



The Saddle Point Problem of Polynomials

Jiawang Nie¹ · Zi Yang¹ · Guangming Zhou²

Received: 15 February 2019 / Revised: 27 April 2021 / Accepted: 7 June 2021 / Published online: 9 July 2021
© The Author(s) 2021

Abstract

This paper studies the saddle point problem of polynomials. We give an algorithm for computing saddle points. It is based on solving Lasserre’s hierarchy of semidefinite relaxations. Under some genericity assumptions on defining polynomials, we show that: (i) if there exists a saddle point, our algorithm can get one by solving a finite hierarchy of Lasserre-type semidefinite relaxations; (ii) if there is no saddle point, our algorithm can detect its nonexistence.

Keywords Saddle point · Polynomial · Nonsingularity · Lasserre relaxation · Semidefinite program

Mathematics Subject Classification 90C22 · 90C47 · 49K35 · 65K05

1 Introduction

Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$ be two sets (for dimensions $n, m > 0$), and let $F(x, y)$ be a continuous function in $(x, y) \in X \times Y$. A pair $(x^*, y^*) \in X \times Y$ is said to be a saddle point of $F(x, y)$ over $X \times Y$ if

$$F(x^*, y) \leq F(x^*, y^*) \leq F(x, y^*) \quad \forall x \in X, \forall y \in Y. \quad (1.1)$$

Communicated by Michael Overton.

✉ Jiawang Nie
njw@math.ucsd.edu

Zi Yang
ziy109@ucsd.edu

Guangming Zhou
zhougm@xtu.edu.cn

¹ Department of Mathematics, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

² School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, Hunan, China

The above implies that

$$\begin{aligned} F(x^*, y^*) &= \min_{x \in X} F(x, y^*) \leq \max_{y \in Y} \min_{x \in X} F(x, y), \\ F(x^*, y^*) &= \max_{y \in Y} F(x^*, y) \geq \min_{x \in X} \max_{y \in Y} F(x, y). \end{aligned}$$

On the other hand, it always holds that

$$\max_{y \in Y} \min_{x \in X} F(x, y) \leq \min_{x \in X} \max_{y \in Y} F(x, y).$$

Therefore, if (x^*, y^*) is a saddle point, then

$$\min_{x \in X} \max_{y \in Y} F(x, y) = F(x^*, y^*) = \max_{y \in Y} \min_{x \in X} F(x, y). \quad (1.2)$$

All saddle points share the same objective value, although there may exist different saddle points. The definition of saddle points in (1.1) requires the inequalities to hold for all points in the feasible sets X, Y . That is, when y is fixed to y^* , x^* is a global minimizer of $F(x, y^*)$ over X ; when x is fixed to x^* , y^* is a global maximizer of $F(x^*, y)$ over Y . Certainly, x^* must also be a local minimizer of $F(x, y^*)$ and y^* must be a local maximizer of $F(x^*, y)$. So, the local optimality conditions can be applied at (x^*, y^*) . However, they are not sufficient to guarantee that (x^*, y^*) is a saddle point, since (1.1) needs to be satisfied for all feasible points.

The saddle point problem of polynomials (SPPP) is for cases that $F(x, y)$ is a polynomial function in (x, y) and X, Y are semialgebraic sets, i.e., they are described by polynomial equalities and/or inequalities. The SPPP concerns the existence of saddle points and the computation of them if they exist. When F is convex–concave in (x, y) and X, Y are nonempty compact convex sets, there exists a saddle point. We refer to [5, §2.6] for the classical theory for convex–concave-type saddle point problems. The SPPPs have broad applications. They are of fundamental importance in duality theory for constrained optimization, min–max optimization and game theory [5,30,56]. The following are some applications.

- Zero-sum games can be formulated as saddle point problems [1,40,53]. In a zero-sum game with two players, suppose the first player has the strategy vector $x := (x_1, \dots, x_n)$ and the second player has the strategy vector $y := (y_1, \dots, y_m)$. The strategies x, y usually represent probability measures over finite sets, for instance, $x \in \Delta_n, y \in \Delta_m$. (The notation Δ_n denotes the standard simplex in \mathbb{R}^n .) A typical profit function of the first player is

$$f_1(x, y) := x^T A_1 x + y^T A_2 y + x^T B y,$$

for matrices A_1, A_2, B . For the zero-sum game, the profit function $f_2(x, y)$ of the second player is $-f_1(x, y)$. Each player wants to maximize the profit, for the given strategy of the other player. A Nash equilibrium is a point (x^*, y^*) such that the function $f_1(x, y^*)$ in x achieves the maximum value at x^* , while the function

$f_2(x^*, y)$ in y achieves the maximum value at y^* . Since $f_1(x, y) + f_2(x, y) = 0$, the Nash equilibrium (x^*, y^*) is a saddle point of the function $F := -f_1$ over feasible sets $X = \Delta_n, Y = \Delta_m$.

- The image restoration [23] can be formulated as a saddle point problem with the function

$$F(x, y) := x^T Ay + \frac{1}{2} \|Bx - z\|^2$$

and some feasible sets X, Y , for two given matrices A, B and a given vector z . Here, the notation $\| \cdot \|$ denotes the Euclidean norm. We refer to [12,19,23] for related work on this topic.

- The saddle point problem plays an important role in robust optimization [6,21,32,61]. For instance, a statistical portfolio optimization problem is

$$\min_{x \in X} -\mu^T x + x^T Qx,$$

where Q is a covariance matrix, μ is the estimation of some parameters, and X is the feasible set for the decision variable x . In applications, there often exists a perturbation for (μ, Q) . Suppose the perturbation for (μ, Q) is $(\delta\mu, \delta Q)$. There are two types of robust optimization problems

$$\begin{aligned} \min_{x \in X} \quad & \max_{(\delta\mu, \delta Q) \in Y} -(\mu + \delta\mu)^T x + x^T (Q + \delta Q)x, \\ \max_{(\delta\mu, \delta Q) \in Y} \quad & \min_{x \in X} -(\mu + \delta\mu)^T x + x^T (Q + \delta Q)x, \end{aligned}$$

where Y is the feasible set for the perturbation $(\delta\mu, \delta Q)$. People are interested in x^* and $(\delta\mu^*, \delta Q^*)$ such that the above two robust optimization problems are solved simultaneously by them. In view of (1.2), this is equivalent to solving a saddle point problem.

For convex–concave-type saddle point problems, most existing methods are based on gradients, subgradients, variational inequalities, or other related techniques. For these classical methods, we refer to the work by Chen et al. [13], Cox et al. [14], He and Yuan [23], He and Monteiro [24], Korpelevich [29], Maistroskii [38], Monteiro and Svaiter [39], Nemirovski [42], Nedić and Ozdaglar [41], and Zabotin [60]. For more general cases of non convex–concave-type saddle point problems (i.e., F is not convex–concave, and/or one of the sets X, Y is nonconvex), the computational task for solving SPPPs is much harder. A saddle point may, or may not, exist. There is very little work for solving non-convex–concave saddle point problems [16,51]. Obviously, SPPPs can be formulated as a first-order formula over the real field \mathbb{R} . By the Tarski–Seidenberg theorem [2,11], the SPPP is equivalent to a quantifier-free formula. Such quantifier-free formula can be computed symbolically, e.g., by cylindrical algebraic decompositions [2,11,28]. Theoretically, the quantifier elimination (QE) method can solve SPPPs exactly, but it typically has high computational complexity [10,17,54]. Another straightforward approach for solving (1.1) is to compute all its critical points first and then select saddle points among them. The complexity of computing critical

points is given in [55]. This approach typically has high computational cost, because the number of critical points is dramatically high [43] and we need to check the global optimality relation in (1.1) for getting saddle points. In Sect. 6.3, we will compare the performance between these methods and the new one given in this paper.

The basic questions for saddle point problems are: If a saddle point exists, how can we find it? If it does not, how can we detect its nonexistence? This paper discusses how to solve saddle point problems that are given by polynomials and that are non-convex–concave. We give a numerical algorithm to solve SPPPs.

1.1 Optimality Conditions

Throughout the paper, a property is said to hold *generically* in a space if it is true everywhere except a subset of zero Lebesgue measure. We refer to [22] for the notion of genericity in algebraic geometry. Assume X, Y are basic closed semialgebraic sets that are given as

$$X = \{x \in \mathbb{R}^n \mid g_i(x) = 0 (i \in \mathcal{E}_1^X), g_i(x) \geq 0 (i \in \mathcal{E}_2^X)\}, \tag{1.3}$$

$$Y = \{y \in \mathbb{R}^m \mid h_j(y) = 0 (j \in \mathcal{E}_1^Y), h_j(y) \geq 0 (j \in \mathcal{E}_2^Y)\}. \tag{1.4}$$

Here, each g_i is a polynomial in $x := (x_1, \dots, x_n)$ and each h_j is a polynomial in $y := (y_1, \dots, y_m)$. The $\mathcal{E}_1^X, \mathcal{E}_2^X, \mathcal{E}_1^Y, \mathcal{E}_2^Y$ are disjoint labeling sets of finite cardinalities. To distinguish equality and inequality constraints, denote the tuples

$$\boxed{\begin{aligned} g_{eq} &:= (g_i)_{i \in \mathcal{E}_1^X}, h_{eq} := (h_j)_{j \in \mathcal{E}_1^Y}, \\ g_{in} &:= (g_i)_{i \in \mathcal{E}_2^X}, h_{in} := (h_j)_{j \in \mathcal{E}_2^Y}. \end{aligned}} \tag{1.5}$$

When $\mathcal{E}_1^X = \emptyset$ (resp., $\mathcal{E}_2^X = \emptyset$), there is no equality (resp., inequality) constraint for X . The same holds for Y . For convenience, denote the labeling sets

$$\mathcal{E}^X := \mathcal{E}_1^X \cup \mathcal{E}_2^X, \quad \mathcal{E}^Y := \mathcal{E}_1^Y \cup \mathcal{E}_2^Y.$$

Suppose (x^*, y^*) is a saddle point. Then, x^* is a minimizer of

$$\begin{cases} \min_{x \in \mathbb{R}^n} & F(x, y^*) \\ \text{subject to} & g_i(x) = 0 (i \in \mathcal{E}_1^X), \\ & g_i(x) \geq 0 (i \in \mathcal{E}_2^X), \end{cases} \tag{1.6}$$

and y^* is a maximizer of

$$\begin{cases} \max_{y \in \mathbb{R}^m} & F(x^*, y) \\ \text{subject to} & h_j(y) = 0 (j \in \mathcal{E}_1^Y), \\ & h_j(y) \geq 0 (j \in \mathcal{E}_2^Y). \end{cases} \tag{1.7}$$

Under the linear independence constraint qualification (LICQ), or other kinds of constraint qualifications (see [4, §3.3]), there exist Lagrange multipliers λ_i, μ_j such that

$$\nabla_x F(x^*, y^*) = \sum_{i \in \mathcal{E}^X} \lambda_i \nabla_x g_i(x^*), \quad 0 \leq \lambda_i \perp g_i(x^*) \geq 0 \ (i \in \mathcal{E}_2^X), \quad (1.8)$$

$$\nabla_y F(x^*, y^*) = \sum_{j \in \mathcal{E}^Y} \mu_j \nabla_y h_j(y^*), \quad 0 \geq \mu_j \perp h_j(y^*) \geq 0 \ (j \in \mathcal{E}_2^Y). \quad (1.9)$$

