



ON THE HARDNESS OF DECIDING THE FINITE CONVERGENCE OF LASSERRE HIERARCHIES

LUIS FELIPE VARGAS

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
(IDSIA USI-SUPSI), Switzerland

(Communicated by Didier Henrion)

ABSTRACT. A polynomial optimization problem (POP) asks for minimizing a polynomial function given a finite set of polynomial constraints (equations and inequalities). This problem is well-known to be hard in general, as it encodes many hard combinatorial problems. The Lasserre hierarchy is a sequence of semidefinite relaxations for solving (POP). Under the standard Archimedean condition, this hierarchy is guaranteed to converge asymptotically to the optimal value of (POP) (Lasserre, 2001) and, moreover, finite convergence holds generically (Nie, 2012). In this paper, we aim to investigate whether there is an efficient algorithmic procedure to decide whether the Lasserre hierarchy of (POP) has finite convergence. We show that unless $P=NP$ there cannot exist such an algorithmic procedure that runs in polynomial time. We show this already for the standard quadratic programs. Our approach relies on characterizing when finite convergence holds for the so-called Motzkin-Straus formulation (and some variations of it) for the stability number of a graph.

1. Introduction. A polynomial optimization problem asks for minimizing a polynomial function subject to a set of polynomial constraints (equations and inequalities). More precisely, given a polynomial f , and two sets of polynomials $\mathbf{g} = \{g_1, \dots, g_m\}$ and $\mathbf{h} = \{h_1, \dots, h_l\}$, the task is to find (or approximate) the infimum of the following problem:

$$f^* = \inf_{x \in K} f(x), \quad (\text{POP})$$

where

$$K = \left\{ x \in \mathbb{R}^n : g_i(x) \geq 0 \text{ for } i = 1, \dots, m \text{ and } h_i(x) = 0 \text{ for } i = 1, \dots, l \right\}. \quad (1)$$

Polynomial optimization captures many hard problems in many areas such as combinatorial optimization, portfolio and energy optimization (see, for example, [13] and references therein). Consequently, solving (POP) is hard in general. A common approach for approximating (and solving) (POP) is given by the so-called

2020 *Mathematics Subject Classification.* Primary: 90C23, 90c20, 68Q17, 11E25.

Key words and phrases. Lasserre relaxation, finite convergence, positive polynomial, sum of squares, polynomial optimization, semidefinite programming, complexity, Motzkin Straus, stability number, standard quadratic program.

The author is supported by [European Union's Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie Actions Grant Agreement No. 813211 (POEMA) and by the Swiss National Science Foundation project n. 200021_207429 / 1 "Ideal Membership Problems and the Bit Complexity of Sum of Squares Proofs"].

Lasserre sum-of-squares hierarchy, which is a sequence of semidefinite programs of increasing size as we now describe.

We first introduce some notation. Throughout the article $\mathbb{N} = \{1, 2, \dots\}$ and \mathbb{R} will denote, respectively, the set of positive integers and real numbers, and for an integer $n \in \mathbb{N}$, we set $[n] = \{1, 2, \dots, n\}$. We denote by $\mathbb{R}[x] := \mathbb{R}[x_1, x_2, \dots, x_n]$ the ring of polynomials with real coefficients in the variables x_1, \dots, x_n . For $d \in \mathbb{N}$, $\mathbb{R}[x]_d$ denotes the set of polynomials with degree at most d . A polynomial p is a *sum of squares* if we have $p = \sum_{i=1}^m q_i^2$ for some other polynomials q_1, \dots, q_m . The cone of sums of squares is denoted by Σ . For an integer $r \in \mathbb{N}$, we let $\Sigma_r = \Sigma \cap \mathbb{R}[x]_r$, and we define the r -th *truncated quadratic module* generated by the set of polynomials $\mathbf{g} = \{g_1, \dots, g_m\}$ as

$$\mathcal{M}(\mathbf{g})_r := \left\{ \sum_{i=0}^r \sigma_i g_i : \sigma_i \in \Sigma, \deg(\sigma_i g_i) \leq r \text{ for } i = 0, 1, \dots, m, \text{ and } g_0 = 1 \right\}.$$

The union

$$\mathcal{M}(\mathbf{g}) := \bigcup_{r=0}^{\infty} \mathcal{M}(\mathbf{g})_r$$

is called the *quadratic module* generated by \mathbf{g} . Given a set of polynomials $\mathbf{h} = \{h_1, \dots, h_l\}$, the r -th *truncated ideal* generated by \mathbf{h} is defined as

$$I(\mathbf{h})_r := \left\{ \sum_{i=1}^l q_i h_i : q_i \in \mathbb{R}[x], \deg(q_i h_i) \leq r \text{ for } i = 1, 2, \dots, l \right\}.$$

The ideal generated by \mathbf{h} is the union

$$I(\mathbf{h}) = \bigcup_{r=0}^{\infty} I(\mathbf{h})_r.$$

The *Lasserre sum-of-squares hierarchy* of problem (POP) [12] is defined as the following sequence of relaxations: for $r \in \mathbb{N}$,

$$f^{(r)} = \sup\{\lambda : f - \lambda \in \mathcal{M}(\mathbf{g})_r + I(\mathbf{h})_r\}. \quad (2)$$

Clearly, for all $r \geq 1$, we have $f^{(r)} \leq f^{(r+1)} \leq f^*$. We say that the sets of polynomials $\mathbf{g} = \{g_1, \dots, g_m\}$ and $\mathbf{h} = \{h_1, \dots, h_l\}$ satisfy the *Archimedean condition* if we have

$$N - \sum_{i=1}^n x_i^2 \in \mathcal{M}(\mathbf{g}) + I(\mathbf{h}) \quad \text{for some } N \in \mathbb{N}. \quad (3)$$

If the Archimedean condition is satisfied, then the hierarchy $f^{(r)}$ converges asymptotically to f^* as $r \rightarrow \infty$. This is a consequence of the celebrated *Putinar's Positivstellensatz* [23].

