Globally Optimizing Owing to Tensor Decomposition

Arthur Marmin*, Marc Castella†, Jean-Christophe Pesquet*

* *Université Paris-Saclay, CentraleSupélec, Center for Visual Computing, Inria*, Gif-sur-Yvette, France
{arthur.marmin, jean-christophe.pesquet}@centralesupelec.fr
† *SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris,* France
marc.castella@telecom-sudparis.eu

***Abstract*—While global optimization is a challenging topic in the nonconvex setting, a recent approach for optimizing polynomials reformulates the problem as an equivalent problem on measures, which is called a moment problem. It is then relaxed into a convex semidefinite programming problem whose solution gives the first moments of a measure supporting the optimal points. However, extracting the global solutions to the polynomial problem from those moments is still difficult, especially if the latter are poorly estimated. In this paper, we address the issue of extracting optimal points and interpret it as a tensor decomposition problem. By leveraging tools developed for noisy tensor decomposition, we propose a method to find the global solutions to a polynomial optimization problem from a noisy estimation of the solution of its corresponding moment problem. Finally, the interest of tensor decomposition methods for global polynomial optimization is shown through a detailed case study.**

I. Introduction

Tensors are pervasive mathematical tools that have been shown valuable in many scientific areas ranging from engineering or medical imaging to chemistry and quantum physics [1]. At the heart of their success lie the different kinds of tensor decompositions that factorize any tensor into smaller chunks that are interpreted more easily. Among those factorizations, Canonical Polyadic Decomposition (CPD) is of special interest as it decomposes a tensor into a sum of rank-1 tensors [2], [3]. We aim here to broaden the scope of tensor decomposition applications to the challenging area of global polynomial optimization.

For a long time, polynomial functions, and more generally ratios of polynomial functions, have been playing a key role in signal processing. Indeed, their high flexibility as a modelling tool is of great interest when approximating practical quantities which are often intricate. For instance, one can mention their importance in filter design [4], remote-sensing [5], communication networks [6] and process control in pooling [7]. Nevertheless, optimizing a multivariate polynomial function under polynomial constraints is a difficult problem when no convexity property holds [8]. Noticeably, some recent and important mathematical breakthroughs [9]–[12] happened.

More precisely, the moment-based approach [10] transforms the original problem into an equivalent moment problem, whose variable is a measure. The latter is then solved through convex relaxations. However, several practical difficulties remain. First, due to a very demanding computational load, problems are limited in size or require to take into account their specific structure [13], [14]. Another potential difficulty is the extraction of the global solutions to the initial polynomial problem from the solution to the equivalent moment problem. We propose here to tackle this issue through the use of tensors.

In our previous work [15], we used tools from moment problems to perform the CPD of a symmetric tensor. More precisely, the CPD was based on the algebraic method developed in [16], which is very accurate but also sensitive to perturbation. As a consequence, the corresponding CPD method showed strong theoretical guarantees, but was limited to low-rank tensors corrupted with low noise level.

In this work, we adopt a reverse viewpoint: we use CPD algorithms to solve a moment problem, namely, searching for an R-atomic measure supported on a compact set, some of its moments being known. As previously mentioned, this is especially of interest in the context of polynomial (or more generally piecewise rational [17]) optimization while performing the extraction of the global optima. Indeed, minimizing a polynomial function under polynomial constraints is equivalent to an optimization problem on measures. Solving a relaxation of the latter, we obtain some of the moments of the optimal measure, from which we extract the global optima of the initial polynomial. We show here that this extraction is equivalent to performing a CPD on a specific tensor.

In the extraction method used so far [16], if the recovered moments are subject to a small perturbation, inexact results are found. In this work, we propose to use the prolific literature on tensor decomposition, and especially on CPD approximation methods which have been proved to be robust to noise, to perform the extraction of the global minimizers. An additional benefit of using those CPD algorithms is the reduction in dimensions of the relaxation to be solved, which results in a lighter computational load.