In the above, $a \perp b$ means the product $a \cdot b = 0$ and $\nabla_x F$ (resp., $\nabla_y F$) denotes the gradient of $F(x, y)$ with respect to x (resp., y). When g, h are nonsingular (see the below for the definition), we can get explicit expressions for λ_i, μ_j in terms of x^*, y^* (see [49]). For convenience, write the labeling sets as

$$\mathcal{E}^X = \{1, \dots, \ell_1\}, \quad \mathcal{E}^Y = \{1, \dots, \ell_2\}.$$

Then, the constraining polynomial tuples can be written as

$$g = (g_1, \dots, g_{\ell_1}), \quad h = (h_1, \dots, h_{\ell_2}).$$

The Lagrange multipliers can be written as vectors

$$\lambda = (\lambda_1, \dots, \lambda_{\ell_1}), \quad \mu = (\mu_1, \dots, \mu_{\ell_2}).$$

Denote the matrices

$$G(x) := \begin{bmatrix} \nabla_x g_1(x) & \nabla_x g_2(x) & \cdots & \nabla_x g_{\ell_1}(x) \\ g_1(x) & 0 & \cdots & 0 \\ 0 & g_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g_{\ell_1}(x) \end{bmatrix}, \quad (1.10)$$

$$H(y) := \begin{bmatrix} \nabla_y h_1(y) & \nabla_y h_2(y) & \cdots & \nabla_y h_{\ell_2}(y) \\ h_1(y) & 0 & \cdots & 0 \\ 0 & h_2(y) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{\ell_2}(y) \end{bmatrix}. \quad (1.11)$$

The tuple g is said to be *nonsingular* if $\text{rank } G(x) = \ell_1$ for all $x \in \mathbb{C}^n$. Similarly, h is nonsingular if $\text{rank } H(y) = \ell_2$ for all $y \in \mathbb{C}^m$. Note that if g is nonsingular, then LICQ must hold at x^* . Similarly, the LICQ holds at y^* if h is nonsingular. When g, h have generic coefficients (i.e., g, h are generic), the tuples g, h are nonsingular. The nonsingularity is a property that holds generically. We refer to the work [49] for more details.

1.2 Contributions

This paper discusses how to solve saddle point problems of polynomials. We assume that the sets X, Y are given as in (1.3)–(1.4) and the defining polynomial tuples g, h are nonsingular, i.e., the matrices $G(x), H(y)$ have full column rank everywhere. Then, as shown in [49], there exist matrix polynomials $G_1(x), H_1(y)$ such that (I_ℓ denotes the $\ell \times \ell$ identity matrix)

$$G_1(x)G(x) = I_{\ell_1}, \quad H_1(y)H(y) = I_{\ell_2}. \quad (1.12)$$

When g, h have generic coefficients, they are nonsingular. Clearly, the above and (1.8)–(1.9) imply that

$$\lambda_i = G_1(x^*)_{i,1:n} \nabla_x F(x^*, y^*), \quad \mu_j = H_1(y^*)_{j,1:m} \nabla_y F(x^*, y^*).$$

(For a matrix A , the notation $A_{i,1:n}$ denotes its i th row with column indices from 1 to n .) Denote the Lagrange polynomial tuples

$$\lambda(x, y) := G_1(x)_{:,1:n} \nabla_x F(x, y), \quad (1.13)$$

$$\mu(x, y) := H_1(y)_{:,1:m} \nabla_y F(x, y). \quad (1.14)$$

(The notation $A_{:,1:n}$ denotes the submatrix of A consisting of its first n columns.) At each saddle point (x^*, y^*) , the Lagrange multiplier vectors λ, μ in (1.8)–(1.9) can be expressed as

$$\lambda = \lambda(x^*, y^*), \quad \mu = \mu(x^*, y^*).$$

Therefore, (x^*, y^*) is a solution to the polynomial system

$$\begin{cases} g_i(x) = 0 (i \in \mathcal{E}_1^X), \quad h_j(y) = 0 (j \in \mathcal{E}_1^Y), \\ \nabla_x F(x, y) = \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i(x), \\ \nabla_y F(x, y) = \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j(y), \\ 0 \leq \lambda_i(x, y) \perp g_i(x) \geq 0 (i \in \mathcal{E}_2^X), \\ 0 \geq \mu_j(x, y) \perp h_j(y) \geq 0 (j \in \mathcal{E}_2^Y). \end{cases} \quad (1.15)$$

However, not every solution (x^*, y^*) to (1.15) is a saddle point. This is because x^* might not be a minimizer of (1.6), and/or y^* might not be a maximizer of (1.7). How can we use (1.15) to get a saddle point? What further conditions do saddle points satisfy? When a saddle point does not exist, what is an appropriate certificate for the nonexistence? This paper addresses these questions. We give an algorithm for computing saddle points. First, we compute a candidate saddle point (x^*, y^*) . If it is verified to be a saddle point, then we are done. If it is not, then either x^* is not a minimizer of (1.6) or y^* is not a maximizer of (1.7). For either case, we add a new valid constraint to exclude such (x^*, y^*) , while all true saddle points are not excluded.

Then, we solve a new optimization problem, together with the newly added constraints. Repeating this process, we get an algorithm (i.e., Algorithm 3.1) for solving SPPPs. When the SPPP is given by generic polynomials, we prove that Algorithm 3.1 is able to compute a saddle point if it exists, and it can detect nonexistence if there does not exist one. The candidate saddle points are optimizers of certain polynomial optimization problems. We also show that these polynomial optimization problems can be solved exactly by Lasserre’s hierarchy of semidefinite relaxations, under some genericity conditions on defining polynomials. Since semidefinite programs are usually solved numerically (e.g., by `SeDuMi`) in practice, the computed solutions are correct up to numerical errors.

The paper is organized as follows: Sect. 2 reviews some basics for polynomial optimization. Section 3 gives an algorithm for solving SPPPs. We prove its finite convergence when the polynomials are generic. Section 4 discusses how to solve the optimization problems that arise in Sect. 3. Under some genericity conditions, we prove that Lasserre-type semidefinite relaxations can solve those optimization problems exactly. Proofs of some core theorems are given in Sect. 5. Numerical examples are given in Sect. 6. Conclusions and some discussions are given in Sect. 7.

2 Preliminaries

This section reviews some basics in polynomial optimization. We refer to [9,33,34,36,37,57] for the books and surveys in this field.

2.1 Notation

The symbol \mathbb{N} (resp., \mathbb{R} , \mathbb{C}) denotes the set of nonnegative integral (resp., real, complex) numbers. Denote by $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ the ring of polynomials in $x := (x_1, \dots, x_n)$ with real coefficients. The notation $\mathbb{R}[x]_d$ stands for the set of polynomials in $\mathbb{R}[x]$ with degrees $\leq d$. Sometimes, we need to work with polynomials in $y := (y_1, \dots, y_m)$ or $(x, y) := (x_1, \dots, x_n, y_1, \dots, y_m)$. The notation $\mathbb{R}[y]$, $\mathbb{R}[y]_d$, $\mathbb{R}[x, y]$, $\mathbb{R}[x, y]_d$ is similarly defined. For a polynomial p , $\deg(p)$ denotes its total degree. For $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer $\geq t$. For an integer $k > 0$, denote $[k] := \{1, 2, \dots, k\}$. For $\alpha := (\alpha_1, \dots, \alpha_l) \in \mathbb{N}^l$ with an integer $l > 0$, denote $|\alpha| := \alpha_1 + \dots + \alpha_l$. For an integer $d > 0$, denote

$$\mathbb{N}_d^l := \{\alpha \in \mathbb{N}^l \mid |\alpha| \leq d\}.$$

For $z = (z_1, \dots, z_l)$ and $\alpha = (\alpha_1, \dots, \alpha_l)$, denote

$$z^\alpha := z_1^{\alpha_1} \cdots z_l^{\alpha_l}, \quad [z]_d := [1 \ z_1 \ \cdots \ z_l \ z_1^2 \ z_1 z_2 \ \cdots \ z_l^d]^T.$$

In particular, we often use the notation $[x]_d$, $[y]_d$ or $[(x, y)]_d$. The superscript T denotes the transpose of a matrix/vector. The notation e_i denotes the i th standard unit vector, while e denotes the vector of all ones. The notation I_k denotes the k -by- k

identity matrix. By writing $X \geq 0$ (resp., $X > 0$), we mean that X is a symmetric positive semidefinite (resp., positive definite) matrix. For matrices X_1, \dots, X_r , $\text{diag}(X_1, \dots, X_r)$ denotes the block diagonal matrix whose diagonal blocks are X_1, \dots, X_r . For a vector z , $\|z\|$ denotes its standard Euclidean norm. For a function f in x , in y , or in (x, y) , $\nabla_x f$ (resp., $\nabla_y f$) denotes its gradient vector in x (resp., in y). In particular, F_{x_i} denotes the partial derivative of $F(x, y)$ with respect to x_i .

2.2 Positive Polynomials

In this subsection, we review some basic results about positive polynomials in the ring $\mathbb{R}[x, y]$. The same kind of results hold for positive polynomials in $\mathbb{R}[x]$ or $\mathbb{R}[y]$. An ideal I of $\mathbb{R}[x, y]$ is a subset such that $I \cdot \mathbb{R}[x, y] \subseteq I$ and $I + I \subseteq I$. For a tuple $p = (p_1, \dots, p_k)$ of polynomials in $\mathbb{R}[x, y]$, $\text{Ideal}(p)$ denotes the smallest ideal containing all p_i , which is the set

$$p_1 \cdot \mathbb{R}[x, y] + \dots + p_k \cdot \mathbb{R}[x, y].$$

In computation, we often need to work with the *truncation*:

$$\text{Ideal}(p)_{2k} := p_1 \cdot \mathbb{R}[x, y]_{2k - \deg(p_1)} + \dots + p_k \cdot \mathbb{R}[x, y]_{2k - \deg(p_k)}.$$

For an ideal $I \subseteq \mathbb{R}[x, y]$, its complex and real varieties are defined, respectively, as

$$\begin{aligned} \mathcal{V}_{\mathbb{C}}(I) &:= \{(u, v) \in \mathbb{C}^n \times \mathbb{C}^m \mid f(u, v) = 0 \forall f \in I\}, \\ \mathcal{V}_{\mathbb{R}}(I) &:= \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^m \mid f(u, v) = 0 \forall f \in I\}. \end{aligned}$$

A polynomial σ is said to be a sum of squares (SOS) if $\sigma = s_1^2 + \dots + s_k^2$ for some real polynomials s_1, \dots, s_k . Whether or not a polynomial is SOS can be checked by solving a semidefinite program (SDP) [31,50]. Clearly, if a polynomial is SOS, then it is nonnegative everywhere. However, the reverse may not be true. Indeed, there are significantly more nonnegative polynomials than SOS ones [8,9]. The set of all SOS polynomials in (x, y) is denoted as $\Sigma[x, y]$, and its d th truncation is

$$\Sigma[x, y]_d := \Sigma[x, y] \cap \mathbb{R}[x, y]_d.$$

For a tuple $q = (q_1, \dots, q_t)$ of polynomials in (x, y) , its *quadratic module* is

$$\text{Qmod}(q) := \Sigma[x, y] + q_1 \cdot \Sigma[x, y] + \dots + q_t \cdot \Sigma[x, y].$$

We often need to work with the truncation

$$\text{Qmod}(q)_{2k} := \Sigma[x, y]_{2k} + q_1 \cdot \Sigma[x, y]_{2k - \deg(q_1)} + \dots + q_t \cdot \Sigma[x, y]_{2k - \deg(q_t)}.$$

For two tuples $p = (p_1, \dots, p_k)$ and $q = (q_1, \dots, q_l)$ of polynomials in (x, y) , for convenience, we denote

$$\begin{cases} \text{IQ}(p, q) & := \text{Ideal}(p) + \text{Qmod}(q), \\ \text{IQ}(p, q)_{2k} & := \text{Ideal}(p)_{2k} + \text{Qmod}(q)_{2k}. \end{cases} \tag{2.1}$$

The set $\text{IQ}(p, q)$ (resp., $\text{IQ}(p, q)_{2k}$) is a convex cone that is contained in $\mathbb{R}[x, y]$ (resp., $\mathbb{R}[x, y]_{2k}$).