We say that the Lasserre hierarchy $f^{(r)}$ has *finite convergence* if

$$f^{(r)} = f^* \text{ for some } r \in \mathbb{N}.$$

Finite convergence does not hold in general (see, e.g., Example 6.19 in [15]). However, it has been shown that finite convergence holds under some mild conditions. We recall some known results from the literature. Laurent [14] showed that if the set of common *complex* zeros of h_1, \dots, h_l is finite (i.e., if the set $\{x \in \mathbb{C}^n : h_1(x) = \dots = h_l(x) = 0\}$ is finite), then the Lasserre hierarchy $f^{(r)}$ has finite convergence. Later, Nie [22] extended this result to the case when the set of common *real* zeros is finite. That is, if $\{x \in \mathbb{R}^n : h_1(x) = \dots = h_l(x) = 0\}$

is finite, then the Lasserre hierarchy $f^{(r)}$ has finite convergence. In addition, Nie [21] showed that under the Archimedean condition the Lasserre hierarchy has finite convergence *generically* (here, *generically* means that the property holds in the entire space of input data except on a set of Lebesgue measure zero). It has also been observed by Henrion and Lasserre [6, 7] that finite convergence holds in many numerical experiments.

The above results raise the question of whether it is possible to fully identify the cases in which finite convergence holds (or fails). More specifically, (given f , \mathbf{g} , and \mathbf{h}) does there exist an algorithmic procedure for deciding if the Lasserre hierarchy of problem (POP) has finite convergence. We show that such a procedure cannot run in polynomial time unless $P=NP$. This is the main result of this paper.

Theorem 1.1. *The problem of deciding whether the Lasserre hierarchy of a polynomial optimization problem has finite convergence is NP-hard.*

We remark that our complexity result holds, in fact, under the Archimedean condition, and more specifically for the class of standard quadratic programs, i.e., when f is a quadratic function and $K = \Delta_n$ is the standard simplex (when $\mathbf{g} = \{x_1, x_2, \dots, x_n\}$, and $\mathbf{h} = \{\sum_{i=1}^n x_i - 1\}$).

Our approach is based on considering variations of the so-called Motzkin-Straus formulation, which is a quadratic optimization problem for computing (the inverse of) the stability number of a graph, as we briefly now describe. Our results also permit us to fully characterize the graphs for which the Lasserre hierarchy of the (original) Motzkin-Straus formulation has finite convergence. This forms the second main result of this article (see Theorem 1.3).

1.1. Strategy of the proof. Given a graph $G = (V = [n], E)$, a subset of vertices $S \subseteq V$ is called *stable* in G if $\{i, j\} \notin E$ for all distinct $i, j \in S$. The *stability number* of G , denoted by $\alpha(G)$, is the maximum cardinality of a stable set in G . Computing $\alpha(G)$ is a well-known NP-hard problem [8]. The starting point of our approach is to consider the following formulation of $\frac{1}{\alpha(G)}$ as the minimization of a quadratic function over the standard simplex

$$\Delta_n = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, x_i \geq 0 \text{ for } i \in [n] \right\}.$$

This formulation is due to Motzkin and Straus [19]. Let $G = ([n], E)$ be a graph. Then we have

$$\frac{1}{\alpha(G)} = \min \left\{ x^T (A_G + I)x : x \in \Delta_n \right\}. \quad (\text{M-S})$$

Here, I is the $n \times n$ identity matrix, and A_G is the adjacency matrix of G (i.e., $(A_G)_{i,j} = 1$ if $\{i, j\} \in E$ and $(A_G)_{i,j} = 0$ if $\{i, j\} \notin E$).

It has been observed that certain variations of problem (M-S) also find $\frac{1}{\alpha(G)}$ (Section 2, see also [16, 5]). In particular, we consider the following reformulation (M-S-e) of $\frac{1}{\alpha(G)}$. Let $e \in E$ be an edge of G , and let $G \setminus e = (V, E \setminus \{e\})$ be the graph obtained by deleting the edge e from the graph G . Then, we have

$$\frac{1}{\alpha(G)} = \min \left\{ x^T (A_G + I + A_{G \setminus e})x : x \in \Delta_n \right\}. \quad (\text{M-S-e})$$

In this paper, we study the finite convergence of the Lasserre hierarchy for problem (M-S) and some variations of it, in particular the ones of the form (M-S-e). The

notion of critical edges plays a crucial role in this analysis, as previously observed in [16]. An edge $e \in E$ in a graph G is *critical* if $\alpha(G \setminus e) = \alpha(G) + 1$. For example, in the *odd cycles* C_{2n+1} (for $n \geq 1$) all edges are critical, while no edge is critical in an *even cycle* C_{2n} (for $n \geq 2$). A pair of vertices (i, j) is called a *twin pair* if $\{i, j\} \in E$ and their sets of neighbours are equal (excluding i and j), i.e., $N_G(i) \setminus \{j\} = N_G(j) \setminus \{i\}$. If (i, j) is a twin pair in G , then contracting the edge $\{i, j\}$ in G produces the same graph as deleting the node i , denoted by $G \setminus i$ (and isomorphic to $G \setminus j$).

Under the assumption that G has no twin pairs, we will characterize when the Lasserre hierarchy of problem (M-S-e) has finite convergence.

Proposition 1.2. *Assume G has no twin pairs and let e be an edge of G . The Lasserre hierarchy of problem (M-S-e) has finite convergence if and only if the edge e is not critical in G .*

The result of Proposition 1.2 combined with the fact that deciding whether an edge is critical in a graph without twin pairs is NP-hard (see Theorem 4.1) gives the main result of our paper (Theorem 1.1). We also characterize when the Lasserre hierarchy of (M-S) has finite convergence. This forms our second main result.

Theorem 1.3. *Let G be a graph. The Lasserre hierarchy of problem (M-S) has finite convergence if and only if the graph obtained after successively contracting all the twin pairs of G does not have critical edges.*

We remark that the “if part” in Proposition 1.2 and Theorem 1.3 essentially follows from the previous work [16]. This relies on using a result by Nie [21] (see Theorem 2.3) claiming that finite convergence holds if the classical optimality conditions hold at every global minimizer. We describe these results in Section 2 for the sake of completeness.