Section II sets up the polynomial optimization problem and its connection with the moment problem. Section III presents shortly the reindexing and the CPD of a symmetric tensor before exploring the link between the moment problem and tensor decomposition. Section IV explains how polynomial optimization can benefit from tensor decomposition tools.

Section V illustrates this benefit on a study case. Section VI concludes our work.

We use the following notation: upper case calligraphic letters denote tensors ($\mathcal{T}$) and fraktur letter ($\mathfrak{T}$) their values after reindexing (see Section III). Bold upper case letters ($\mathbf{M}$) denote matrices, bold lower case letters ($\mathbf{v}$) denote vectors, and lower case letters (s) denote scalars. For a multi-index $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$ in $\mathbb{N}^n$ of length n, we define its absolute value $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n$ and we denote by $\mathbb{N}_k^n$ the set of multi-indices of n elements whose absolute value is smaller than or equal to k.

II. Global polynomial optimization

We consider the following optimization problem:

$$\underset{\mathbf{x} \in \mathcal{K}}{\text{minimize}} \;\; p(\mathbf{x}) \,, \tag{1}$$

where p is a polynomial function and the feasible set $\mathcal{K}$ is a compact subset of $\mathbb{R}^n$ defined by a finite number of polynomial inequalities. More specifically, we want to retrieve the global minimizers of (1) by getting rid of spurious local ones, which makes the problem difficult. Let us emphasize that a similar issue is encountered in more general situations such as rational optimization or piecewise rational functions [17]. Addressing these cases is however out of the scope of this paper.

Problem (1) is equivalent to a linear optimization problem on the infinite dimensional space of probability measures [10],

$$\begin{aligned} \underset{\mu \in \mathcal{M}_+(\mathcal{K})}{\text{minimize}} \;\; & \int_{\mathcal{K}} p(\mathbf{x}) \mu(\mathrm{d}\mathbf{x}) \\ \text{s.t.} \; & \int_{\mathcal{K}} \mu(\mathrm{d}\mathbf{x}) = 1 \,, \end{aligned} \tag{2}$$

where $\mathcal{M}_+(\mathcal{K})$ denotes the set of positive finite measures supported on $\mathcal{K}$. In the following, we assume that (1) admits a finite number R of global minimizers. It follows that a solution μ_* to (2) is an atomic measure concentrated on a subset of the global minimizers of (1). Here, we especially look for an optimal measure μ_* supported on the R global optima of (1) and consequently we assume that μ_* is R-atomic. This choice is justified in Section IV-B.

Using recent techniques [10], the moments of μ_* up to a given even degree d are estimated. Indeed, Problem (2) is reformulated as an infinite-dimensional problem on the moments of μ_*. For numerical tractability, the number of considered moments is truncated up to a given degree $d = 2k$ which yields a Semi-Definite Program (SDP) of finite dimensions. The integer k is the index for successive SDP problems of increasing size, which are known as Lasserre's hierarchy [10]. The optimal solution to the order k SDP problem is a vector $\mathbf{y}^*(k) = (y^*_{\boldsymbol{\alpha}}(k))_{\boldsymbol{\alpha} \in \mathbb{N}_d^n}$ corresponding to all the moments up to degree d. Note that the hierarchy is indexed by k instead of $d = 2k$ because of the moment matrix that intervenes in the semidefinite constraint of the SDP problem: as detailed in (10), this matrix contains all the moments up to degree d and thus has its rows and columns indexed by moments up to degree k. Ideally, $\mathbf{y}^*(k)$ contains the moments of the measure μ_* solution to (2) and the extraction step consists then in recovering the support points of this atomic measure from the vector $\mathbf{y}^*(k)$. These support points are the global optimal solutions to (1).

In the next section, we interpret $\mathbf{y}^*(k)$ as a symmetric tensor and use the tools developed for tensor decomposition to perform the extraction phase.