The set $\text{IQ}(p, q)$ is said to be *archimedean* if there exists $\sigma \in \text{IQ}(p, q)$ such that $\sigma(x, y) \geq 0$ defines a compact set in $\mathbb{R}^n \times \mathbb{R}^m$. If $\text{IQ}(p, q)$ is archimedean, then the set $K := \{p(x, y) = 0, q(x, y) \geq 0\}$ must be compact. The reverse is not always true. However, if K is compact, say, $K \subseteq B(0, R)$ (the ball centered at 0 with radius R), then $\text{IQ}(p, \tilde{q})$ is always archimedean, with $\tilde{q} = (q, R - \|x\|^2 - \|y\|^2)$, while $\{p(x, y) = 0, \tilde{q}(x, y) \geq 0\}$ defines the same set K . Under the assumption that $\text{IQ}(p, q)$ is archimedean, every polynomial in (x, y) , which is strictly positive on K , must belong to $\text{IQ}(p, q)$. This is the so-called Putinar’s Positivstellensatz [52]. Interestingly, under some optimality conditions, if a polynomial is nonnegative (but not strictly positive) over K , then it belongs to $\text{IQ}(p, q)$. This is shown in [46].

The above is for polynomials in (x, y) . For polynomials in only x or y , the ideals, sum of squares, quadratic modules, and their truncations are defined in the same way. The notations $\Sigma[x], \Sigma[x]_d, \Sigma[y], \Sigma[y]_d$ are similarly defined.

2.3 Localizing and Moment Matrices

Let $\xi := (\xi_1, \dots, \xi_l)$ be a subvector of $(x, y) := (x_1, \dots, x_n, y_1, \dots, y_m)$. Throughout the paper, the vector ξ is either x , or y , or (x, y) . Denote by $\mathbb{R}^{\mathbb{N}_d^l}$ the space of real sequences indexed by $\alpha \in \mathbb{N}_d^l$. A vector in $w := (w_\alpha)_{\alpha \in \mathbb{N}_d^l} \in \mathbb{R}^{\mathbb{N}_d^l}$ is called a *truncated multi-sequence* (tms) of degree d . It gives the *Riesz functional* \mathcal{R}_w acting on $\mathbb{R}[\xi]_d$ as (each $f_\alpha \in \mathbb{R}$)

$$\mathcal{R}_w \left(\sum_{\alpha \in \mathbb{N}_d^l} f_\alpha \xi^\alpha \right) := \sum_{\alpha \in \mathbb{N}_d^l} f_\alpha w_\alpha. \tag{2.2}$$

For $f \in \mathbb{R}[\xi]_d$ and $w \in \mathbb{R}^{\mathbb{N}_d^l}$, we denote

$$\langle f, w \rangle := \mathcal{R}_w(f). \tag{2.3}$$

Consider a polynomial $q \in \mathbb{R}[\xi]_{2k}$ with $\text{deg}(q) \leq 2k$. The k th *localizing matrix* of q , generated by a tms $w \in \mathbb{R}^{\mathbb{N}_{2k}^l}$, is the symmetric matrix $L_q^{(k)}(w)$ such that

$$\text{vec}(a_1)^T \left(L_q^{(k)}(w) \right) \text{vec}(a_2) = \mathcal{R}_w(q a_1 a_2) \tag{2.4}$$

for all $a_1, a_2 \in \mathbb{R}[\xi]_{k-\lceil \deg(q)/2 \rceil}$. (The $vec(a_i)$ denotes the coefficient vector of a_i .) For instance, when $n = 2$ and $k = 2$ and $q = 1 - x_1^2 - x_2^2$, we have

$$L_q^{(2)}[w] = \begin{bmatrix} w_{00} - w_{20} - w_{02} & w_{10} - w_{30} - w_{12} & w_{01} - w_{21} - w_{03} \\ w_{10} - w_{30} - w_{12} & w_{20} - w_{40} - w_{22} & w_{11} - w_{31} - w_{13} \\ w_{01} - w_{21} - w_{03} & w_{11} - w_{31} - w_{13} & w_{02} - w_{22} - w_{04} \end{bmatrix}.$$

When $q = 1$ (the constant one polynomial), $L_q^{(k)}(w)$ is called the *moment matrix* and we denote

$$M_k(w) := L_1^{(k)}(w). \tag{2.5}$$

The columns and rows of $L_q^{(k)}(w)$, as well as $M_k(w)$, are labeled by $\alpha \in \mathbb{N}^l$ with $2|\alpha| \leq 2k - \deg(q)$. When $q = (q_1, \dots, q_t)$ is a tuple of polynomials, then we define

$$L_q^{(k)}(w) := \text{diag}\left(L_{q_1}^{(k)}(w), \dots, L_{q_t}^{(k)}(w)\right), \tag{2.6}$$

which is a block diagonal matrix. Moment and localizing matrices are important tools for constructing semidefinite programming relaxations for solving moment and polynomial optimization problems [20,25,31,47]. Moreover, moment matrices are also useful for computing tensor decompositions [48]. We refer to [59] for a survey on semidefinite programming and applications.

3 An Algorithm for Solving SPPPs

Let F, g, h be the polynomial tuples for the saddle point problem (1.1). Assume g, h are nonsingular. So the Lagrange multiplier vectors $\lambda(x, y), \mu(x, y)$ can be expressed as in (1.13)–(1.14). We have seen that each saddle point (x^*, y^*) must satisfy (1.15). This leads us to consider the optimization problem

$$\left\{ \begin{array}{l} \min_{x \in X, y \in Y} F(x, y) \\ \text{subject to } \nabla_x F(x, y) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i(x) = 0, \\ \qquad \qquad \nabla_y F(x, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j(y) = 0, \\ \qquad \qquad 0 \leq \lambda_i(x, y) \perp g_i(x) \geq 0 \ (i \in \mathcal{E}_2^X), \\ \qquad \qquad 0 \geq \mu_j(x, y) \perp h_j(y) \geq 0 \ (j \in \mathcal{E}_2^Y), \end{array} \right. \tag{3.1}$$

where $\lambda_i(x, y)$ and $\mu_j(x, y)$ are Lagrange polynomials given as in (1.13)–(1.14). The saddle point problem (1.1) is not equivalent to (3.1). However, the optimization problem (3.1) can be used to get a candidate saddle point. Suppose (x^*, y^*) is a minimizer of (3.1). If x^* is a minimizer of $F(x, y^*)$ over X and y^* is a maximizer of $F(x^*, y)$ over Y , then (x^*, y^*) is a saddle point; otherwise, such (x^*, y^*) is not a saddle point, i.e., there exists $u \in X$ and/or there exists $v \in Y$ such that

$$F(u, y^*) - F(x^*, y^*) < 0 \quad \text{and/or} \quad F(x^*, v) - F(x^*, y^*) > 0.$$

The points u, v can be used to give new constraints

$$F(u, y) - F(x, y) \geq 0 \quad \text{and/or} \quad F(x, y) - F(x, v) \geq 0. \tag{3.2}$$

Every saddle point (x, y) must satisfy (3.2), so (3.2) can be added to the optimization problem (3.1) without excluding any true saddle points. For generic polynomials F, g, h , problem (3.1) has only finitely many feasible points (see Theorem 3.3). Therefore, by repeatedly adding new inequalities like (3.2), we can eventually get a saddle point or detect nonexistence of saddle points. This results in the following algorithm.

Algorithm 3.1 (An algorithm for solving saddle point problems.)

Input: The polynomials F, g, h as in (1.1), (1.3), (1.4) and Lagrange multiplier expressions as in (1.13)–(1.14).

Step 0: Let $K_1 = K_2 = S_a := \emptyset$ be empty sets.

Step 1: If problem (3.1) is infeasible, then (1.1) does not have a saddle point and stop; otherwise, solve (3.1) for a set K^0 of minimizers. Let $k := 0$.

Step 2: For each $(x^*, y^*) \in K^k$, do the following:

(a): (Lower-level minimization) Solve the problem

$$\begin{cases} \vartheta_1(y^*) := \min_{x \in X} F(x, y^*) \\ \text{subject to } \nabla_x F(x, y^*) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y^*) \nabla_x g_i(x) = 0, \\ 0 \leq \lambda_i(x, y^*) \perp g_i(x) \geq 0 \ (i \in \mathcal{E}_2^X), \end{cases} \tag{3.3}$$

and get a set of minimizers $S_1(y^*)$. If $F(x^*, y^*) > \vartheta_1(y^*)$, update

$$K_1 := K_1 \cup S_1(y^*).$$

(b): (Lower-level maximization) Solve the problem

$$\begin{cases} \vartheta_2(x^*) := \max_{y \in Y} F(x^*, y) \\ \text{subject to } \nabla_y F(x^*, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x^*, y) \nabla_y h_j(y) = 0, \\ 0 \geq \mu_j(x^*, y) \perp h_j(y) \geq 0 \ (j \in \mathcal{E}_2^Y) \end{cases} \tag{3.4}$$

and get a set of maximizers $S_2(x^*)$. If $F(x^*, y^*) < \vartheta_2(x^*)$, update

$$K_2 := K_2 \cup S_2(x^*).$$

(c): If $\vartheta_1(y^*) = F(x^*, y^*) = \vartheta_2(x^*)$, update:

$$S_a := S_a \cup \{(x^*, y^*)\}.$$

Step 3: If $S_a \neq \emptyset$, then each point in S_a is a saddle point and stop; otherwise go to Step 4.

