{\hat{i}j\hat{k}}, \quad ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_2 \\
f_3^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{x}) &= -x_{ijk} x_{\hat{i}\hat{j}\hat{k}} + x_{ij\hat{k}} x_{\hat{i}\hat{j}k}, \quad ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_3 \\
g(\mathbf{x}) &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk}^2 - 1
\end{aligned}$$

with the corresponding sets of subscripts

$$\begin{aligned}
\mathcal{S}_1 &= \left\{ijk\hat{i}\hat{j}\hat{k} : i < \hat{i},\, j < \hat{j},\, k \leq \hat{k}\right\}, \\
\mathcal{S}_2 &= \left\{ijk\hat{i}\hat{j}\hat{k} : i \leq \hat{i},\, j < \hat{j},\, k < \hat{k}\right\}, \\
\mathcal{S}_3 &= \left\{ijk\hat{i}\hat{j}\hat{k} : i < \hat{i},\, j \leq \hat{j},\, k < \hat{k}\right\},
\end{aligned}$$

where $i, \hat{i} \in [n_1]$, $j, \hat{j} \in [n_2]$ and $k, \hat{k} \in [n_3]$. These polynomials correspond to the order two minors of the unfoldings of a tensor.

Lemma 1 ([22]): A tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a rank one, Frobenius norm one tensor if and only if

$$g(\mathbf{X}) = 0 \text{ and } f_\ell^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{X}) = 0, \quad \text{for all } ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_\ell,\, \ell \in [3].$$

A third order tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a rank-one tensor if and only if all three unfoldings $\mathbf{X}^{\{1\}} \in \mathbb{R}^{n_1 \times n_2 n_3}$, $\mathbf{X}^{\{2\}} \in \mathbb{R}^{n_2 \times n_1 n_3}$, and $\mathbf{X}^{\{3\}} \in \mathbb{R}^{n_3 \times n_1 n_2}$ are rank-one matrices. Notice that $f_\ell^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{X}) = 0$ for all $ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_\ell$ is equivalent to the statement that the ℓ-th unfolding $\mathbf{X}^{\{\ell\}}$ is a rank one matrix, i.e., that all its minors of order two vanish.

Our aim is to find a relaxation of the tensor unit nuclear norm ball. In order to apply the concept of theta bodies, we need to come up with a polynomial ideal $J_3 \subset \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_{111}, x_{112}, \ldots, x_{n_1 n_2 n_3}]$ such that its algebraic variety is of the form

$$\nu_{\mathbb{R}}(J_3) = \Big\{ \mathbf{x} : g(\mathbf{x}) = 0 \text{ and } f_\ell^{ijk\hat{i}\hat{j}\hat{k}}(\mathbf{x}) = 0, \text{ for all } ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_\ell, \ell \in [3] \Big\}.$$

To this end, we define the polynomial ideal $J_3 = \langle \mathcal{G}_{3rd} \rangle$, where

$$\mathcal{G}_{3rd} = \Big\{ f_\ell^{ijk\hat{i}\hat{j}\hat{k}} : ijk\hat{i}\hat{j}\hat{k} \in \mathcal{S}_\ell, \ell \in [3] \Big\} \cup \{g\}. \tag{5}$$

Theorem 2 ([22]): The basis $\mathcal{G}_{3rd}$ defined in (5) forms the reduced Groebner basis of the ideal $J_3 = \langle \mathcal{G}_{3rd} \rangle$ with respect to the grevlex order.

Based on the moment matrix $\mathbf{M}_{\mathcal{B}_1}(\mathbf{y})$, the θ_1-norm of a tensor $\mathbf{X}$ can be computed via the semidefinite program

$$\min_{t,\mathbf{y}} t \quad \text{s.t.} \quad \mathbf{M}(t, \mathbf{y}, \mathbf{X}) \succeq 0,$$

where

$$\mathbf{M}(t, \mathbf{y}, \mathbf{Z}) = t\mathbf{M}_0 + \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} Z_{ijk} \mathbf{M}_{ijk} + \sum_{i=2}^{9} \sum_{j=1}^{|\mathbf{M}^i|} y_\ell \mathbf{M}_j^i, \tag{6}$$

with $\ell = \sum_{k=3}^{i} \left| \mathbf{M}^{(k-1)} \right| + j$ and with matrices $\mathbf{M}_0$, $\mathbf{M}_{ijk}, \mathbf{M}_j^i \in \mathbb{R}^{(n_1 n_2 n_3 + 1) \times (n_1 n_2 n_3 + 1)}$ as defined in Table II.