III. Moment problem as a tensor decomposition

A. Symmetric tensor and reindexing

Let $\mathcal{T}$ denote a symmetric tensor of order d on $\mathbb{R}^{n+1}$ with $d \geq 4$ an even integer. By symmetry, the entries $(\mathcal{T}_{i_1,\ldots,i_d})_{0 \leq i_1,\ldots,i_d \leq n}$ of $\mathcal{T}$ are unchanged by any permutation of the indices. Therefore, those entries are uniquely defined by specifying the number of times each index value appears in $\mathbf{i}$. More precisely, to any d-tuple $\mathbf{i} = (i_1, \ldots, i_d)$, we associate an n-tuple $\boldsymbol{\alpha}(\mathbf{i}) = (\alpha_1(\mathbf{i}), \ldots, \alpha_n(\mathbf{i}))$ of $\mathbb{N}_d^n$, where for each l in $[\![1, n]\!]$, $\alpha_l(\mathbf{i})$ is the number of times the index value l appears in $\mathbf{i}$. Note that, since the order of the tensor is d, the number of times the index 0 appear is uniquely defined by $\boldsymbol{\alpha}$ through $d - \sum_{l=1}^n \alpha_l(\mathbf{i})$. We then index our tensor with $\boldsymbol{\alpha}$ in $\mathbb{N}_d^n$ instead of $\mathbf{i}$ and define the tensor values

$$\mathcal{T}_{\mathbf{i}} = \mathfrak{T}_{\boldsymbol{\alpha}(\mathbf{i})} \,. \tag{3}$$

In the following, we will drop the index $\mathbf{i}$ in $\boldsymbol{\alpha}$ and simply write $\mathfrak{T}_{\boldsymbol{\alpha}}$.

B. Canonical Polyadic Decomposition

A tensor is said to be symmetric rank-1 if it can be expressed as

$$\mathbf{v}^{\otimes d} = \underbrace{\mathbf{v} \otimes \cdots \otimes \mathbf{v}}_{d \text{ times}}$$

for a vector $\mathbf{v} = (v_i)_{i \in [\![0,n]\!]}$ of $\mathbb{R}^{n+1}$, that is $\left[\mathbf{v}^{\otimes d}\right]_{i_1,\ldots,i_d} = v_{i_1} \ldots v_{i_d}$. The CPD problem consists then in finding a decomposition of $\mathcal{T}$ into a sum of rank-1 tensors, $\mathcal{T} = \sum_{r=1}^R \mathbf{v}(r)^{\otimes d}$, or equivalently

$$\mathcal{T}_{i_1,\ldots,i_d} = \sum_{r=1}^R v_{i_1}(r) \ldots v_{i_d}(r) \,. \tag{4}$$

The symmetric rank is the minimum number of terms required in any representation of $\mathcal{T}$ as above. Classical steps to determine a CPD of $\mathcal{T}$ consist of first detecting its rank R and then looking for an approximation of $\mathcal{T}$ as a tensor of rank R, that is determining the vectors $(\mathbf{v}(r))_{r \in [\![1,R]\!]}$. Notice that there are ambiguities in defining the vectors of Decomposition (4) as the order of the vectors in the sum is arbitrary.

Furthermore, if we assume that

$$(\forall r \in [\![1, R]\!]) \quad v_0(r) \neq 0 \,,$$

we can normalize each vector $\mathbf{v}(r)$ along its first coordinate and Decomposition (4) reads

$$\mathcal{T} = \sum_{r=1}^R \lambda_r \left(\frac{\mathbf{v}(r)}{v_0(r)}\right)^{\otimes d} = \sum_{r=1}^R \lambda_r \mathbf{u}(r)^{\otimes d} \,, \tag{5}$$

where $\mathbf{u}(r) = \left(\frac{v_1(r)}{v_0(r)}, \dots, \frac{v_n(r)}{v_0(r)}\right)$ is the dehomogenization of $\mathbf{v}(r)$ and $\lambda_r = v_0(r)^d$ is a positive coefficient. Performing the reindexing in (3), Decomposition (5) reads

$$\mathfrak{T}_{\boldsymbol{\alpha}} = \sum_{r=1}^{R} \lambda_r u_1(r)^{\alpha_1} \dots u_n(r)^{\alpha_n} . \tag{6}$$