Step 4: (Upper-level minimization) Solve the optimization problem

$$\left\{ \begin{array}{l} \min_{x \in X, y \in Y} F(x, y) \\ \text{subject to } \nabla_x F(x, y) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i(x) = 0, \\ \nabla_y F(x, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j(y) = 0, \\ 0 \leq \lambda_i(x, y) \perp g_i(x) \geq 0 (i \in \mathcal{E}_2^X), \\ 0 \geq \mu_j(x, y) \perp h_j(y) \geq 0 (j \in \mathcal{E}_2^Y), \\ F(u, y) - F(x, y) \geq 0 (u \in K_1), \\ F(x, v) - F(x, y) \leq 0 (v \in K_2). \end{array} \right. \quad (3.5)$$

If (3.5) is infeasible, then (1.1) has no saddle points and stop; otherwise, compute a set K^{k+1} of optimizers for (3.5). Let $k := k + 1$ and go to Step 2.

Output: If \mathcal{S}_a is nonempty, every point in \mathcal{S}_a is a saddle point; otherwise, output that there is no saddle point.

For generic polynomials, the feasible set \mathcal{K}_0 of (3.1), as well as each K^k in Algorithm 3.1, is finite. The convergence of Algorithm 3.1 is shown as follows.

Theorem 3.2 Let \mathcal{K}_0 be the feasible set of (3.1) and let \mathcal{S}_a be the set of saddle points for (1.1). If the complement set of \mathcal{S}_a in \mathcal{K}_0 (i.e., the set $\mathcal{K}_0 \setminus \mathcal{S}_a$) is finite, then Algorithm 3.1 must terminate after finitely many iterations. Moreover, if $\mathcal{S}_a \neq \emptyset$, then each $(x^*, y^*) \in \mathcal{S}_a$ is a saddle point; if $\mathcal{S}_a = \emptyset$, then there is no saddle point.

Proof At an iteration, if $\mathcal{S}_a \neq \emptyset$, then Algorithm 3.1 terminates. For each iteration with $\mathcal{S}_a = \emptyset$, each point $(x^*, y^*) \in K^k$ is not feasible for (3.5). When the k th iteration goes to the $(k + 1)$ th one, the nonempty sets

$$K^0, K^1, K^2, K^3, \dots, K^k$$

are disjoint from each other. All the points in K^i are not saddle points, so

$$\bigcup_{i=0}^k K^i \subseteq \mathcal{K}_0 \setminus \mathcal{S}_a.$$

Therefore, when the set $\mathcal{K}_0 \setminus \mathcal{S}_a$ is finite, Algorithm 3.1 must terminate after finitely many iterations.

When $\mathcal{S}_a \neq \emptyset$, each point $(x^*, y^*) \in \mathcal{S}_a$ is verified as a saddle point in Step 2. When $\mathcal{S}_a = \emptyset$, Algorithm 3.1 stops in Step 4 at some iteration, with the case that (3.5) is infeasible. Since every saddle point is feasible for both (3.1) and (3.5), there does not exist a saddle point if $\mathcal{S}_a = \emptyset$. \square

The number of iterations required by Algorithm 3.1 to terminate is bounded above by the cardinality of the complement set $\mathcal{K}_0 \setminus \mathcal{S}_a$, which is always less than or equal to the cardinality $|\mathcal{K}_0|$ of the feasible set of (3.1). Generally, it is hard to count $|\mathcal{K}_0 \setminus \mathcal{S}_a|$ or $|\mathcal{K}_0|$ accurately. When the polynomials F, g, h are generic, we can prove that the number of solutions for equality constraints in (3.1) is finite. For degrees $a_0, b_0 > 0$, denote the set product $\mathbb{C}[x, y]_{a_0, b_0} := \mathbb{C}[x]_{a_0} \cdot \mathbb{C}[y]_{b_0}$.

Theorem 3.3 *Let a_0, b_0 and $a_i, b_j > 0$ be positive degrees, for $i \in \mathcal{E}^X$ and $j \in \mathcal{E}^Y$. If $F(x, y) \in \mathbb{C}[x, y]_{a_0, b_0}$, $g_i \in \mathbb{C}[x]_{a_i}$, $h_j \in \mathbb{C}[y]_{b_j}$ are generic polynomials, then the polynomial system*

$$\begin{cases} \nabla_x F(x, y) = \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i(x), \\ g_i(x) = 0 \ (i \in \mathcal{E}_1^X), \ \lambda_i(x, y) g_i(x) = 0 \ (i \in \mathcal{E}_2^X), \\ \nabla_y F(x, y) = \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j(y), \\ h_j(y) = 0 \ (j \in \mathcal{E}_1^Y), \ \mu_j(x, y) h_j(y) = 0 \ (j \in \mathcal{E}_2^Y) \end{cases} \tag{3.6}$$

has only finitely many complex solutions in $\mathbb{C}^n \times \mathbb{C}^m$.

The proof for Theorem 3.3 will be given in Sect. 5. One would like to know what is the number of complex solutions to the polynomial system (3.6) for generic polynomials F, g, h . That number is an upper bound for $|\mathcal{K}_0|$ and so is also an upper bound for the number of iterations required by Algorithm 3.1 to terminate. The following theorem gives an upper bound for $|\mathcal{K}_0|$.

Theorem 3.4 *For the degrees a_i, b_j as in Theorem 3.3, let*

$$M := \sum_{\substack{\{i_1, \dots, i_{r_1}\} \subseteq [\ell_1], 0 \leq r_1 \leq n \\ \{j_1, \dots, j_{r_2}\} \subseteq [\ell_2], 0 \leq r_2 \leq m}} a_{i_1} \cdots a_{i_{r_1}} b_{j_1} \cdots b_{j_{r_2}} \cdot s \tag{3.7}$$

where in the above the number s is given as

$$s = \sum_{\substack{k_0 + \dots + k_{r_1+r_2} = n+m-r_1-r_2 \\ k_0, \dots, k_{r_1+r_2} \in \mathbb{N}}} (a_0 + b_0)^{k_0} (a_{i_1})^{k_1} \cdots (a_{i_{r_1}})^{k_{r_1}} (b_{j_1})^{k_{r_1+1}} \cdots (b_{j_{r_2}})^{k_{r_1+r_2}}.$$

If $F(x, y), g_i, h_j$ are generic, then (3.6) has at most M complex solutions, and hence, Algorithm 3.1 must terminate within M iterations.

The proof for Theorem 3.4 will be given in Sect. 5. We remark that the upper bound M given in (3.7) is not sharp. In our computational practice, Algorithm 3.1 typically terminates after a few iterations. It is an interesting question to obtain accurate upper bounds for the number of iterations required by Algorithm 3.1 to terminate.

4 Solving Optimization Problems

We discuss how to solve the optimization problems that appear in Algorithm 3.1. Under some genericity assumptions on F, g, h , we show that their optimizers can be computed by solving Lasserre-type semidefinite relaxations. Let X, Y be feasible sets given as in (1.3)–(1.4). Assume g, h are nonsingular, so $\lambda(x, y), \mu(x, y)$ can be expressed as in (1.13)–(1.14).

4.1 The Upper-Level Optimization

The optimization problem (3.1) is a special case of (3.5), with $K_1 = K_2 = \emptyset$. It suffices to discuss how to solve (3.5) with finite sets K_1, K_2 . For convenience, we rewrite (3.5) explicitly as

$$\left\{ \begin{array}{l} \min_{(x,y)} F(x, y) \\ \text{subject to } \nabla_x F(x, y) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i(x) = 0, \\ \nabla_y F(x, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j(y) = 0, \\ g_i(x) = 0, h_j(y) = 0 \ (i \in \mathcal{E}_1^X, j \in \mathcal{E}_1^Y), \\ \lambda_i(x, y) g_i(x) = 0, \mu_j(x, y) h_j(y) = 0 \ (i \in \mathcal{E}_2^X, j \in \mathcal{E}_2^Y), \\ g_i(x) \geq 0, \lambda_i(x, y) \geq 0 \ (i \in \mathcal{E}_2^X), \\ h_j(y) \geq 0, -\mu_j(x, y) \geq 0 \ (j \in \mathcal{E}_2^Y), \\ F(u, y) - F(x, y) \geq 0 \ (\forall u \in K_1), \\ F(x, y) - F(x, v) \geq 0 \ (\forall v \in K_2). \end{array} \right. \quad (4.1)$$

Recall that $\lambda_i(x, y), \mu_j(x, y)$ are Lagrange polynomials as in (1.13)–(1.14). Denote by ϕ the tuple of equality constraining polynomials

$$\begin{aligned} \phi := & \left\{ \nabla_x F - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y) \nabla_x g_i \right\} \cup \left\{ \nabla_y F - \sum_{j \in \mathcal{E}^Y} \mu_j(x, y) \nabla_y h_j \right\} \\ & \cup \left\{ g_i, h_j \right\}_{i \in \mathcal{E}_1^X, j \in \mathcal{E}_1^Y} \cup \left\{ \lambda_i(x, y) g_i, \mu_j(x, y) h_j \right\}_{i \in \mathcal{E}_2^X, j \in \mathcal{E}_2^Y}, \end{aligned} \quad (4.2)$$

and denote by ψ the tuple of inequality constraining ones

$$\begin{aligned} \psi := & \left\{ g_i, h_j, \lambda_i(x, y), -\mu_j(x, y) \right\}_{i \in \mathcal{E}_2^X, j \in \mathcal{E}_2^Y} \\ & \cup \left\{ F(u, y) - F(x, y), F(x, y) - F(x, v) \right\}_{u \in K_1, v \in K_2}. \end{aligned} \quad (4.3)$$

They are polynomials in (x, y) . Let

$$d_0 := \lceil \frac{1}{2} \max \{ \deg F(x, y), \deg(\phi), \deg(\psi) \} \rceil. \quad (4.4)$$

Then, the optimization problem (4.1) can be simply written as

$$\left\{ \begin{array}{l} f_* := \min F(x, y) \\ \text{subject to } \phi(x, y) = 0, \psi(x, y) \geq 0. \end{array} \right. \quad (4.5)$$

We apply Lasserre’s hierarchy of semidefinite relaxations to solve (4.5). For integers $k = d_0, d_0 + 1, \dots$, the k th-order semidefinite relaxation is

$$\left\{ \begin{array}{l} F_k := \min \langle F, w \rangle \\ \text{subject to } (w)_0 = 1, M_k(w) \geq 0, \\ L_\phi^{(k)}(w) = 0, L_\psi^{(k)}(w) \geq 0, w \in \mathbb{R}^{\mathbb{N}_{2k}^{n+m}}. \end{array} \right. \quad (4.6)$$

The number k is called a relaxation order. We refer to (2.4) for the localizing and moment matrices used in (4.6).

Algorithm 4.1 (An algorithm for solving the optimization (4.1).)

Input: Polynomials F, ϕ, ψ as in (4.2)–(4.3).

Step 0: Let $k := d_0$.

Step 1: Solve the semidefinite relaxation (4.6).

Step 2: If the relaxation (4.6) is infeasible, then (1.1) has no saddle points and stop; otherwise, solve it for a minimizer w^* . Let $t := d_0$.

Step 3 Check whether or not w^* satisfies the rank condition

$$\text{rank } M_t(w^*) = \text{rank } M_{t-d_0}(w^*). \tag{4.7}$$

Step 4 If (4.7) holds, extract $r := \text{rank } M_t(w^*)$ minimizers for (4.1) and stop.