For the “only if” part in Proposition 1.2 and Theorem 1.3, we will exploit the structure of the (infinitely many) minimizers of problems (M-S) and (M-S-e) to obtain information about the polynomials appearing in the sum-of-squares decompositions for the finite convergence. In fact, we show the result for a more general class of problems (see formulation (M-S-B) in Section 2 and see Theorem 3.1 for the more general statement).

1.2. Link to related literature. The complexity status of several decision problems in polynomial optimization has been studied recently. In [1], Ahmadi and Zhang show that given two sets of polynomials $\mathbf{g} = \{g_1, \dots, g_m\}$ and $\mathbf{h} = \{h_1, \dots, h_l\}$ it is NP-hard to test whether the Archimedean property (3) holds. In the same work, they study the problem of deciding whether the optimum of problem (POP) is attained. In particular, they show that this problem is NP-hard already when the objective is linear, and the constraints are quadratic. In addition, Ahmadi and Zhang [2] show that there is no polynomial-time algorithm that decides whether a quadratic program with a bounded feasible set has a unique local minimizer. In [16], the author and Laurent showed that it is NP-hard to decide whether a polynomial optimization problem has finitely many minimizers. In fact, this hardness result holds already for standard quadratic programs. The proof of the last two results also considers variations of the formulation (M-S) by Motzkin and Straus, as we do in the present paper.

The finite convergence of the Lasserre hierarchy of problem (M-S) has been studied recently. Indeed, in [16] it is shown that finite convergence holds for graphs without critical edges (also recalled later in Theorem 2.4). In fact, this result extends for several formulations for the stability number of *weighted* graphs. The following reformulation of problem (M-S) has been also considered:

$$\frac{1}{\alpha(G)} = \min\{(x^{\circ 2})^T(A_G + I)x^{\circ 2} : \sum_{i=1}^n x_i^2 = 1\},$$

where $x^{\circ 2} = (x_1^2, x_2^2, \dots, x_n^2)$. Recently, the author and Schweighofer [26] showed that the Lasserre hierarchy of this problem *always* has finite convergence, for every graph G .

The Lasserre hierarchy of a (POP) has finite convergence if and only if a certain polynomial (namely, $f - f^*$) which is nonnegative (*with zeros*) on the semialgebraic set K belongs to the quadratic module $\mathcal{M}(\mathbf{g}) + I(\mathbf{h})$. The representation of nonnegative polynomials in quadratic modules has been intensively studied in the literature. We highlight the works by Putinar [23], Scheiderer [24, 25], Marshall [17, 18], Kriel and Schweighofer [10, 11], and references therein.

1.3. Organization of the paper. In Section 2 we recall the formulation (M-S) by Motzkin and Straus and some variations of it. Then, we recap some known results about their minimizers and their Lasserre hierarchies. In addition, we recall a result by Nie that permits to show finite convergence of the Lasserre hierarchy under the classical optimality conditions. In Section 3, we prove Theorem 1.3, that is, we characterize the graphs for which the Lasserre hierarchy of problem (M-S) has finite convergence. In fact, we show a more general result (Theorem 3.1) that will be used later for showing our main complexity result. Finally, in Section 4 we show our main result. Namely, we show that the problem of deciding whether the Lasserre hierarchy of a given polynomial optimization has finite convergence is NP-hard.

The main results of this article were included in the PhD thesis [27].

2. Preliminaries.

2.1. Stability number of a graph. We start by recalling the result by Motzkin and Straus that gives the following formulation (M-S) for $\frac{1}{\alpha(G)}$, already mentioned in the Introduction.

Theorem 2.1. *Let $G = ([n], E)$ be a graph, then*

$$\frac{1}{\alpha(G)} = \min \left\{ x^T(A_G + I)x : x \in \Delta_n \right\}. \quad (\text{M-S})$$

Observe that given a graph G and a stable set S of G of size $\alpha(G)$ the vector $x = \frac{1}{\alpha(G)}\chi^S$ is a minimizer of problem (M-S). Here, $\chi^S \in \mathbb{R}^n$ is the characteristic vector of S (defined by $(\chi^S)_i = 1$ for $i \in S$ and $(\chi^S)_i = 0$ for $i \in [n] \setminus S$). In general, problem (M-S) might have infinitely many minimizers as shown in [5] (see also [16] and Theorem 2.2).

In this paper, we consider some variations of problem (M-S). For this, given a graph G we consider the following set of matrices

$$\mathcal{B}(G) = \{B \in \mathcal{S}^n : B_{ii} = 1, B_{ij} \geq 1 \text{ for } \{i, j\} \in E, B_{ij} = 0 \text{ for } \{i, j\} \notin E\}. \quad (4)$$

So, $A_G + I \in \mathcal{B}(G)$. Moreover, for all $B \in \mathcal{B}(G)$ we have $B \geq A_G + I$, where the inequality is understood entry-wise. Then, for all $B \in \mathcal{B}(G)$ the following equality holds as a consequence of Theorem 2.1:

$$\frac{1}{\alpha(G)} = \min \left\{ x^T B x : x \in \Delta_n \right\}. \quad (\text{M-S-B})$$

Indeed, the vectors of the form $x = \frac{1}{\alpha(G)} \chi^S$ (with S stable in G of size $\alpha(G)$) give $x^T B x = \frac{1}{\alpha(G)}$, and for all $x \in \Delta_n$ we have $x^T B x \geq x^T (A_G + I)x$. We now consider the Lasserre hierarchy corresponding to problem (M-S-B).