C. Extraction step as a Canonical Polyadic Decomposition

We assumed in Section II that the sought measure μ_* is R-atomic, i.e. supported on R points $(\mathbf{u}(r))_{r\in[\![1,R]\!]}$ that are global minimizers of (1)

$$\mu_* = \sum_{r=1}^{R} \lambda_r \delta_{\mathbf{u}(r)} . \tag{7}$$

Moreover, the moments contained in the vector $\mathbf{y}^*(k)$ are expressed as

$$(\forall \boldsymbol{\alpha} \in \mathbb{N}_d^n) \quad y^*_{\boldsymbol{\alpha}}(k) = \int_{\mathcal{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mu_*(\mathrm{d}\mathbf{x}) . \tag{8}$$

Replacing (7) in the right hand side of (8), we obtain

$$(\forall \boldsymbol{\alpha} \in \mathbb{N}_d^n) \quad y^*_{\boldsymbol{\alpha}}(k) = \sum_{r=1}^{R} \lambda_r \mathbf{u}(r)^{\boldsymbol{\alpha}} ,$$

which is, following (6), exactly the expression of the CPD of a symmetric tensor $\mathcal{T}$ of order d on $\mathbb{R}^{n+1}$ defined through the indexing (3) as

$$(\forall \boldsymbol{\alpha} \in \mathbb{N}_d^n) \quad \mathfrak{T}_{\boldsymbol{\alpha}} = y^*_{\boldsymbol{\alpha}}(k) . \tag{9}$$

Determining the R-atomic measure μ_* solution to (2), and thus the global minimizers of (1), is equivalent to finding the vectors $(\mathbf{u}(r))_{r\in[\![1,R]\!]}$ and the coefficients $(\lambda_r)_{r\in[\![1,R]\!]}$ of the CPD of the symmetric tensor $\mathcal{T}$ built from the moment vector $\mathbf{y}(k)$. Note that the link between tensor decomposition and moment problem can be found in [18], [19] but with a different formalism.

IV. Impact of tensor decomposition on global solutions extraction

A. Convergence of Lasserre's hierarchy

In Section II, we saw that (2) is relaxed into a hierarchy of SDP problems. Solving each SDP problem yields a truncated vector of moments $\mathbf{y}^*(k)$ and also a lower bound $\mathcal{J}_k^*$ on the optimal value of (1). The higher the order k, the tighter the bound $\mathcal{J}_k^*$ and the higher the number of available moments of μ_* in $\mathbf{y}^*(k)$ but the higher also the dimensions of the SDP problem. It has been proved that the sequence $(\mathcal{J}_k^*)_{k\in\mathbb{N}}$ is increasing and converges to the optimal value of (1) [10]. Moreover, this convergence happens at a finite relaxation order generically [20], i.e. for any instance of Problem (1) where the coefficients of the polynomials are drawn from an absolutely continuous probability distribution, there exists almost surely a finite relaxation order k for which the optimal value of the SDP relaxation is equal to the optimal value of (1).

A sufficient condition to detect convergence is given in [10] using the rank of the moment matrix associated to $\mathbf{y}^*(k)$. The moment matrix is a convenient tool to study moments contained in $\mathbf{y}^*(k)$. It is built by arranging those moments such that

$$(\forall (\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathbb{N}_k^n \times \mathbb{N}_k^n) \;\; (\mathbf{M}_k)_{(\boldsymbol{\alpha},\boldsymbol{\beta})} = y^*_{\boldsymbol{\alpha}+\boldsymbol{\beta}}(k) , \tag{10}$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ are multi-indices indexing $\mathbf{M}_k$ following a given monomial ordering, usually the graded lexicographic one. Note that, following [21], knowing whether this rank condition holds is equivalent to knowing the rank of the symmetric tensor built from $\mathbf{y}^*(k)$. Hence, the tensor associated to $\mathbf{y}^*(k)$ has rank R if convergence occurs for order k. However, this condition is only sufficient and thus convergence can happen for a lower relaxation order. In practice, we often compare the lowest criterion value obtained by our method for Problem (1) and the lower bound $\mathcal{J}_k^*$ to determine whether convergence occurs.