Step 5 If $t < k$, let $t := t + 1$ and go to Step 3; otherwise, let $k := k + 1$ and go to Step 1.

Output: Minimizers of the optimization problem (4.1) or a certificate for the infeasibility of (4.1).

The conclusions in the Steps 2 and 3 are justified by Proposition 4.2. The rank condition (4.7) is called flat extension or flat truncation [15,44]. It is a sufficient and also almost necessary criterion for checking convergence of Lasserre-type relaxations [44]. When it is satisfied, the method in [27] can be applied to extract minimizers in Step 4. It was implemented in the software `GloptiPoly 3` [26].

Proposition 4.2 Suppose g, h are nonsingular polynomial tuples. For the hierarchy of relaxations (4.6), we have the properties:

- (i) If (4.6) is infeasible for some k , then (4.1) is infeasible and (1.1) has no saddle points.
- (ii) If (4.6) has a minimizer w^* satisfying (4.7), then $F_k = f_*$ and there are $r := \text{rank } M_t(w^*)$ minimizers for (4.1).

Proof Since g, h are nonsingular, every saddle point must be a critical point, and Lagrange multipliers can be expressed as in (1.13)–(1.14).

- (i) For each (u, v) that is feasible for (4.1), $[(u, v)]_{2k}$ satisfies all the constraints of (4.6), for all k . Therefore, if (4.6) is infeasible for some k , then (4.1) is infeasible.
- (ii) The conclusion follows from the classical results in [15,27,35,44].

□

We refer to (2.1) for the notation IQ , which is the sum of an ideal and a quadratic module. The polynomial tuples ϕ, ψ are from (4.2)–(4.3). Algorithm 4.1 is able to solve (4.1) successfully after finitely many iterations, under the following genericity conditions.

Condition 4.3 The polynomial tuples g, h are nonsingular and F, g, h satisfy one (not necessarily all) of the following:

- (1) $IQ(g_{eq}, g_{in}) + IQ(h_{eq}, h_{in})$ is archimedean;
- (2) the equation $\phi(x, y) = 0$ has finitely many real solutions;
- (3) $IQ(\phi, \psi)$ is archimedean.

In the above, the item (1) is almost the same as that X, Y are compact sets; the item (2) is the same as that (3.6) has only finitely many real solutions. Also note that the item (1) or (2) implies (3). In Theorem 3.3, we have shown that (3.6) has only finitely many complex solutions when F, g, h are generic. Therefore, Condition 4.3 holds generically. Under Condition 4.3, Algorithm 4.1 can be shown to have finite convergence.

Theorem 4.4 *Under Condition 4.3, we have that:*

- (i) *If problem (4.1) is infeasible, then the semidefinite relaxation (4.6) must be infeasible for all k big enough.*
- (ii) *Suppose (4.1) is feasible. If (4.1) has only finitely many minimizers and each of them is an isolated critical point (i.e., an isolated real solution of (3.6)), then, for all k big enough, (4.6) has a minimizer and each minimizer must satisfy the rank condition (4.7).*

We would like to remark that when F, g, h are generic, every minimizer of (4.1) is an isolated real solution of (3.6). This is because (3.6) has only finitely many complex solutions for generic F, g, h . Therefore, Algorithm 4.1 has finite convergence for generic cases. We would also like to remark that Proposition 4.2 and Theorem 4.4 assume that the semidefinite relaxation (4.6) is solved exactly. However, semidefinite programs are usually solved numerical (e.g., by SeDuMi), for better computational performance. Therefore, in computational practice, the optimizers obtained by Algorithm 4.1 are correct up to numerical errors. This is a common feature of all numerical methods.

4.2 Lower-Level Minimization

For a given pair (x^*, y^*) that is feasible for (3.1) or (3.5), we need to check whether or not x^* is a minimizer of $F(x, y^*)$ over X . This requires us to solve the minimization problem

$$\begin{cases} \min_{x \in \mathbb{R}^n} & F(x, y^*) \\ \text{subject to} & g_i(x) = 0 \ (i \in \mathcal{E}_1^X), \\ & g_i(x) \geq 0 \ (i \in \mathcal{E}_2^X). \end{cases} \tag{4.8}$$

When g is nonsingular, if it has a minimizer, the optimization (4.8) is equivalent to (by adding necessary optimality conditions)

$$\begin{cases} \min_{x \in \mathbb{R}^n} & F(x, y^*) \\ \text{subject to} & \nabla_x F(x, y^*) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y^*) \nabla_x g_i(x) = 0, \\ & g_i(x) = 0 \ (i \in \mathcal{E}_1^X), \ \lambda_i(x, y^*) g_i(x) = 0 \ (i \in \mathcal{E}_2^X), \\ & g_i(x) \geq 0, \ \lambda_i(x, y^*) \geq 0 \ (i \in \mathcal{E}_2^X). \end{cases} \tag{4.9}$$

Denote the tuple of equality constraining polynomials

$$\phi_{y^*} := \left\{ \begin{aligned} &\nabla_x F(x, y^*) - \sum_{i \in \mathcal{E}^X} \lambda_i(x, y^*) \nabla_x g_i \\ &\cup \{g_i\}_{i \in \mathcal{E}^X} \cup \{\lambda_i(x, y^*) \cdot g_i\}_{i \in \mathcal{E}_2^X}, \end{aligned} \right\} \tag{4.10}$$

and denote the tuple of inequality ones

$$\psi_{y^*} := \left\{ g_i, \lambda_i(x, y^*) \right\}_{i \in \mathcal{E}_2^X}. \tag{4.11}$$

They are polynomials in x but not in y , depending on the value of y^* . Let

$$d_1 := \left\lceil \frac{1}{2} \max\{\deg F(x, y^*), \deg(\phi_{y^*}), \deg(\psi_{y^*})\} \right\rceil. \tag{4.12}$$

We can rewrite (4.9) equivalently as

$$\left\{ \begin{aligned} &\min_{x \in \mathbb{R}^n} F(x, y^*) \\ &\text{subject to } \phi_{y^*}(x) = 0, \psi_{y^*}(x) \geq 0. \end{aligned} \right. \tag{4.13}$$

Lasserre’s hierarchy of semidefinite relaxations for solving (4.13) is

$$\left\{ \begin{aligned} &\min_z \langle F(x, y^*), z \rangle \\ &\text{subject to } (z)_0 = 1, M_k(z) \geq 0, \\ &\quad L_{\phi_{y^*}}^{(k)}(z) = 0, L_{\psi_{y^*}}^{(k)}(z) \geq 0, z \in \mathbb{R}^{\mathbb{N}_{2k}^n}, \end{aligned} \right. \tag{4.14}$$

for relaxation orders $k = d_1, d_1 + 1, \dots$. Since (x^*, y^*) is a feasible pair for (3.1) or (3.5), problems (4.8) and (4.13) are also feasible; hence, (4.14) is also feasible. A standard algorithm for solving (4.13) is as follows.

Algorithm 4.5 (An algorithm for solving the problem (4.13).)

Input: The point y^* and polynomials $F(x, y^*), \phi_{y^*}, \psi_{y^*}$ as in (4.10)–(4.11).

Step 0: Let $k := d_1$.

Step 1: Solve the semidefinite relaxation (4.14) for a minimizer z^* . Let $t := d_1$.

Step 2: Check whether or not z^* satisfies the rank condition

$$\text{rank } M_t(z^*) = \text{rank } M_{t-d_1}(z^*). \tag{4.15}$$

Step 3: If (4.15) holds, extract $r := \text{rank } M_t(z^*)$ minimizers and stop.

Step 4: If $t < k$, let $t := t + 1$ and go to Step 3; otherwise, let $k := k + 1$ and go to Step 1.

Output: Minimizers of the optimization problem (4.13).

Similar conclusions as in Proposition 4.2 hold for Algorithm 4.5. For cleanness of the paper, we do not state them again. The method in [27] can be applied to extract minimizers in the Step 3. Moreover, Algorithm 4.5 also terminates within finitely many iterations, under some genericity conditions.

Condition 4.6 *The polynomial tuple g is nonsingular and the point y^* satisfies one (not necessarily all) of the following:*

- (1) $IQ(g_{eq}, g_{in})$ is archimedean;
- (2) the equation $\phi_{y^*}(x) = 0$ has finitely many real solutions;
- (3) $IQ(\phi_{y^*}, \psi_{y^*})$ is archimedean.

Since (x^*, y^*) is feasible for (3.1) or (3.5), Condition 4.3 implies Condition 4.6, which also holds generically. The finite convergence of Algorithm 4.5 is summarized as follows.

Theorem 4.7 *Assume the optimization problem (4.8) has a minimizer and Condition 4.6 holds. If each minimizer of (4.8) is an isolated critical point, then, for all k big enough, (4.14) has a minimizer and each of them must satisfy (4.15).*

The proof of Theorem 4.7 will be given in Sect. 5. We would like to remark that every minimizer of (4.13) is an isolated critical point of (4.8), when F, g, h are generic. This is implied by Theorem 3.3.

4.3 Lower-Level Maximization

For a given pair (x^*, y^*) that is feasible for (3.1) or (3.5), we need to check whether or not y^* is a maximizer of $F(x^*, y)$ over Y . This requires us to solve the maximization problem

$$\begin{cases} \max_{y \in \mathbb{R}^m} & F(x^*, y) \\ \text{subject to} & h_j(y) = 0 \ (j \in \mathcal{E}_1^Y), \ h_j(y) \geq 0 \ (j \in \mathcal{E}_2^Y). \end{cases} \tag{4.16}$$

When h is nonsingular, if it has a minimizer, the optimization (4.16) is equivalent to (by adding necessary optimality conditions) the problem

$$\begin{cases} \max_{y \in \mathbb{R}^m} & F(x^*, y) \\ \text{subject to} & \nabla_y F(x^*, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x^*, y) \nabla_y h_j(y) = 0, \\ & h_j(y) = 0 \ (j \in \mathcal{E}_1^Y), \ \mu_j(x^*, y) \cdot h_j(y) = 0 \ (j \in \mathcal{E}_2^Y), \\ & h_j(y) \geq 0, \ -\mu_j(x^*, y) \geq 0 \ (j \in \mathcal{E}_2^Y). \end{cases} \tag{4.17}$$

Denote the tuple of equality constraining polynomials

$$\begin{aligned} \phi_{x^*} := & \left\{ \nabla_y F(x^*, y) - \sum_{j \in \mathcal{E}^Y} \mu_j(x^*, y) \nabla_y h_j \right\} \\ & \cup \{h_j\}_{j \in \mathcal{E}_1^Y} \cup \{\mu_j(x^*, y) h_j\}_{j \in \mathcal{E}_2^Y}, \end{aligned} \tag{4.18}$$

and denote the tuple of inequality ones

$$\psi_{x^*} := \left\{ h_j, -\mu_j(x^*, y) \right\}_{j \in \mathcal{E}_2^Y}. \tag{4.19}$$

They are polynomials in y but not in x , depending on the value of x^* . Let

$$d_2 := \left\lceil \frac{1}{2} \max \{ \deg F(x^*, y), \deg(\phi_{x^*}), \deg(\psi_{x^*}) \} \right\rceil. \tag{4.20}$$

Hence, (4.17) can be simply expressed as

$$\begin{cases} \max_{y \in \mathbb{R}^m} & F(x^*, y) \\ \text{subject to} & \phi_{x^*}(y) = 0, \psi_{x^*}(y) \geq 0. \end{cases} \tag{4.21}$$

Lasserre’s hierarchy of semidefinite relaxations for solving (4.21) is

$$\begin{cases} \max_z & \langle F(x^*, y), z \rangle \\ \text{subject to} & (z)_0 = 1, M_k(z) \geq 0, \\ & L_{\phi_{x^*}}^{(k)}(z) = 0, L_{\psi_{x^*}}^{(k)}(z) \geq 0, \\ & z \in \mathbb{R}^{\mathbb{N}_{2k}^m}, \end{cases} \tag{4.22}$$

for relaxation orders $k = d_2, d_2 + 1, \dots$. Since (x^*, y^*) is feasible for (3.1) or (3.5), problems (4.16) and (4.21) must also be feasible. Hence, the relaxation (4.22) is always feasible. Similarly, an algorithm for solving (4.21) is as follows.