We then propose to recover a low-rank tensor $\mathbf{X}$ from underdetermined linear measurements $\mathbf{b} = \mathbf{\Phi}(\mathbf{X})$ via θ_k-minimization, i.e.,

$$\arg \min_{\mathbf{Z}} \|\mathbf{Z}\|_{\theta_k} \quad \text{s.t.} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b},$$

which is equivalent to

$$\arg \min_{t, \mathbf{y}, \mathbf{Z}} \quad \text{s.t.} \quad \mathbf{M}(t, y, \mathbf{Z}) \succeq 0 \text{ and } \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}.$$

Remark 2: As already mentioned before, the above Groebner basis $\mathcal{G}_{3rd}$ can be obtained by taking all minors of order two of all three unfoldings of the tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ (not considering the same minor twice). One might think that the θ_1-norm obtained in this way corresponds to a (weighted) sum of the nuclear norms of the unfoldings, which has already been treated in the papers [6], [11]. That is, there exist $\alpha, \beta, \gamma \in \mathbb{R}$ such that

$$\alpha \cdot \|\mathbf{X}^{\{1\}}\|_* + \beta \cdot \|\mathbf{X}^{\{2\}}\|_* + \gamma \cdot \|\mathbf{X}^{\{3\}}\|_* = \|\mathbf{X}\|_{\theta_1}.$$

The example of a cubic tensor $\mathbf{X} \in \mathbb{R}^{2 \times 2 \times 2}$ presented in Table I shows that this is not the case. From the first and the second tensor in Table I we obtain $\gamma = 0$. Similarly, the first and the third tensor, and the first and fourth tensor give $\beta = 0$ and $\alpha = 0$, respectively. Thus, the θ_1-norm is not a weighted sum of the nuclear norms of the unfoldings. In addition, the last tensor shows that the θ_1-norm is not the maximum of the norms of the unfoldings.

	$\mathbf{X} \in \mathbb{R}^{2\times2\times2}$	$\Vert\mathbf{X}^{\{1\}}\Vert_*$	$\Vert\mathbf{X}^{\{2\}}\Vert_*$	$\Vert\mathbf{X}^{\{3\}}\Vert_*$	$\Vert\mathbf{X}\Vert_{\theta_1}$
1	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	2	2	2	2
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	2	2	$\sqrt{2}$	2
3	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$	2	$\sqrt{2}$	2	2
4	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$\sqrt{2}$	2	2	2
5	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$	$\sqrt{2}+1$	$\sqrt{2}+1$	$\sqrt{2}+1$	3

TABLE I

TENSORS $\mathbf{X} \in \mathbb{R}^{2\times2\times2}$ ARE REPRESENTED IN THE SECOND COLUMN AS $\mathbf{X} = [\mathbf{X}(:,:,1) \,|\, \mathbf{X}(:,:,2)]$. THE THIRD, FOURTH AND FIFTH COLUMN REPRESENT THE NUCLEAR NORMS OF THE FIRST, SECOND AND THE THIRD UNFOLDING OF A TENSOR $\mathbf{X}$, RESPECTIVELY. THE LAST COLUMN CONTAINS THE θ_1-NORM WHICH WAS COMPUTED NUMERICALLY.

Remark 3 (complexity): The positive semidefinite matrix $\mathbf{M}$ used either for low-rank tensor recovery or computing the θ_1-norm of a third order tensor $\mathbf{X} \in \mathbb{R}^{n \times n \times n}$ is of dimension $(1+n^3) \times (1+n^3)$. If $a := n^3$ denotes the total number of entries of a tensor $\mathbf{X}$, then $\mathbf{y}$ is a vector of at most $\frac{a(a+1)}{2} \sim O(a^2)$ variables. Therefore, the semidefinite program for computing the θ_1-norm as well as the semidefinite program for low-rank tensor recovery has polynomial complexity.