B. Benefit of using interior point methods to solve SDP problems

Many state-of-the-art SDP solvers are relying on interior point methods. Those methods have the advantage to return as a solution an interior point, i.e. a point $\mathbf{y}^*(k)$ in the relative interior of the optimum face of the feasible set. It can be proved [8, Lemma 1.4] that the corresponding moment matrix returned by the SDP solver then has maximum rank among all the matrices which are solutions to the SDP problem. This is an important feature as it guarantees [8] that all global minimizers of (1) can be recovered from $\mathbf{y}^*(k)$. Indeed, the rank of the moment matrix, and thus of the tensor $\mathcal{T}$ built from $\mathbf{y}^*(k)$, is equal to the number of global optima of (1). Having a moment matrix of lower rank implies that some of those minimizers will be missed after performing the extraction. The choice of the SDP solver is therefore important for the extraction.

C. Extraction of solutions: issues with the current method

When convergence in the hierarchy is reached, an algebraic extraction method is currently used [16] in order to recover the measure μ_* from the moment vector $\mathbf{y}^*(k)$, and thus the global minimizers of (1). Nevertheless this method has two main drawbacks: it requires the knowledge of enough moments in order to recover μ_*, i.e. a sufficiently high relaxation order, and it is highly sensitive to noise on the moments.

Those disadvantages raise in several applications. Indeed, since the dimensions of SDP problems increase exponentially with the order of relaxation k, solving them with a sufficient relaxation order k is often computationally too heavy for state-of-the-art SDP solvers and one has to settle for the solutions at the first orders of relaxations. As a consequence, we have access only to a limited numbers of moments which are moreover approximations of the true moments of μ_*. In this context, the extraction method in [16] either fails due to the lack of some moments, or extract minimizers far from the global optima of (1) due to the perturbation on the moments.

Therefore, a robust extraction method is required for many practical applications of polynomial optimization in order to retrieve the exact global minimizers.

D. Robust extraction methods

We showed in Section III that the extraction problem is equivalent to finding the CPD of a symmetric tensor. Thereby robust tensor CPD methods can be used to perform the extraction instead of the algorithm in [16]. Especially, many algorithms relying on the minimization of a fit function instead of algebraic tools have been developed to recover faithfully tensors corrupted by noise, even when some elements are missing [22]. Examples of such algorithms include unconstrained nonlinear optimization [23] (OPT), alternating least square (ALS) [3], and nonlinear least square (NLS) [3].

Hence, after solving the SDP relaxations of order k, we build the symmetric tensor $\mathcal{T}$ out of the solutions $\mathbf{y}^*(k)$ of the SDP problem. This tensor can be seen as a noisy sub-tensor of the infinite tensor of rank R containing all the moments of the sought measure μ_*. Therefore, we apply robust CPD algorithms to get a better approximation of the rank R infinite tensor and enhance the quality of the global optimum of the polynomial problem. Hence, using a moment vector from a low order of the hierarchy, we can avoid huge computational burden while obtaining accurate global minimizers of Problem (1). In practice indeed, the computational time becomes longer for SDP relaxations at an order higher than 4 or 5 [10], [14], [24]. This highlights the interest of inferring the solutions from a lower relaxation order.

An important remark is that many CPD algorithms require the prior knowledge of the rank R of the tensor decomposition which implies, in our context, to know the number of solutions to Problem (1). If the number of solutions R is not known, a rank estimation method, such as the one from Tensorlab [25], can be applied on $\mathcal{T}$.

Let us summarize the overall procedure to solve (1):

- Reformulate (1) into the moment problem (2),
- Relax (2) into an SDP problem by replacing the measure μ with its moments and truncating them up to degree $2k$,
- Solve the SDP relaxation with an interior point method,
- Build the tensor $\mathcal{T}$ from $\mathbf{y}^*(k)$ obtained as a solution of the SDP relaxation,
- Perform the CPD of $\mathcal{T}$,
- Deduce the global minima of (1) from the previous CPD.