Algorithm 4.8 (An algorithm for solving the problem (4.21).)

Input: The point x^* and polynomials $F(x^*, y), \phi_{x^*}, \psi_{x^*}$ as in (4.18)–(4.19).

Step 0: Let $k := d_2$.

Step 1: Solve the semidefinite relaxation (4.22) for a maximizer z^* . Let $t := d_2$.

Step 2: Check whether or not z^* satisfies the rank condition

$$\text{rank } M_t(z^*) = \text{rank } M_{t-d_2}(z^*). \tag{4.23}$$

Step 3: If (4.23) holds, extract $r := \text{rank } M_t(z^*)$ maximizers for (4.21) and stop.

Step 4: If $t < k$, let $t := t + 1$ and go to Step 3; otherwise, let $k := k + 1$ and go to Step 1.

Output: Maximizers of the optimization problem (4.21).

The same kind of conclusions like in Proposition 4.2 hold for Algorithm 4.8. The method in [27] can be applied to extract maximizers in Step 3. We can show that it must also terminate within finitely many iterations, under some genericity conditions.

Condition 4.9 The polynomial tuple h is nonsingular and the point x^* satisfies one (not necessarily all) of the following:

- (1) $IQ(h_{eq}, h_{in})$ is archimedean;

- (2) the equation $\phi_{x^*}(y) = 0$ has finitely many real solutions;
- (3) $IQ(\phi_{x^*}, \psi_{x^*})$ is archimedean.

By the same argument as for Condition 4.6, we can also see that Condition 4.9 holds generically. Similarly, Algorithm 4.8 also terminates within finitely many iterations under some genericity conditions.

Theorem 4.10 Assume that (4.16) has a maximizer and Condition 4.9 holds. If each maximizer of (4.16) is an isolated critical point, then, for all k big enough, (4.22) has a maximizer and each of them must satisfy (4.23).

The proof of Theorem 4.10 will be given in Sect. 5. Similarly, when F, g, h are generic, each maximizer of (4.16) is an isolated critical point of (4.16).

5 Some Proofs

This section gives the proofs for some theorems in the previous sections.

Proof of Theorem 3.3 Under the genericity assumption, the polynomial tuples g, h are nonsingular, so the Lagrange multipliers in (1.8)–(1.9) can be expressed as in (1.13)–(1.14). Hence, (3.6) is equivalent to the polynomial system in (x, y, λ, μ) :

$$\begin{cases} \nabla_x F(x, y) = \sum_{i \in \mathcal{E}^X} \lambda_i \nabla_x g_i(x), \\ \nabla_y F(x, y) = \sum_{j \in \mathcal{E}^Y} \mu_j \nabla_y h_j(y), \\ g_i(x) = 0 \ (i \in \mathcal{E}_1^X), \ \lambda_i g_i(x) = 0 \ (i \in \mathcal{E}_2^X), \\ h_j(y) = 0 \ (j \in \mathcal{E}_1^Y), \ \mu_j h_j(y) = 0 \ (j \in \mathcal{E}_2^Y). \end{cases} \tag{5.1}$$

Due to the complementarity conditions, $g_i(x) = 0$ or $\lambda_i = 0$ for each $i \in \mathcal{E}_2^X$, and $h_j(y) = 0$ or $\mu_j = 0$ for each $j \in \mathcal{E}_2^Y$. Note that if $g_i(x) \neq 0$ then $\lambda_i = 0$ and if $h_j(y) \neq 0$ then $\mu_j = 0$. Since $\mathcal{E}_2^X, \mathcal{E}_2^Y$ are finite labeling sets, there are only finitely many cases of $g_i(x) = 0$ or $g_i(x) \neq 0, h_j(y) = 0$ or $h_j(y) \neq 0$. We prove the conclusion is true for every case. Moreover, if $g_i(x) = 0$ for $i \in \mathcal{E}_2^X$, then the inequality $g_i(x) \geq 0$ can be counted as an equality constraint. The same is true for $h_j(y) = 0$ with $j \in \mathcal{E}_2^Y$. Therefore, we only need to prove the conclusion is true for the case that has only equality constraints. Without loss of generality, assume $\mathcal{E}_2^X = \mathcal{E}_2^Y = \emptyset$ and write the labeling sets as

$$\mathcal{E}_1^X = \{1, \dots, \ell_1\}, \quad \mathcal{E}_1^Y = \{1, \dots, \ell_2\}.$$

When all g_i are generic polynomials, the equations $g_i(x) = 0 \ (i \in \mathcal{E}_1^X)$ have no solutions if $\ell_1 > n$. Similarly, the equations $h_j(y) = 0 \ (j \in \mathcal{E}_1^Y)$ have no solutions if $\ell_2 > m$ and all h_j are generic. Therefore, we only consider the case that $\ell_1 \leq n$ and $\ell_2 \leq m$. When F, g, h are generic, we show that (5.1) cannot have infinitely many solutions. System (5.1) is the same as

$$\begin{cases} \nabla_x F(x, y) = \sum_{i=1}^{\ell_1} \lambda_i \nabla_x g_i(x), \ g_1(x) = \dots = g_{\ell_1}(x) = 0, \\ \nabla_y F(x, y) = \sum_{j=1}^{\ell_2} \mu_j \nabla_y h_j(y), \ h_1(y) = \dots = h_{\ell_2}(y) = 0. \end{cases} \tag{5.2}$$

Let $\tilde{x} = (x_0, x_1, \dots, x_n)$ and $\tilde{y} = (y_0, y_1, \dots, y_m)$. Denote by $\tilde{g}_i(\tilde{x})$ (resp., $\tilde{h}_j(\tilde{y})$) the homogenization of $g_i(x)$ (resp., $h_j(y)$). Let \mathbb{P}^n denote the n -dimensional complex projective space. Consider the projective variety

$$\mathcal{U} := \{(\tilde{x}, \tilde{y}) \in \mathbb{P}^n \times \mathbb{P}^m : \tilde{g}_i(\tilde{x}) = 0 (i \in \mathcal{E}^X), \tilde{h}_j(\tilde{y}) = 0 (j \in \mathcal{E}^Y)\}.$$

It is smooth, by Bertini’s theorem [22], under the genericity assumption on g_i, h_j . Denote the bi-homogenization of $F(x, y)$

$$\tilde{F}(\tilde{x}, \tilde{y}) := x_0^{a_0} y_0^{b_0} \tilde{F}(x/x_0, y/y_0).$$

When $F(x, y)$ is generic, the projective variety

$$\mathcal{V} := \mathcal{U} \cap \{\tilde{F}(\tilde{x}, \tilde{y}) = 0\}$$

is also smooth. One can directly verify that (for homogeneous polynomials)

$$\begin{aligned} x^T \nabla_x \tilde{F}(\tilde{x}, \tilde{y}) + x_0 \partial_{x_0} \tilde{F}(\tilde{x}, \tilde{y}) &= a_0 \tilde{F}(\tilde{x}, \tilde{y}), \\ x^T \nabla_x \tilde{g}_i(\tilde{x}) + x_0 \partial_{x_0} \tilde{g}_i(\tilde{x}) &= a_i \tilde{g}_i(\tilde{x}), \\ y^T \nabla_y \tilde{F}(\tilde{x}, \tilde{y}) + y_0 \partial_{y_0} \tilde{F}(\tilde{x}, \tilde{y}) &= b_0 \tilde{F}(\tilde{x}, \tilde{y}), \\ y^T \nabla_y \tilde{h}_j(\tilde{y}) + y_0 \partial_{y_0} \tilde{h}_j(\tilde{y}) &= b_j \tilde{h}_j(\tilde{y}). \end{aligned}$$

(They are called Euler’s identities.) Consider the determinantal variety

$$W := \left\{ (x, y) \in \mathbb{C}^n \times \mathbb{C}^m \left| \begin{array}{l} \text{rank} \begin{bmatrix} \nabla_x F(x, y) & \nabla_x g_1(x) & \dots & \nabla_x g_{\ell_1}(x) \end{bmatrix} \leq \ell_1 \\ \text{rank} \begin{bmatrix} \nabla_y F(x, y) & \nabla_y h_1(y) & \dots & \nabla_y h_{\ell_2}(y) \end{bmatrix} \leq \ell_2 \end{array} \right. \right\}.$$

Its homogenization is

$$\tilde{W} := \left\{ (\tilde{x}, \tilde{y}) \in \mathbb{P}^n \times \mathbb{P}^m \left| \begin{array}{l} \text{rank} \begin{bmatrix} \nabla_x \tilde{F}(\tilde{x}, \tilde{y}) & \nabla_x \tilde{g}_1(\tilde{x}) & \dots & \nabla_x \tilde{g}_{\ell_1}(\tilde{x}) \end{bmatrix} \leq \ell_1 \\ \text{rank} \begin{bmatrix} \nabla_y \tilde{F}(\tilde{x}, \tilde{y}) & \nabla_y \tilde{h}_1(\tilde{y}) & \dots & \nabla_y \tilde{h}_{\ell_2}(\tilde{y}) \end{bmatrix} \leq \ell_2 \end{array} \right. \right\}.$$

The projectivization of (5.2) is the intersection

$$\tilde{W} \cap \mathcal{U}.$$

If (3.6) has infinitely many complex solutions, so does (5.2). Then, $\tilde{W} \cap \mathcal{U}$ must intersect the hypersurface $\{\tilde{F}(\tilde{x}, \tilde{y}) = 0\}$. This means that there exists $(\tilde{x}, \tilde{y}) \in \mathcal{V}$ such that

$$\nabla_x \tilde{F}(\tilde{x}, \tilde{y}) = \sum_{i=1}^{\ell_1} \lambda_i \nabla_x \tilde{g}_i(\tilde{x}), \quad \nabla_y \tilde{F}(\tilde{x}, \tilde{y}) = \sum_{j=1}^{\ell_2} \mu_j \nabla_y \tilde{h}_j(\tilde{y}),$$

for some λ_i, μ_j . Also note $\tilde{g}_i(\bar{x}) = \tilde{h}_j(\bar{y}) = \tilde{F}(\bar{x}, \bar{y}) = 0$. Write

$$\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_n), \quad \bar{y} = (\bar{y}_0, \bar{y}_1, \dots, \bar{y}_m).$$

– If $\bar{x}_0 \neq 0$ and $\bar{y}_0 \neq 0$, by Euler’s identities, we can further get

$$\partial_{x_0} \tilde{F}(\bar{x}, \bar{y}) = \sum_{i=1}^{\ell_1} \lambda_i \partial_{x_0} \tilde{g}_i(\bar{x}), \quad \partial_{y_0} \tilde{F}(\bar{x}, \bar{y}) = \sum_{j=1}^{\ell_2} \mu_j \partial_{y_0} \tilde{h}_j(\bar{y}).$$

This implies that \mathcal{V} is singular, which is a contradiction.