For convenience, we denote by $\mathcal{M}(x_1, \dots, x_n)$ (resp. $\mathcal{M}(x_1, \dots, x_n)_r$) the quadratic module (resp. the r -th truncated quadratic module) generated by the set of polynomials $\{x_1, x_2, \dots, x_n\}$. Also, we denote by I_{Δ_n} (resp. $I_{\Delta_n, r}$) the ideal (resp. r -th truncated ideal) generated by the polynomial $\sum_{i=1}^n x_i - 1$. The Lasserre relaxation of order r for problem (M-S-B) reads

$$f_{G,B}^{(r)} = \max \{ \lambda : x^T B x - \lambda \in \mathcal{M}(x_1, x_2, \dots, x_n)_r + I_{\Delta_n, r} \}. \quad (5)$$

When $B = A_G + I$, we simply write $f_G^{(r)}$. That is, we define $f_G^{(r)} := f_{G, A_G + I}^{(r)}$, i.e.,

$$f_G^{(r)} = \max \{ \lambda : x^T (A_G + I)x - \lambda \in \mathcal{M}(x_1, x_2, \dots, x_n)_r + I_{\Delta_n, r} \}. \quad (6)$$

It is a well-known fact that the Archimedean condition (3) is satisfied by the sets $\mathbf{g} = \{x_1, \dots, x_n\}$ and $\mathbf{h} = \{\sum_{i=1}^n x_i - 1\}$. We repeat here the easy argument for the sake of completeness. We have the following identities

$$1 - x_i = \left(1 - \sum_{k=1}^n x_k \right) + \sum_{k \in [n] \setminus \{i\}} x_k, \quad (7)$$

$$1 - x_i^2 = \frac{(1 + x_i)^2}{2} (1 - x_i) + \frac{(1 - x_i)^2}{2} (1 + x_i). \quad (8)$$

Using identities (7) and (8), we have that $1 - x_i^2 \in \mathcal{M}(x_1, \dots, x_n) + I_{\Delta_n}$ for all $i \in [n]$. Then, summing up the identities (8) for $i \in [n]$ we obtain

$$\sum_{i=1}^n (1 - x_i^2) = n - \sum_{i=1}^n x_i^2 \in \mathcal{M}(x_1, \dots, x_n) + I_{\Delta_n},$$

as desired.

It is also well-known that the program (5) (and thus (6)) is feasible and attains its maximum, thus the ‘sup’ from the definition of the Lasserre hierarchy can be replaced by a ‘max’. We repeat the argument for completeness as well. If $B \succeq 0$ (i.e., all its eigenvalues are nonnegative), then the polynomial $x^T B x$ is a sum of squares and thus $x^T B x \in \mathcal{M}(x_1, \dots, x_n)$. Otherwise, the matrix $B - \mu I$ is positive semidefinite, where $\mu := \lambda_{\min}(B) < 0$. Then, the polynomial

$$x^T B x - n\mu = x^T (B - \mu I)x - \mu \left(n - \sum_{i=1}^n x_i^2 \right)$$

belongs to $\mathcal{M}(x_1, x_2, \dots, x_n) + I_{\Delta_n}$, showing feasibility of program (5). Finally, the optimum of program (5) is attained because the feasible region is a closed set (see [17]).

It has been shown in [16] that the notion of critical edges plays a role in the analysis of the minimizers of problem (M-S-B) and of the convergence of the hierarchies $f_{G,B}^{(r)}$ and $f_G^{(r)}$. Recall that an edge $e \in E$ is critical if $\alpha(G \setminus e) = \alpha(G) + 1$. For a vector $x \in \mathbb{R}^n$, we denote by $\text{supp}(x) = \{i \in [n] : x_i > 0\}$ its support. We have the following result.

Theorem 2.2 (particular case of Corollary 4.4. in [16]). *Let $G = (V, E)$ be a graph, and let $B \in \mathcal{B}(G)$. Let $x \in \Delta_n$ with support $S = \text{supp}(x)$, and let C_1, C_2, \dots, C_k denote the connected components of the graph $G[S]$. Then, x is a minimizer of problem (M-S-B) if and only if the following holds:*

- (i): $B_{ij} = 1$ for all edges $\{i, j\}$ of $G[S]$,
- (ii): C_h is a clique of G for all $h \in [k]$,
- (iii): $\sum_{i \in C_h} x_i = \frac{1}{\alpha(G)}$ for all $h \in [k]$.

In this case all edges of $G[S]$ are critical in G .

This result is used to study the finite convergence of the hierarchies $f_{G,B}^{(r)}$, as described in the following sections.

2.2. Optimality conditions and finite convergence. In this section, we recall a result of Nie for proving the finite convergence of Lasserre hierarchies when the classical optimality conditions hold. We first do a quick recap of the classical optimality conditions applied to problem (POP) (although it holds for a more general class of problems as described in [3]):

Let u be a local minimizer of problem (POP), and let

$$J(u) = \{j \in [m] : g_j(u) = 0\}$$

be the index set of the active inequality constraints at u . We say that the *constraint qualification condition (CQC)* holds at u if the vectors

$$\{\nabla g_j(u) : j \in J(u)\} \cup \{\nabla h_i(u) : i \in [l]\}$$

are linearly independent. If (CQC) holds at u , then there exist multipliers $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_m \in \mathbb{R}$ such that

$$\begin{aligned} \nabla f(u) &= \sum_{i=1}^l \lambda_i \nabla h_i(u) + \sum_{j=1}^m \mu_j \nabla g_j(u), \\ \mu_1 g_1(u) &= 0, \dots, \mu_m g_m(u) = 0, \mu_1 \geq 0, \mu_m \geq 0. \end{aligned}$$

If it holds that

$$\mu_j > 0 \text{ for every } j \in J(u), \text{ and } \mu_j = 0 \text{ for } j \in [m] \setminus J(u),$$

then we say that the *strict complementarity condition (SCC)* holds at u . We define now the Lagrangian function

$$L(x) = f(x) - \sum_{i=1}^l \lambda_i h_i(x) - \sum_{j \in J(u)} \mu_j g_j(x).$$

A necessary condition for u to be a local minimizer is the following:

$$v^T \nabla^2 L(u) v \geq 0 \text{ for all } v \in G(u)^\perp,$$

where $G(u)$ is defined as the matrix with rows the gradients of the active constraints at u , and $G(u)^\perp$ is its kernel. Hence, we have

$$G(u)^\perp = \{x \in \mathbb{R}^n : x^T \nabla g_j(u) = 0 \text{ for all } j \in J(u) \text{ and}$$

$$x^T \nabla h_i(u) = 0 \text{ for all } i \in [l].$$

If it holds that

$$v^T \nabla^2 L(u) v > 0 \text{ for all } 0 \neq v \in G(u)^\perp,$$

then we say that the *second order sufficiency condition (SOSC)* holds at u .