V. Case Study

Let us take the following simple polynomial optimization problem from [16]

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^3}{\text{minimize}} \quad & -(x_1-1)^2-(x_2-1)^2-(x_3-1)^2 \\ \text{s.t.} \quad & 1-(x_1-1)^2 \geq 0 \\ & 1-(x_1-x_2)^2 \geq 0 \\ & 1-(x_2-3)^2 \geq 0\,. \end{aligned} \tag{11}$$

Problem (11) has eight global minima

$$\begin{array}{lll} x_1^* = (0,0,0) & x_2^* = (2,0,0) & x_3^* = (0,2,0) \\ x_4^* = (0,0,2) & x_5^* = (2,2,0) & x_6^* = (2,0,2) \\ x_7^* = (0,2,2) & x_8^* = (2,2,2)\,, & \end{array}$$

and the optimal value is -3. Notice that it is not a convex problem. We use GloptiPoly [26] to perform the relaxation into SDP problems and to extract solutions using the algebraic method [16]. Those SDP relaxations are solved by the SDP solver SDPT3 [27]. We compare the extraction method from GloptiPoly with the implementation of NLS from Tensorlab [25]. We choose NLS as it gives better results than ALS and OPT.

A. Extraction method from GloptiPoly

We first apply directly Lasserre's framework in GloptiPoly to solve Problem (11). At the relaxation orders $k=1$, $k=2$, and $k=3$, we do not have convergence in Lasserre's hierarchy and the extraction method fails. Indeed, at those orders, there are not enough moments in $\mathbf{y}^*(k)$ and thus the extraction procedure from [16] fails while extracting the multiplication matrices from the moment matrix. More precisely, in [15, Proposition 1], the index $\beta_r + e_i$ may be greater than the dimension of the moment matrix $\mathbf{M}_k$ and thus the multiplication matrix $\mathbf{N}_i$ cannot be extracted from it. The certificate of convergence is obtained for the relaxation order $k=4$ and the algebraic extraction procedure hence retrieves the eight solutions and the optimal value with a precision higher than 10^{-4}.

B. Robust Extraction using NLS

On the other hand, instead of the algebraic method [16], we extract the solutions by applying NLS algorithm on the tensor generated by the moment vector $\mathbf{y}^*(k)$ for $k=1$, $k=2$, and $k=3$. At each order, we retrieve the eight approximate solutions listed in Table I. Furthermore, Table II shows the value of the criterion at optimality for the solutions extracted with NLS and GloptiPoly, respectively. We observe that, by applying NLS on the tensor corresponding to $\mathbf{y}^*(3)$, we retrieve the eight global solutions and the correct optimal value at a precision higher than 10^{-4}. Therefore, there is no need to solve the SDP problem of order $k=4$. Moreover, at the order $k=2$, NLS gives already a good approximation within an accuracy of 10^{-1} that can be enough in several applications. The same conclusion holds for the optimal value: it is clear from Table II that the optimality is not reached in the hierarchy at $k=3$ since the obtained optimal value is -2.94 instead of -3. However, using NLS as an extraction method yields the correct optimal value with an accuracy higher than 10^{-4}. We yet remark that the lower bounds $\mathcal{J}_1^*$, $\mathcal{J}_2^*$, and $\mathcal{J}_3^*$ all equal the optimal value -3.