– If $x_0 = 0$ but $y_0 \neq 0$, by Euler’s identities, we can also get

$$\partial_{y_0} \tilde{F}(\bar{x}, \bar{y}) = \sum_{j=1}^{\ell_2} \mu_j \partial_{y_0} \tilde{h}_j(\bar{y}).$$

This means the linear section $\mathcal{V} \cap \{x_0 = 0\}$ is singular, which is a contradiction again, by the genericity assumption on F, g, h .

– If $x_0 \neq 0$ but $y_0 = 0$, then we can have

$$\partial_{x_0} \tilde{F}(\bar{x}, \bar{y}) = \sum_{i=1}^{\ell_1} \lambda_i \partial_{x_0} \tilde{g}_i(\bar{x}).$$

So the linear section $\mathcal{V} \cap \{y_0 = 0\}$ is singular, which is again a contradiction.

– If $x_0 = y_0 = 0$, then $\mathcal{V} \cap \{x_0 = 0, y_0 = 0\}$ is singular. It is also a contradiction, under the genericity assumption on F, g, h .

For every case, we obtain a contradiction. Therefore, the polynomial system (3.6) must have only finitely many complex solutions, when F, g, h are generic. □

(Proof of Theorem 3.4) Each solution of (3.6) is a critical point of $F(x, y)$ over the set $X \times Y$. We count the number of critical points by enumerating all possibilities of active constraints. For an active labeling set $\{i_1, \dots, i_{r_1}\} \subseteq [\ell_1]$ (for X) and an active labeling set $\{j_1, \dots, j_{r_2}\} \subseteq [\ell_2]$ (for Y), an upper bound for the number is critical points $a_{i_1} \cdots a_{i_{r_1}} b_{j_1} \cdots b_{j_{r_2}} \cdot s$, which is given by Theorem 2.2 of [43]. Summing this upper bound for all possible active constraints, we eventually get the bound M . Since \mathcal{K}_0 is a subset of (3.6), Algorithm 3.1 must terminate within M iterations, for generic polynomials. □

(Proof of Theorem 4.4) In Condition 4.3, the item (1) or (2) implies (3). Note that the dual optimization problem of (4.6) is

$$\begin{cases} \max & \gamma \\ \text{subject to} & F - \gamma \in \text{IQ}(\phi, \psi)_{2k}. \end{cases} \tag{5.3}$$

(i) When (4.1) is infeasible, the set $\{\phi(x, y) = 0, \psi(x, y) \geq 0\}$ is empty. Since $\text{IQ}(\phi, \psi)$ is archimedean, by the classical Positivstellensatz [7] and Putinar’s Positivstellensatz [52], we have $-1 \in \text{IQ}(\phi, \psi)$. So, $-1 \in \text{IQ}(\phi, \psi)_{2k}$ for all such k big enough. Hence, (5.3) is unbounded from above for all big k . By weak duality, we know (4.6) must be infeasible.

(ii) When (4.1) is feasible, every feasible point is a critical point. By Lemma 3.3 of [18], $F(x, y)$ achieves finitely many values on $\phi(x, y) = 0$, say,

$$c_1 < c_2 < \dots < c_N.$$

Recall that f_* is the minimum value of (4.5). So, f_* is one of the c_i , say, $c_\ell = f_*$. Since (4.1) has only finitely many minimizers, we can list them as the set

$$O := \{(u_1, v_1), \dots, (u_B, v_B)\}.$$

If (x, y) is a feasible point of (4.1), then either $F(x, y) = c_k$ with $k > \ell$, or (x, y) is one of $(u_1, v_1), \dots, (u_B, v_B)$. Define the polynomial

$$P(x, y) := \left(\prod_{i=\ell+1}^N (F(x, y) - c_i)^2 \right) \cdot \left(\prod_{(u_j, v_j) \in O} (\|x - u_j\|^2 + \|y - v_j\|^2) \right).$$

We partition the set $\{\phi(x, y) = 0\}$ into four disjoint ones:

$$\begin{aligned} U_1 &:= \{\phi(x, y) = 0, c_1 \leq F(x, y) \leq c_{\ell-1}\}, \\ U_2 &:= \{\phi(x, y) = 0, F(x, y) = c_\ell, (x, y) \notin O\}, \\ U_3 &:= \{\phi(x, y) = 0, F(x, y) = c_\ell, (x, y) \in O\}, \\ U_4 &:= \{\phi(x, y) = 0, c_{\ell+1} \leq F(x, y) \leq c_N\}. \end{aligned}$$

Note that U_3 is the set of minimizers for (4.5).

- For all $(x, y) \in U_1$ and $i = \ell + 1, \dots, N$,

$$(F(x, y) - c_i)^2 \geq (c_{\ell-1} - c_{\ell+1})^2.$$

The set U_1 is closed and each $(u_j, v_j) \notin U_1$. The distance from (u_j, v_j) to U_1 is positive. Hence, there exists $\epsilon_1 > 0$ such that $P(x, y) > \epsilon_1$ for all $(x, y) \in U_1$.

- For all $(x, y) \in U_2$, $(F(x, y) - c_i)^2 = (c_\ell - c_i)^2$. For each $(u_j, v_j) \in O$, its distance to U_2 is positive. This is because each $(u_i, v_i) \in O$ is an isolated real critical point. So, there exists $\epsilon_2 > 0$ such that $P(x, y) > \epsilon_2$ for all $(x, y) \in U_2$.

Denote the new polynomial

$$q(x, y) := \min(\epsilon_1, \epsilon_2) - P(x, y).$$

On the set $\{\phi(x, y) = 0\}$, the inequality $q(x, y) \geq 0$ implies $(x, y) \in U_3 \cup U_4$. Therefore, (4.1) is equivalent to the optimization problem

$$\begin{cases} \min_{x,y} & F(x, y) \\ \text{subject to} & \phi(x, y) = 0, q(x, y) \geq 0. \end{cases} \tag{5.4}$$

Note that $q(x, y) > 0$ on the feasible set of (4.1). (This is because if (x, y) is a feasible point of (4.1), then $F(x, y) \geq f_* = c_\ell$, so $(x, y) \notin U_1$. If $F(x, y) = c_\ell$, then $(x, y) \in O$ and $P(x, y) = 0$, so $q(x, y) = \min(\epsilon_1, \epsilon_2) > 0$. If $F(x, y) > c_\ell$, then $P(x, y) = 0$ and we also have $q(x, y) = \min(\epsilon_1, \epsilon_2) > 0$.) By Condition 4.3 and Putinar’s Positivstellensatz, it holds that $q \in \text{IQ}(\phi, \psi)$. Now, we consider the hierarchy of Lasserre’s relaxations for solving (5.4):

$$\begin{cases} f'_k := \min \langle F, w \rangle \\ \text{subject to} & (w)_0 = 1, M_k(w) \geq 0, \\ & L_\phi^{(k)}(w) = 0, L_q^{(k)}(w) \geq 0. \end{cases} \tag{5.5}$$

Its dual optimization problem is

$$\begin{cases} f_k := \max \gamma \\ \text{subject to} & F - \gamma \in \text{IQ}(\phi, q)_{2k}. \end{cases} \tag{5.6}$$

□

Claim: For all k big enough, it holds that $f_k = f'_k = f_*$.

Proof The possible objective values of (5.4) are c_ℓ, \dots, c_N . Let p_1, \dots, p_N be real univariate polynomials such that $p_i(c_j) = 0$ for $i \neq j$ and $p_i(c_j) = 1$ for $i = j$. Let

$$s_i := (c_i - f_*)(p_i(F))^2, \quad i = \ell, \dots, N.$$

Then $s := s_\ell + \dots + s_N \in \Sigma[x]_{2k_1}$ for some order $k_1 > 0$. Let

$$\hat{F} := F - f_* - s.$$

Note that $\hat{F}(x) \equiv 0$ on the set

$$\mathcal{K}_2 := \{\phi(x, y) = 0, q(x, y) \geq 0\}.$$

It has a single inequality. By the Positivstellensatz [7, Corollary 4.1.8], there exist $0 < t \in \mathbb{N}$ and $Q = b_0 + qb_1$ ($b_0, b_1 \in \Sigma[x]$) such that $\hat{F}^{2t} + Q \in \text{Ideal}(\phi)$. Note that $Q \in \text{Qmod}(q)$. For all $\epsilon > 0$ and $\tau > 0$, we have $\hat{F} + \epsilon = \phi_\epsilon + \theta_\epsilon$ where

$$\begin{aligned} \phi_\epsilon &= -\tau\epsilon^{1-2t}(\hat{F}^{2t} + Q), \\ \theta_\epsilon &= \epsilon\left(1 + \hat{F}/\epsilon + \tau(\hat{F}/\epsilon)^{2t}\right) + \tau\epsilon^{1-2t}Q. \end{aligned}$$

By Lemma 2.1 of [45], when $\tau \geq \frac{1}{2t}$, there exists k_2 such that, for all $\epsilon > 0$,

$$\phi_\epsilon \in \text{Ideal}(\phi)_{2k_2}, \quad \theta_\epsilon \in \text{Qmod}(q)_{2k_2}.$$

Hence, we can get

$$F - (f_* - \epsilon) = \phi_\epsilon + \sigma_\epsilon,$$

where $\sigma_\epsilon = \theta_\epsilon + s \in \text{Qmod}(q)_{2k_2}$ for all $\epsilon > 0$. For all $\epsilon > 0$, $\gamma = f_* - \epsilon$ is feasible in (5.6) for the order k_2 , so $f_{k_2} \geq f_*$. Because $f_k \leq f_{k+1} \leq \dots \leq f_*$, we have $f_k = f'_k = f_*$ for all $k \geq k_2$. □

Because $q \in \text{Qmod}(\psi)$, each w , which is feasible for (4.6), is also feasible for (5.5). This can be implied by [44, Lemma 2.5]. So, when k is big, each w is also a minimizer of (5.5). The problem (5.4) also has only finitely many minimizers. By Theorem 2.6 of [44], the condition (4.7) must be satisfied for some $t \in [d_0, k]$, when k is big enough. □

(Proof of Theorem 4.7) The proof is the same as the one for Theorem 4.4. This is because the Lasserre’s relaxations (4.14) are constructed by using optimality conditions of (4.8), which is the same as for Theorem 4.4. In other words, Theorem 4.7 can be thought of a special version of Theorem 4.4 with $K_1 = K_2 = \emptyset$, without variable y . The assumptions are the same. Therefore, the same proof can be used. □

(Proof of Theorem 4.10) The proof is the same as the one for Theorem 4.7. □

6 Numerical Experiments

This section presents numerical examples of applying Algorithm 3.1 to solve saddle point problems. The computation is implemented in MATLAB R2012a, on a Lenovo Laptop with CPU@2.90GHz and RAM 16.0G. The Lasserre-type moment semidefinite relaxations are solved by the software GloptiPoly 3 [26], which calls the semidefinite program solver SeDuMi [58]. For cleanness, only four decimal digits are displayed for computational results.