The relation between the optimality conditions just described and the finite convergence of the Lasserre hierarchy of (POP) is given by the following result of Nie [21].

Theorem 2.3 ([21]). *Assume the Archimedean condition is satisfied by the sets $\mathbf{g} = \{g_1, \dots, g_m\}$ and $\mathbf{h} = \{h_1, \dots, h_l\}$. If the conditions (CQC), (SCC) and (SOSC) hold at every global minimizer of (POP), then the Lasserre hierarchy $f^{(r)}$ has finite convergence.*

2.3. Finite convergence for graphs with no critical edges. As an application of Theorem 2.2 and Theorem 2.3 we obtain the following result. This a particular case of a more general result obtained in the previous work [16]. Specifically, it follows by considering a particular case of Proposition 4.5. and Theorem 5.1. in [16].

Theorem 2.4 ([16]). *Let $G = ([n], E)$ be a graph, and let $B \in \mathcal{B}(G)$. The following three assertions are equivalent:*

- (i): $B_{lm} > 1$ for every critical edge $\{l, m\}$ of G .
- (ii): Problem (M-S-B) has finitely many minimizers.
- (iii): All minimizers of problem (M-S-B) satisfy the optimality conditions (CQC), (SCC), (SOSC).

In addition, any of these three assertions imply the following:

- (iv): The hierarchy $f_{G,B}^{(r)}$ has finite convergence.

In the next section, we will study situations in which assertion (iv) is equivalent to assertions (i)-(ii)-(iii).

3. Characterizing finite convergence. In this section, we aim to characterize when the hierarchy $f_{G,B}^{(r)}$ (and $f_G^{(r)}$) has finite convergence. As a main result, we characterize the graphs for which the hierarchy $f_G^{(r)}$ (of problem (M-S)) has finite convergence (Theorem 1.3). For this, we will show a more general result (Theorem 3.1) that we will use also in Section 4 for showing our main complexity result (Theorem 1.1).

In the absence of twin pairs. The result of Theorem 2.4 gives *sufficient* conditions for the hierarchy $f_{G,B}^{(r)}$ to have finite convergence. We show that when the graph G does not have twin pairs these *sufficient* conditions are in fact *necessary*.

Theorem 3.1. *Let $G = ([n], E)$ be a graph without twin pairs and let $B \in \mathcal{B}(G)$. The Lasserre hierarchy $f_{G,B}^{(r)}$ has finite convergence to $\frac{1}{\alpha(G)}$ if and only if, for every critical edge $\{l, m\}$ of G , we have $B_{lm} > 1$.*

Proof. The ‘if’ part follows from Theorem 2.4. For the ‘only if’ part we proceed by contradiction as follows. Assume there is a critical edge $\{l, m\}$ such that $B_{lm} = 1$.

We assume, moreover, that $f_{G,B}^{(r)}$ has finite convergence, that is, there exist $\sigma, \sigma_i \in \Sigma$ for $i \in [n]$, and $q \in \mathbb{R}[\mathbf{x}]$ such that

$$x^T Bx - \frac{1}{\alpha(G)} = \sigma + \sum_{i=1}^n x_i \sigma_i + q \cdot \left(\sum_{i=1}^n x_i - 1 \right). \quad (9)$$

Since the edge $\{l, m\}$ is critical, we observe that there exists $S \subseteq V$ such that $S \cup \{l\}$ and $S \cup \{m\}$ are stable of size $\alpha(G)$ in G . By Theorem 2.2, for $t \in (0, 1)$, the vector

$$u_t = \frac{1}{\alpha(G)} (t\chi^{S \cup \{l\}} + (1-t)\chi^{S \cup \{m\}})$$

is an optimal solution of problem (M-S-B), i.e., $u_t^T B u_t = \frac{1}{\alpha(G)}$ and $u_t \in \Delta_n$. We evaluate relation (9) at $x + u_t$ and we obtain

$$x^T Bx + 2x^T B u_t = \sigma(x + u_t) + \sum_{i=1}^n (x + u_t)_i \sigma_i(x + u_t) + q(x + u_t) \left(\sum_{i=1}^n x_i \right). \quad (10)$$

Now, we will reach a contradiction by comparing coefficients at both sides in relation (10). First, since there is no constant term on the left-hand side, the constant term on the right-hand side is equal to zero. That is,

$$\sigma(u_t) + \frac{1}{\alpha(G)} \sum_{s \in S} \sigma_s(u_t) + \frac{t}{\alpha(G)} \sigma_l(u_t) + \frac{1-t}{\alpha(G)} \sigma_m(u_t) = 0.$$

Since all terms are nonnegative we have that, for every $t \in (0, 1)$, $\sigma(u_t) = 0$ and $\sigma_i(u_t) = 0$ for $i \in S \cup \{l, m\}$. Then, for every $t \in (0, 1)$, the polynomials $\sigma(x + u_t)$, and $\sigma_i(x + u_t)$ (for $i \in S \cup \{l, m\}$) do not have a constant term and therefore do not have linear terms. Now, we compare the coefficient of x_s , where $s \in S$. On the left hand side of (10), it is equal to

$$2 \sum_{i \in S \cup \{l, m\}} B_{si}(u_t)_i = 2B_{ss}(u_t)_s = \frac{2}{\alpha(G)}.$$