We remark that this approach for defining relaxations of nuclear norm can also be extended to general dth order tensors, see [22].

VI. NUMERICAL EXPERIMENTS

In this section we provide recovery results for third order tensors obtained by minimizing the θ_1-norm.

We directly build the matrix $\mathbf{M}$ defined in (6) where matrices $\mathbf{M}_0, \mathbf{M}_{ijk}, \mathbf{M}_j^i$ are listed in Table II. Due to symmetry, only the non-zero entries of the upper triangle part of these matrices are specified. The elements of the θ-basis are given via their representative in the second column. The function $f : \mathbb{Z}^3 \to \mathbb{R}$ is defined as $f(i, j, k) = (i-1)n_2 n_3 + (j-1)n_3 + k + 1$. The last column lists the set $\mathcal{T}_{p_1, p_2, p_3} = \{(i, j, k, (p_i)_{i \in \mathcal{Q}}) : 1 \leq i < p_1 \leq n_1, 1 \leq j < p_2 \leq n_2, 1 \leq k < p_3 \leq n_3\}$ where $\mathcal{Q} := \{i : p_i \neq n_i\}$ and $\hat{n}_i = n_i + 1$, for all $i \in [3]$.

We present the recovery results for third order tensors $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ in Table III. We consider cubic and non cubic tensors of ranks one and two.

For fixed dimensions $n_1 \times n_2 \times n_3$, number of measurements m and fixed rank, we performed 200 simulations.

We say that our algorithm succeeds to recover the original tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ if the element-wise difference between the original tensor $\mathbf{X}_0$ and the tensor $\mathbf{X}^* = \arg\min_{\mathbf{Z}:\mathbf{\Phi}(\mathbf{Z}) = \mathbf{\Phi}(\mathbf{X})} \|\mathbf{Z}\|_{\theta_1}$ is at most 10^{-6}. With $m_{\max}$ we

$\mathbf{M}$	θ-basis	$M_{pq} \cdot (p,q)$	Range of $i, \hat{i}, j, \hat{j}, k, \hat{k}$
$\mathbf{M}_0$	1	$1 \cdot (1,1), 1 \cdot (2,2)$	
$\mathbf{M}_{ijk}$	x_{ijk}	$1 \cdot (1, f(i,j,k))$	$\mathcal{T}_{\hat{n}_1,\hat{n}_2,\hat{n}_3}$
$\mathbf{M}^2_{f_2}$	x^2_{ijk}	$-1 \cdot (2,2)$	
		$1 \cdot (f(i,j,k), f(i,j,k))$	$\mathcal{T}_{\hat{n}_1,\hat{n}_2,\hat{n}_3} \backslash \{(1,1,1)\}$
$\mathbf{M}^3_{f_3}$	$x_{i\hat{j}k} x_{ij\hat{k}}$	$1 \cdot (f(i,j,k), f(i,\hat{j},\hat{k}))$	$\mathcal{T}_{\hat{n}_1,\hat{j},\hat{k}}$
		$1 \cdot (f(i,j,\hat{k}), f(i,\hat{j},k))$	$\mathcal{T}_{\hat{n}_1,\hat{j},\hat{k}}$
$\mathbf{M}^4_{f_4}$	$x_{ijk} x_{\hat{i}\hat{j}\hat{k}}$	$1 \cdot (f(i,j,k), f(\hat{i},\hat{j},\hat{k}))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{k}}$
		$1 \cdot (f(i,\hat{j},k), f(\hat{i},j,\hat{k}))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{k}}$
		$1 \cdot (f(i,\hat{j},\hat{k}), f(\hat{i},j,k))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{k}}$
		$1 \cdot (f(i,j,\hat{k}), f(\hat{i},\hat{j},k))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{k}}$
$\mathbf{M}^5_{f_5}$	$x_{ijk} x_{\hat{i}j\hat{k}}$	$1 \cdot (f(i,j,k), f(\hat{i},j,\hat{k}))$	$\mathcal{T}_{\hat{i},\hat{n}_2,\hat{k}}$
		$1 \cdot (f(i,j,\hat{k}), f(\hat{i},j,k))$	$\mathcal{T}_{\hat{i},\hat{n}_2,\hat{k}}$
$\mathbf{M}^6_{f_6}$	$x_{ijk} x_{\hat{i}\hat{j}k}$	$1 \cdot (f(i,j,k), f(\hat{i},\hat{j},k))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{n}_3}$
		$1 \cdot (f(i,\hat{j},k), f(\hat{i},j,k))$	$\mathcal{T}_{\hat{i},\hat{j},\hat{n}_3}$
$\mathbf{M}^7_{f_7}$	$x_{\hat{i}jk} x_{ijk}$	$1 \cdot (f(i,j,k), f(\hat{i},j,k))$	$\mathcal{T}_{\hat{i},\hat{n}_2,\hat{n}_3}$
$\mathbf{M}^8_{f_8}$	$x_{i\hat{j}k} x_{ijk}$	$1 \cdot (f(i,j,k), f(i,\hat{j},k))$	$\mathcal{T}_{\hat{n}_1,\hat{j},\hat{n}_3}$
$\mathbf{M}^9_{f_9}$	$x_{ij\hat{k}} x_{ijk}$	$1 \cdot (f(i,j,k), f(i,j,\hat{k}))$	$\mathcal{T}_{\hat{n}_1,\hat{n}_2,\hat{k}}$