VI. Conclusion

We have addressed the problem of extracting the global solutions to a polynomial optimization problem from the estimation of the truncated vector of moments obtained through Lasserre's hierarchy. We have proposed an alternative approach to the standard method for performing this extraction. By interpreting the extraction as a tensor decomposition problem, we use robust methods for tensor CPD such as NLS algorithm in order to extract the global solutions

TABLE I
EXTRACTION OF SOLUTION IN POLYNOMIAL OPTIMIZATION USING NLS (ACCURACY OF 10^{-4})

	$k=1$	$k=2$	$k=3$
x_1^*	$\begin{pmatrix}-0.0116\\-0.0110\\-0.0459\end{pmatrix}$	$\begin{pmatrix}0\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\0\\0\end{pmatrix}$
x_2^*	$\begin{pmatrix}1.0835\\-0.1286\\-0.2224\end{pmatrix}$	$\begin{pmatrix}1.9998\\0.0006\\0.0015\end{pmatrix}$	$\begin{pmatrix}2\\0\\0\end{pmatrix}$
x_3^*	$\begin{pmatrix}-0.2187\\1.1305\\-0.2055\end{pmatrix}$	$\begin{pmatrix}0.0002\\1.9994\\0.0015\end{pmatrix}$	$\begin{pmatrix}0\\2\\0\end{pmatrix}$
x_4^*	$\begin{pmatrix}-0.1728\\-0.1758\\0.9500\end{pmatrix}$	$\begin{pmatrix}0.0002\\0.0006\\1.9985\end{pmatrix}$	$\begin{pmatrix}0\\0\\2\end{pmatrix}$
x_5^*	$\begin{pmatrix}1.3728\\1.2453\\-0.3417\end{pmatrix}$	$\begin{pmatrix}2.0006\\2.0017\\-0.0044\end{pmatrix}$	$\begin{pmatrix}2\\2\\0\end{pmatrix}$
x_6^*	$\begin{pmatrix}1.4839\\-0.3589\\1.5066\end{pmatrix}$	$\begin{pmatrix}2.0006\\-0.0017\\2.0044\end{pmatrix}$	$\begin{pmatrix}2\\0\\2\end{pmatrix}$
x_7^*	$\begin{pmatrix}-0.3968\\1.5726\\1.5603\end{pmatrix}$	$\begin{pmatrix}-0.0006\\2.0017\\2.0044\end{pmatrix}$	$\begin{pmatrix}0\\2\\2\end{pmatrix}$
x_8^*	$\begin{pmatrix}4.7932\\5.2755\\4.7767\end{pmatrix}$	$\begin{pmatrix}1.9984\\1.9957\\1.9862\end{pmatrix}$	$\begin{pmatrix}2\\2\\2\end{pmatrix}$

TABLE II
VALUE OF THE CRITERION AT THE EXTRACTED GLOBAL MINIMA (ACCURACY OF 10^{-4})

	$k=1$	$k=2$	$k=3$
Using NLS	-3.1395	-3.0001	-3.0000
	-2.7751	-2.9954	-3.0000
	-2.9556	-2.9954	-3.0000
	-2.7604	-2.9954	-3.0000
	-1.9994	-3.0135	-3.0000
	-2.3375	-3.0135	-3.0000
	-2.5927	-3.0135	-3.0000
	-4.6932	-2.9666	-3.0000
Using GloptiPoly	-1.2762	-2.5297	-2.9401
Lower bound from GloptiPoly	-3	-3	-3

before the convergence in the hierarchy, thus alleviating the computational burden. Finally, a case study shows the benefit of tensor decomposition for the extraction of global solutions to polynomial optimization problems.

REFERENCES

[1] L. Qi, H. Chen, and Y. Chen, *Tensor Eigenvalues and Their Applications*. Springer Singapore, 2018.

[2] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[3] N. D. Sidiropoulos, L. D. Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[4] B. A. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer Netherlands, 2007.

[5] C. V. Tao and Y. Hu, "A comprehensive study of the rational function model for photogrammetric processing," *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 12, pp. 1347–1358, 2001.

[6] M. Chiang, "Nonconvex optimization for communication networks," in *Advances in Mechanics and Mathematics*. Springer US, 2009, pp. 137–196.

[7] A. Marandi, J. Dahl, and E. de Klerk, "A numerical evaluation of the bounded degree sum-of-squares hierarchy of Lasserre, Toh, and Yang on the pooling problem," *Ann. Oper. Res.*, vol. 265, no. 1, pp. 67–92, Feb. 2017.