On the right hand side of (10), the polynomials $\sigma(x + u_t)$ and $(x + u_t)_i \sigma_i(x + u_t)$ for $i \in S \cup \{l, m\}$ have no linear term, and for $i \in [n] \setminus (S \cup \{l, m\})$, the polynomials $(x + u_t)_i \sigma_i(x + u_t)$ are divisible by x_i . Hence, the coefficient of x_s is $q(u_t)$. Therefore, $q(u_t) = \frac{2}{\alpha(G)}$. Let $j \in [n]$ be such that $j \in N_G(l)$ and $j \notin N_G(m)$. Here, we use that l and m are not twin nodes (we switch l and m if necessary). We compare the coefficient of x_j at both sides of (10). In the left hand side, the coefficient of x_j is $2 \sum_{i \in S \cup \{l, m\}} B_{ij}(u_t)_i = \frac{2}{\alpha(G)} B_{lj} t + \frac{2}{\alpha(G)} \sum_{i \in S} B_{ij}$. Finally, on the right-hand side, the coefficient of x_j is $\sigma_j(u_t) + q(u_t) = \sigma_j(u_t) + \frac{2}{\alpha(G)}$. Hence, we obtain

$$\sigma_j(u_t) = \frac{2}{\alpha(G)} B_{lj} t + \frac{2}{\alpha(G)} \sum_{i \in S} B_{ij} - \frac{2}{\alpha(G)}.$$

This is a contradiction because $\sigma_j(u_t)$ is a sum of squares of polynomials in t , while the polynomial in the right hand has degree 1 (since $B_{lj} \geq 1$). \square

Finite convergence for the Motzkin-Straus formulation (M-S). Now we show Theorem 1.3. Namely, we characterize the graphs G for which the Lasserre hierarchy $f_G^{(r)}$ (of problem (M-S)) has finite convergence. For this, we first show that the finite convergence of the hierarchy $f_G^{(r)}$ is not affected by contracting twin pairs.

Lemma 3.2. *Let $G = \{[n], E\}$ be a graph such that $(1, 2)$ is a twin pair. The hierarchy $f_G^{(r)}$ has finite convergence if and only if the hierarchy $f_{G \setminus 1}^{(r)}$ has finite convergence.*

Proof. Clearly, we have that $\alpha(G \setminus 1) = \alpha(G)$. We now show that $f_G^{(r)} = \frac{1}{\alpha(G)}$ if and only if $f_{G \setminus 1}^{(r)} = \frac{1}{\alpha(G)}$.

The ‘only if’ part follows directly by replacing $x_1 = 0$ in the definition of $f_G^{(r)}$ (see relation (6)). Assume now that $f_{G \setminus 1}^{(r)} = \frac{1}{\alpha(G)}$. Then, we have

$$x^T(A_{G \setminus 1} + I)x - \frac{1}{\alpha(G)} = \sigma_0 + \sum_{i=2}^n x_i \sigma_i + q \cdot \left(\sum_{i=2}^n x_i - 1 \right), \quad (11)$$

for some $\sigma_0 \in \Sigma_r$, $\sigma_2, \dots, \sigma_n \in \Sigma_{r-1}$, $q \in \mathbb{R}[x]_r$. Now, we set

$$f_G = x^T(A_G + I)x \in \mathbb{R}[x_1, x_2, \dots, x_n], \text{ and}$$

$$f_{G \setminus 1} = x^T(A_{G \setminus 1} + I)x \in \mathbb{R}[x_2, x_3, \dots, x_n].$$

Observe that the following equality holds:

$$f_G(x_1, x_2, \dots) = f_{G \setminus 1}(x_1 + x_2, x_3, \dots, x_n).$$

Then, by replacing x_2 by $x_1 + x_2$ in relation (11) we obtain

$$x^T(A_G + I)x - \frac{1}{\alpha(G)} = \tilde{\sigma}_0 + (x_1 + x_2)\tilde{\sigma}_2 + \sum_{i=3}^n x_i \tilde{\sigma}_i + \tilde{q} \cdot \left(\sum_{i=1}^n x_i - 1 \right),$$

where

$$\tilde{\sigma}_i(x_1, x_2, \dots, x_n) := \sigma(x_1 + x_2, x_3, \dots, x_n), \text{ and}$$

$$\tilde{q}(x_1, x_2, \dots, x_n) := q(x_1 + x_2, x_3, \dots, x_n)$$

Clearly, $\tilde{\sigma}_0 \in \Sigma_r$, $\tilde{\sigma}_2, \dots, \tilde{\sigma}_n \in \Sigma_{r-1}$ and $\tilde{q} \in \mathbb{R}[x]_r$. This shows $p_G^{(r)} = \frac{1}{\alpha(G)}$. \square

Now, we are able to prove Theorem 1.3. We repeat the statement for the sake of clarity.

Theorem 1.3. Let G be a graph. The Lasserre hierarchy of problem (M-S) has finite convergence if and only if the graph obtained after successively contracting all the twin pairs of G does not have critical edges.

Proof. Let \tilde{G} be the graph obtained by successively contracting all the twin pairs from G , so that \tilde{G} does not have twin pairs. By Lemma 3.2, the Lasserre hierarchy of problem (M-S) for the graph G has finite convergence if and only if the Lasserre hierarchy of problem (M-S) for the graph \tilde{G} has finite convergence. By Theorem 3.1 (using $B = I + A_G$), the Lasserre hierarchy of problem (M-S) for the graph \tilde{G} has finite convergence if and only if \tilde{G} does not have critical edges. \square

4. Complexity result. In this section, we show the main result of the paper (Theorem 1.1). Namely, we resolve the complexity status of the problem **FINITE-CONV**, defined as follows.

FINITE-CONV: Given a polynomial optimization problem (POP), does the corresponding Lasserre hierarchy $f^{(r)}$ have finite convergence to f^* ?

We will show that (FINITE-CONV) is NP-hard, already for standard quadratic programs. For this, we will consider the class of problems (M-S-B), where $B \in \mathcal{B}(G)$ and G is a graph. As mentioned in the introduction, the formulation by Motzkin and Straus for $\frac{1}{\alpha(G)}$ has already been used for resolving the complexity status of several problems in continuous optimization.