TABLE II

MATRICES USED IN THE DEFINITION OF $\mathbf{M}$ IN (6)

denote the maximal number of measurements for which we do not recovery any of the 200 generated tensors and $m_{\min}$ denotes the minimal number of measurements for which we recovered all 200 tensors (the success of recovery is $200/200$).

We use linear mappings $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times n_3} \to \mathbb{R}^m$ defined with tensors $\mathbf{\Phi}_k \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ via $[\mathbf{\Phi}(\mathbf{X})](k) = \langle \mathbf{X}, \mathbf{\Phi}_k \rangle$, for $k \in [m]$. We choose the $\mathbf{\Phi}_k$ as stochastically independent tensors with i.i.d. Gaussian $\mathcal{N}\left(0, \frac{1}{m}\right)$ entries. We generate tensors $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ of rank $r = 1$ via their decomposition. If $\mathbf{X} = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ is its CP-decomposition, each entry of the vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is taken independently from the normal distribution $\mathcal{N}(0,1)$. Rank two tensors are obtained by summing two rank one tensors.

The last column in Table III represents the number of independent measurements which are always enough for the recovery of a tensor of an arbitrary rank.

$n_1 \times n_2 \times n_3$	rank	$m_{\max}$	$m_{\min}$	$n_1 n_2 n_3$
$2 \times 2 \times 3$	1	4	12	12
$3 \times 3 \times 3$	1	6	19	27
$3 \times 4 \times 5$	1	11	30	60
$4 \times 4 \times 4$	1	11	32	64
$4 \times 5 \times 6$	1	18	42	120
$5 \times 5 \times 5$	1	18	43	125
$3 \times 4 \times 5$	2	27	56	60
$4 \times 4 \times 4$	2	26	56	64
$4 \times 5 \times 6$	2	41	85	120

TABLE III

NUMERICAL RESULTS FOR LOW-RANK TENSOR RECOVERY IN $\mathbb{R}^{n_1 \times n_2 \times n_3}$

We used MATLAB (R2008b) for these numerical experiments, including SeDuMi_1.3 for solving the semidefinite programs.

ACKNOWLEDGEMENTS

We would like to thank Bernd Sturmfels, Daniel Plaumann and Shmuel Friedland for helpful discussions and useful inputs to this paper. We would also like to thank James Saunderson and Hamza Fawzi for deeper insights into the matrix case scenario. We acknowledge funding by the European Research Council through the Starting Grant StG 258926.