[8] M. Laurent, "Sums of squares, moment matrices and optimization over polynomials," in *Emerging Applications of Algebraic Geometry*. Springer New York, Sep. 2008, pp. 157–270.

[9] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Math. Programm.*, vol. 96, no. 2, pp. 293–320, May 2003.

[10] J.-B. Lasserre, *Moments, Positive Polynomials and Their Applications*. London, U.K.: Imperial College Press, 2009.

[11] A. Ahmadi and G. Hall, "Sum of squares basis pursuit with linear and second order cone programming," in *Algebraic and Geometric Methods in Discrete Mathematics*. American Mathematical Society, 2017, pp. 27–53.

[12] M. Dressler, S. Iliman, and T. de Wolff, "A positivstellensatz for sums of nonnegative circuit polynomials," *SIAM J. Appl. Algebr. Geom.*, vol. 1, no. 1, pp. 536–555, Jan. 2017.

[13] F. Bugarin, D. Henrion, and J.-B. Lasserre, "Minimizing the sum of many rational functions," *Math. Program. Comput.*, vol. 8, no. 1, pp. 83–111, Aug. 2015.

[14] M. Castella, J.-C. Pesquet, and A. Marmin, "Rational optimization for nonlinear reconstruction with approximate ℓ_0 penalization," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1407–1417, Mar. 2019.

[15] A. Marmin, M. Castella, and J.-C. Pesquet, "A moment-based approach for guaranteed tensor decomposition," in *Proc. Int. Conf. Acoust. Speech Signal Process.* IEEE, May 2020, pp. 3927–3931.

[16] D. Henrion and J.-B. Lasserre, "Detecting global optimality and extracting solutions in GloptiPoly," in *Positive Polynomials in Control*. Springer Berlin Heidelberg, Sep. 2005, vol. 312, pp. 293–310.

[17] A. Marmin, M. Castella, and J.-C. Pesquet, "How to globally solve non-convex optimization problems involving an approximate ℓ_0 penalization," in *Proc. Int. Conf. Acoust. Speech Signal Process.* IEEE, May 2019, pp. 5601–5605.

[18] J. Brachat, P. Comon, B. Mourrain, and E. Tsigaridas, "Symmetric tensor decomposition," *Lin. Algebra Appl.*, vol. 433, no. 11-12, pp. 1851–1872, Dec. 2010.

[19] A. Bernardi, J. Brachat, P. Comon, and B. Mourrain, "General tensor decomposition, moment matrices and applications," *J. Symbolic Computat.*, vol. 52, pp. 51–71, May 2013.

[20] J. Nie, "Optimality conditions and finite convergence of Lasserre's hierarchy," *Math. Programm.*, vol. 146, no. 1-2, pp. 97–121, May 2013.

[21] A. Marmin, M. Castella, and J.-C. Pesquet, "Detecting the rank of a symmetric tensor," in *Proc. European Signal Processing Conference*. IEEE, Sep. 2019, pp. 1–5.

[22] N. Vervliet, O. Debals, L. Sorber, and L. D. Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 71–79, Sep. 2014.

[23] L. Sorber, M. V. Barel, and L. D. Lathauwer, "Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-$(l_r, l_r, 1)$ terms, and a new generalization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 695–720, Jan. 2013.

[24] A. Marmin, M. Castella, J.-C. Pesquet, and L. Duval, "Signal reconstruction from sub-sampled and nonlinearly distorted observations," in *Proc. European Signal Processing Conference*. IEEE, Sep. 2018, pp. 1970–1974.

[25] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer. (2016, Mar.) Tensorlab 3.0. Available online. [Online]. Available: https://www.tensorlab.net

[26] D. Henrion, J.-B. Lasserre, and J. Löfberg, "GloptiPoly 3: moments, optimization and semidefinite programming," *Optim. Methods Softw.*, vol. 24, no. 4-5, pp. 761–779, Oct. 2009.

[27] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3 — a Matlab software package for semidefinite programming, version 1.3," *Optim. Methods Softw.*, vol. 11, no. 1-4, pp. 545–581, Jan. 1999.