4.1. Linear programs. Consider first the case when (POP) is a linear optimization problem of the form

$$p^* = \inf \{c^T x : Ax \geq b, Dx = f\}, \tag{L-P}$$

where $A \in \mathbb{R}^{n \times m}$, $D \in \mathbb{R}^{n \times l}$, $b \in \mathbb{R}^m$, $f \in \mathbb{R}^l$, $c \in \mathbb{R}^n$. Let a_i (for $i \in [m]$) and d_i (for $i \in [l]$) denote, respectively, the rows of the matrices A and D . Then, problem (L-P) reads

$$p^* = \inf \left\{ c^T x : \begin{aligned} &a_i^T x - b_i \geq 0 \text{ for } i \in [m] \\ &d_i^T x - f_i = 0 \text{ for } i \in [l] \end{aligned} \right\}. \tag{L-P}$$

We show that for the class of problems (L-P) (FINITE-CONV) can be solved in polynomial time. Observe that the first level of the Lasserre sum-of-squares hierarchy for problem (L-P) finds its optimum p^* . Indeed, the first level of the hierarchy $p^{(1)}$ reads

$$p^{(1)} = \sup_{\substack{\lambda \in \mathbb{R}, \alpha \in \mathbb{R}_+^m \\ \beta \in \mathbb{R}^l}} \{ \lambda : c^T x - \lambda = \sum_{i=1}^m \alpha_i (a_i^T x - b_i) + \sum_{i=1}^l \beta_i (d_i^T x - f_i) \}$$

By equating coefficients at both sides we obtain that this problem reads

$$p^{(1)} = \sup_{\alpha \in \mathbb{R}_+^m, \beta \in \mathbb{R}^l} \{ \alpha^T b + \beta^T f : c = A^T \alpha + D^T \beta \}.$$

Note that this is precisely the dual linear program of (L-P). Hence, finite convergence always holds for linear programs, already at level $r = 1$.

4.2. Hardness in standard quadratic programs. We now show that the problem (FINITE-CONV) is NP-hard already for the class of standard quadratic programs. For this, we recall the recall problem (M-S-e) that was considered already in the Introduction. Let $G = (V, E)$ be a graph without twin pairs and let $e \in E$.

$$\frac{1}{\alpha(G)} = \min \left\{ x^T (A_G + I + A_{G \setminus e}) x : x \in \Delta_n \right\}. \tag{M-S-e}$$

We will show that deciding whether the Lasserre hierarchy of problem (M-S-e) has finite convergence is NP-hard. First, we show the following result that was already announced in the Introduction.

Proposition 1.2. Assume G has no twin pairs and let e be an edge of G . The Lasserre hierarchy of problem (M-S-e) has finite convergence if and only if the edge e is not critical in G .

Proof. This is a problem of the form (M-S-B) with $B = A_G + I + A_{G \setminus e}$. As a direct application of Theorem 3.1 we obtain the desired result. \square

To conclude, we show that deciding whether an edge is critical in a graph without twin pairs is an NP-hard problem. For this, we consider the following two problems.

CRITICAL-EDGE*: Given a graph $G = (V, E)$ without twin pairs and an edge $e \in E$, is e a critical edge of G ?

STABLE-SET: Given a graph G and $k \in \mathbb{N}$, does $\alpha(G) \geq k$ hold?

The problem STABLE-SET is well-known to be NP-Complete [8]. From this, we now prove that unless $P=NP$ there is no polynomial-time algorithm for solving CRITICAL-EDGE*. ¹

Theorem 4.1. *If there is a polynomial-time algorithm that solves the problem CRITICAL-EDGE*, then $P=NP$.*

Proof. Assume that there exists an algorithm A that solves CRITICAL-EDGE* in polynomial time. We will describe an algorithm that finds the stability number of a graph $\alpha(G)$ in polynomial time.

For describing the algorithm, we will first describe an operation that will be used repeatedly.

Contract all twin pairs $CTP(G)$:

Input: A graph G ,

Output: The graph G' obtained by successively contracting all the twin pairs from G . We have that G' does not have twin pairs,

$$|V(G')| \leq |V(G)|, E(G') \leq E(G) \text{ and } \alpha(G') = \alpha(G).$$

We observe that $CTP(G)$ can be carried out in polynomial time. Indeed, it can be checked in polynomial time whether G has a twin pair, by checking all pair of nodes and their sets of neighbors. If we find a twin pair (u, v) , then update the graph $G \rightarrow G \setminus u$ by deleting the node u , and we have $\alpha(G) = \alpha(G \setminus u)$. This operation is done successively until we obtain a graph G' that does not have twin pairs. Observe that the remaining graph satisfies that $\alpha(G') = \alpha(G)$.

Now we present an algorithm that finds the stability number of a graph.

Input: A graph G , **Output:** $\alpha(G)$.

Step 0: Initialize $d = 0$, $c = 0$, $H = G$

Step 1: Compute $CTP(H)$,

Update $d \rightarrow d + |H| - |CTP(H)|$,

Update $H \rightarrow CTP(H)$.

Step 2: If H has no edge STOP.

Step 3: Pick an edge e of H and decide if e is critical or not in H .

If e is critical, update $c \rightarrow c + 1$

Update $H \rightarrow H \setminus e$.

Go to Step 1.

Return: $|V(G)| - d - c$.

¹In a previous work of the author with Monique Laurent [16] it was shown that the problem of deciding whether an edge is critical in a graph is NP-hard. Here, we extend this argument to the problem CRITICAL-EDGE*, which has the additional condition that G has no twin pairs.

Observe that this algorithm runs in polynomial time. Indeed, Step 1 runs in polynomial time as explained above. Step 3 runs in polynomial time by assumption, as the graph does not have twin pairs. Now we verify the correctness of the algorithm. Clearly, the algorithm terminates because at each step either the number of nodes or edges is decreasing. Let \tilde{G} be the last update of the graph H in the algorithm. The only step where nodes are deleted is in Step 1, in the operation $CTP(H)$. The parameter d counts the number of nodes that are deleted. Therefore, the number of nodes of the remaining graph \tilde{G} is

$$|V(\tilde{G})| = |V(G)| - d.$$

In Step 1 the stability number $\alpha(H)$ remains invariant. In Step 3 the stability number increases by 1 precisely when e is critical. Hence,

$$\alpha(\tilde{G}) = \alpha(G) + c.$$

On the other hand, when the algorithm terminates the graph H does not have edges. Thus, $\alpha(\tilde{G}) = |V(\tilde{G})|$. In summary, we have found that

$$\alpha(G) = |V(G)| - d - c,$$

which is precisely the output of the algorithm. \square

We finally prove Theorem 1.1, which forms the main contribution of this paper. We repeat the statement for the sake of clarity.