REFERENCES

[1] G. Blekherman, P. A. Parrilo, R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*. SIAM, 2013.

[2] E. J. Candès , Y. Plan, *Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements*. IEEE Transactions on Information Theory, 57(4): 2342–2359, 2009.

[3] V. Chandrasekaran, B. Recht, P. A. Parrilo, A. S. Willsky, *The Convex Geometry of Linear Inverse Problems*. Foundations of Computational Mathematics, 12(6): 805–849, 2012.

[4] D. A. Cox, J. Little, D. O'Shea, *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 1991.

[5] D. A. Cox, J. Little, D. O'Shea, *Using Algebraic geometry*. Springer, Graduate texts in mathematics 185, 2005.

[6] S. Gandy, B. Recht, I. Yamada, *Tensor completion and low-n-rank tensor recovery via convex optimization*. Inverse Problems, 27(2): 19pp, 2011.

[7] J. Gouveia, P. A. Parrilo, R. R. Thomas, *Theta Bodies for Polynomial Ideals*. SIAM Journal on Optimization, 20(4): 2097–2118, 2010.

[8] W. Hackbusch, S. Kuhn, *A new scheme for the tensor representation*. J. Fourier Anal. Appl., 15(5): 706–722, 2009.

[9] J. Håstad, *Tensor rank is NP-complete*. Journal of Algorithms, 11(4): 644–654, 1990.

[10] S. Holtz, T. Rohwedder, R. Schneider, *The alternating linear scheme for tensor optimization in the tensor train format*. SIAM J. Sci. Comput., 34(2): A683–A713, 2012.

[11] B. Huang, C. Mu, D. Goldfarb, J. Wright, *Provable Low-Rank Tensor Recovery*. To appear in Pacific Journal of Optimization, 2015.

[12] T. Kolda, B. W. Bader, *Tensor Decompositions and Applications*. SIAM Rev., 51(3): 455–500, 2009.

[13] D. Kressner, M. Steinlechner, B. Vandereycken, *Low-rank tensor completion by Riemannian optimization*. BIT Numerical Mathematics, 54(2): 447–468, 2014.

[14] L.-H. Lim, C. J. Hillar, *Most Tensor Problems are NP-Hard*. Journal of the ACM (JACM), 60(6): 45pp, 2013.

[15] L.-H. Lim, S. Friedland, *Computational complexity of tensor nuclear norm*. arXiv 1410.6072, 2014.

[16] J. Liu, P. Musiaski, P. Wonka, J. Ye, *Tensor Completion for Estimating Missing Values in Visual Data*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1): 208–220, 2013.

[17] L. Lovász. *On the Shannon capacity of a graph*. IEEE Trans. Inform. Theory, 25(1): 1–7, 1979.

[18] C. Mu, B. Huang, J. Wright, D. Goldfarb, *Square deal: lower bounds and improved relaxations for tensor recovery*. In: Proceedings of the International conference on machine learning, 2014.

[19] I. V. Oseledets, *Tensor-train decomposition*. SIAM J. Sci. Comput. 33(5): 2295-2317, 2011.

[20] H. Rauhut, R. Schneider, Ž. Stojanac, *Low-rank tensor recovery via Iterative hard thresholding*. In: Proceedings of the International Conference on Sampling Theory and Applications, 2013.

[21] H. Rauhut, R. Schneider, Ž. Stojanac, *Tensor completion in hierarchical tensor representations*. To be published in Compressed Sensing and Its Applications (edited by H. Boche, R. Calderbank, G. Kutyniok, J. Vybiral), 2015.

[22] H. Rauhut, Ž. Stojanac, *Low-rank tensor recovery via convex optimization*. in preparation, 2015.

[23] B. Recht, M. Fazel, P. A. Parrilo, *Guaranteed minimum-rank solution of linear matrix equations via nuclear norm minimization*. SIAM Rev., 52(3): 471–501, 2010.

[24] M. Yuan, C.-H. Zhang, *On tensor completion via nuclear norm minimization*. arXiv 1405.1773, 2014.