Theorem 1.1. The problem of deciding whether the Lasserre hierarchy of a polynomial optimization problem has finite convergence is NP-hard.

Proof. Let G be a graph without twin pairs, and let e be an edge of G . We consider the problem (M-S- e). In Proposition 1.2, we show that the Lasserre hierarchy of this problem has finite convergence if and only if e is not critical. In Theorem 4.1 we show that the problem of deciding whether an edge is critical in a graph without twin pairs is NP-hard. Combining these two results we obtain that it is NP-hard to decide whether the Lasserre hierarchy of problem (M-S- e) has finite convergence. \square

Acknowledgements. The author thanks two anonymous referees for carefully reading this manuscript and for their valuable comments that helped to improve this paper. The author also wants to thank Amir Ali Ahmadi for bringing this problem to his attention and Monique Laurent for the useful discussions and her valuable comments about the presentation of this work. This work is supported by the European Union’s Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie Actions Grant Agreement No. 813211 (PO-EMA) and by the Swiss National Science Foundation project n. 200021_207429 / 1 “Ideal Membership Problems and the Bit Complexity of Sum of Squares Proofs”. Most of this research was carried out when the author was with Centrum Wiskunde & Informatica, Amsterdam.

REFERENCES

- [1] A. Ahmadi and J. Zhang, [On the complexity of finding a local minimizer of a quadratic function over a polytope](#), *Mathematical Programming*, **196** (2022), 783-792.
- [2] A. Ahmadi and J. Zhang, [Complexity aspects of local minima and related notions](#), *Advances in Mathematics*, **397** (2022), Article ID: 108119.
- [3] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.

- [4] I. Bomze and E. de Klerk, [Solving standard quadratic optimization problems via linear, semidefinite and copositive programming](#), *Journal of Global Optimization*, **24** (2002), 163-185.
- [5] L. E. Gibbons, D. W. Hearn, P. M. Pardalos and M. V. Ramana, [Continuous characterizations of the maximum clique problem](#), *Mathematics of Operations Research*, **22** (1997), 754-768.
- [6] D. Henrion and J. B. Lasserre, [GloptiPoly: global optimization over polynomials with Matlab and SeDuMi](#), *ACM Trans. Math. Soft.*, **29** (2003), 165-194.
- [7] D. Henrion and J. B. Lasserre, [Detecting global optimality and extracting solutions in GloptiPoly](#), *Positive Polynomials in Control. Lecture Notes in Control and Information Science*, **312** (2005), 293-310, Springer, Berlin.
- [8] R. Karp, *Reducibility Among Combinatorial Problems*, Plenum Press, New York, 1972.
- [9] E. de Klerk and D. Pasechnik, [Approximation of the stability number of a graph via copositive programming](#), *SIAM Journal on Optimization*, **12** (2002), 875-892.
- [10] T. -L. Kriel and M. Schweighofer, [On the exactness of Lasserre relaxations and pure states over real closed fields](#), *Foundations of Computational Mathematics*, **19** (2018), 1223-1263.
- [11] T. -L. Kriel and M. Schweighofer, [On the exactness of Lasserre relaxations for compact convex basic closed semialgebraic sets](#), *SIAM Journal of Optimization*, **28** (2018), 1796-1816.
- [12] J. B. Lasserre, [Global optimization with polynomials and the problem of moments](#), *SIAM Journal on Optimization*, **11** (2001), 796-817.
- [13] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press, 2009.
- [14] M. Laurent, [Semidefinite representations for finite varieties](#), *Mathematical Programming*, **109** (2007), 1-26.
- [15] M. Laurent, [Sums of squares, moment matrices and optimization over polynomials](#), in *Emerging Applications of Algebraic Geometry*, IMA Vol. Math. Appl. 149, M. Putinar and S. Sullivant, eds., Springer, Cham, (2009), 157-270.
- [16] M. Laurent and L. F. Vargas, [Finite convergence of sum-of-squares hierarchies for the stability number of a graph](#), *SIAM J. Optim.*, **32** (2022), 491-518.
- [17] M. Marshall, [Optimization of polynomial functions](#), *Canad. Math. Bull.*, **46** (2003), 575-587.
- [18] M. Marshall, [Representations of non-negative polynomials having finitely many zeros](#), *Annales de la faculté des sciences de Toulouse Mathématiques*, **15**, 599-609.
- [19] T. S. Motzkin and E. G. Straus, [Maxima for graphs and a new proof of a theorem of Turán](#), *Canadian J. Math.*, **17** (1965), 533-540.
- [20] K. G. Murty and S.N. Kabadi, [Some NP-complete problems in quadratic and nonlinear programming](#), *Math. Programming*, **39** (1987), 117-129.
- [21] J. Nie, [Optimality conditions and finite convergence of Lasserre's hierarchy](#), *Mathematical Programming*, **146** (2012), 97-121.
- [22] J. Nie, [Polynomial optimization with real varieties](#). *SIAM Journal on Optimization*, **23** (2013), 1634-1646.
- [23] M. Putinar, [Positive polynomials on compact semi-algebraic sets](#), *Ind. Univ. Math. J.*, **42** (1993), 969-984.
- [24] C. Scheidere, [Distinguished representations of non-negative polynomials](#), *J. Algebra*, **289** (2005), 558-573.
- [25] C. Scheiderer, [Sums of squares on real algebraic surfaces](#), *Manuscripta Math.*, **119** (2006), 395-410.
- [26] M. Schwieghofer and L. F. Vargas, [Sum-of-squares certificates for copositivity via test states](#), [arXiv:2310.12853](#), 2023.
- [27] L. F. Vargas, [Sum-of-Squares Representations for Copositive Matrices and Independent Sets in Graphs](#), PhD Thesis, Tilburg University, 2023.

Received January 2024; 1st revision April 2024; final revision May 2024; early access June 2024.