In prior existing references, there are very few examples of non-convex–concave-type SPPPs. We construct various examples, with different types of functions and constraints. When g, h are nonsingular tuples, the Lagrange multipliers $\lambda(x, y), \mu(x, y)$ can be expressed by polynomials as in (1.13)–(1.14). Here we give some expressions for $\lambda(x, y)$ that will be frequently used in the examples. The expressions are similar for $\mu(x, y)$. Let $F(x, y)$ be the objective.

- For the simplex $\Delta_n = \{x \in \mathbb{R}^n : e^T x = 1, x \geq 0\}$, $g = (e^T x - 1, x_1, \dots, x_n)$ and we have

$$\lambda(x, y) = (x^T \nabla_x F, F_{x_1} - x^T \nabla_x F, \dots, F_{x_n} - x^T \nabla_x F). \tag{6.1}$$

– For the hypercube set $[-1, 1]^n$, $g = (1 - x_1^2, \dots, 1 - x_n^2)$ and

$$\lambda(x, y) = -\frac{1}{2}(x_1 F_{x_1}, \dots, x_n F_{x_n}). \quad (6.2)$$

– For the box constraint $[0, 1]^n$, $g = (x_1, \dots, x_n, 1 - x_1, \dots, 1 - x_n)$ and

$$\lambda(x, y) = ((1 - x_1)F_{x_1}, \dots, (1 - x_n)F_{x_n}, -x_1 F_{x_1}, \dots, -x_n F_{x_n}). \quad (6.3)$$

– For the ball $B_n(0, 1) = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ or sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, $g = 1 - x^T x$ and we have

$$\lambda(x, y) = -\frac{1}{2}x^T \nabla_x F. \quad (6.4)$$

– For the nonnegative orthant \mathbb{R}_+^n , $g = (x_1, \dots, x_n)$ and we have

$$\lambda(x, y) = (F_{x_1}, \dots, F_{x_n}). \quad (6.5)$$

We refer to [49] for more details about Lagrange multiplier expressions.

6.1 Some Explicit Examples

Example 6.1 Consider the simplex feasible sets $X = \Delta_n$, $Y = \Delta_m$. The Lagrange multipliers can be expressed as in (6.1).

(i) Let $n = m = 3$ and

$$F(x, y) = x_1 x_2 + x_2 x_3 + x_3 y_1 + x_1 y_3 + y_1 y_2 + y_2 y_3.$$

This function is neither convex in x nor concave in y . After 1 iteration by Algorithm 3.1, we got the saddle point:

$$x^* = (0.0000, 1.0000, 0.0000), \quad y^* = (0.2500, 0.5000, 0.2500).$$

It took about 2 s.

(ii) Let $n = m = 3$ and $F(x, y)$ be the function

$$x_1^3 + x_2^3 - x_3^3 - y_1^3 - y_2^3 + y_3^3 + x_3 y_1 y_2 (y_1 + y_2) + x_2 y_1 y_3 (y_1 + y_3) + x_1 y_2 y_3 (y_2 + y_3).$$

This function is neither convex in x nor concave in y . After 2 iterations by Algorithm 3.1, we got the saddle point

$$x^* = (0.0000, 0.0000, 1.0000), \quad y^* = (0.0000, 0.0000, 1.0000).$$

It took about 7.5 s.

(iii) Let $n = m = 4$ and

$$F(x, y) = \sum_{i,j=1}^4 x_i^2 y_j^2 - \sum_{i \neq j} (x_i x_j + y_i y_j).$$

This function is neither convex in x nor concave in y . After 2 iterations by Algorithm 3.1, we got 4 saddle points:

$$x^* = (0.2500, 0.2500, 0.2500, 0.2500), \quad y^* = e_i,$$

with $i = 1, 2, 3, 4$. It took about 99 s.

(iv) Let $n = m = 3$ and

$$F(x, y) := x_1 x_2 y_1 y_2 + x_2 x_3 y_2 y_3 + x_3 x_1 y_3 y_1 - x_1^2 y_3^2 - x_2^2 y_1^2 - x_3^2 y_2^2.$$

This function is neither convex in x nor concave in y . After 4 iterations by Algorithm 3.1, we got that there is no saddle point. It took about 32 s.

Example 6.2 Consider the box constraints $X = [0, 1]^n$ and $Y = [0, 1]^m$. The Lagrange multipliers can be expressed as in (6.3).

(i) Consider $n = m = 2$ and

$$F(x, y) := (x_1 + x_2 + y_1 + y_2 + 1)^2 - 4(x_1 x_2 + x_2 y_1 + y_1 y_2 + y_2 + x_1).$$

This function is convex in x but not concave in y . After 2 iterations by Algorithm 3.1, we got the saddle point

$$x^* = (0.3249, 0.3249), \quad y^* = (1.0000, 0.0000).$$

It took about 3.7 s.

(ii) Let $n = m = 3$ and

$$F(x, y) := \sum_{i=1}^n (x_i + y_i) + \sum_{i < j} (x_i^2 y_j^2 - y_i^2 x_j^2).$$

This function is neither convex in x nor concave in y . After 3 iterations by Algorithm 3.1, we got that there is no saddle point. It took about 12.8 s.

Example 6.3 Consider the cube constraints $X = Y = [-1, 1]^3$. The Lagrange multipliers can be expressed as in (6.2).

(i) Consider the function

$$F(x, y) := \sum_{i=1}^3 (x_i + y_i) - \prod_{i=1}^3 (x_i - y_i).$$

This function is neither convex in x nor concave in y . After 1 iteration by Algorithm 3.1, we got 3 saddle points:

$$\begin{aligned}x^* &= (-1.0000, -1.0000, 1.0000), & y^* &= (1.0000, 1.0000, 1.0000), \\x^* &= (-1.0000, 1.0000, -1.0000), & y^* &= (1.0000, 1.0000, 1.0000), \\x^* &= (1.0000, -1.0000, -1.0000), & y^* &= (1.0000, 1.0000, 1.0000).\end{aligned}$$

It took about 75 s.

(ii) Consider the function

$$F(x, y) := y^T y - x^T x + \sum_{1 \leq i < j \leq 3} (x_i y_j - x_j y_i).$$

This function is neither convex in x nor concave in y . After 4 iterations by Algorithm 3.1, we got the saddle point

$$x^* = (-1.0000, 1.0000, -1.0000), \quad y^* = (-1.0000, 1.0000, -1.0000).$$

It took about 6 s.

Example 6.4 Consider the sphere constraints $X = \mathbb{S}^2$ and $Y = \mathbb{S}^2$. They are not convex. The Lagrange multipliers can be expressed as in (6.4).

(i) Let $F(x, y)$ be the function

$$x_1^3 + x_2^3 + x_3^3 + y_1^3 + y_2^3 + y_3^3 + 2(x_1 x_2 y_1 y_2 + x_1 x_3 y_1 y_3 + x_2 x_3 y_2 y_3).$$

After 2 iterations by Algorithm 3.1, we got 9 saddle points $(-e_i, e_j)$, with $i, j = 1, 2, 3$. It took about 64 s.

(ii) Let $F(x, y)$ be the function

$$\begin{aligned}x_1^2 y_1^2 + x_2^2 y_2^2 + x_3^2 y_3^2 + x_1^2 y_2 y_3 + x_2^2 y_1 y_3 + x_3^2 y_1 y_2 \\+ y_1^2 x_2 x_3 + y_2^2 x_1 x_3 + y_3^2 x_1 x_2.\end{aligned}$$

After 4 iterations by Algorithm 3.1, we got that there is no saddle point. It took about 127 s.

Example 6.5 Let $X = Y = B_3(0, 1)$ be the ball constraints and

$$F(x, y) := x_1^2 y_1 + 2x_2^2 y_2 + 3x_3^2 y_3 - x_1 - x_2 - x_3.$$

The Lagrange multipliers can be expressed as in (6.4). The function F is not convex in x but is concave in y . After 1 iteration by Algorithm 3.1, we got the saddle point:

$$x^* = (0.7264, 0.4576, 0.3492), \quad y^* = (0.6883, 0.5463, 0.4772).$$

It took about 3.3 s.

Example 6.6 Consider the function

$$F(x, y) := x_1^2 y_2 y_3 + y_1^2 x_2 x_3 + x_2^2 y_1 y_3 + y_2^2 x_1 x_3 + x_3^2 y_1 y_2 + y_3^2 x_1 x_2$$

and the sets

$$X := \{x \in \mathbb{R}^3 : x^T x - 1 = 0, x \geq 0\}, \quad Y := \{y \in \mathbb{R}^3 : y^T y - 1 = 0, y \geq 0\}.$$

They are nonnegative portions of spheres. The feasible sets X, Y are non-convex. The Lagrange multipliers are expressed as

$$\begin{aligned} \lambda(x, y) &= \left(\frac{1}{2}x^T \nabla_x F, F_{x_1} - x_1 x^T \nabla_x F, F_{x_2} - x_2 x^T \nabla_x F, F_{x_3} - x_3 x^T \nabla_x F\right), \\ \mu(x, y) &= \left(\frac{1}{2}y^T \nabla_y F, F_{y_1} - y_1 y^T \nabla_y F, F_{y_2} - y_2 y^T \nabla_y F, F_{y_3} - y_3 y^T \nabla_y F\right). \end{aligned}$$

After 3 iterations by Algorithm 3.1, we got that there is no saddle point. It took about 37.3 s.

Example 6.7 Let $X = Y = \mathbb{R}_+^4$ be the nonnegative orthant and $F(x, y)$ be

$$\begin{aligned} & y_1(x_2 + x_3 + x_4 - 1)^2 + y_2(x_1 + x_3 + x_4 - 2)^2 + y_3(x_1 + x_2 + x_4 - 3)^2 \\ & - y_4(x_1 + x_2 + x_3 - 4)^2 - \left(x_1(y_2 + y_3 + y_4 - 1)^2 + x_2(y_1 + y_3 + y_4 - 2)^2\right. \\ & \left. - x_3(y_1 + y_2 + y_4 - 3)^2 + x_4(y_1 + y_2 + y_3 - 4)^2\right). \end{aligned}$$

The Lagrange multipliers can be expressed as in (6.5). The function F is neither convex in x nor concave in y . After 1 iteration by Algorithm 3.1, we got the saddle point

$$x^* = (1.5075, 0.5337, 0.0000, 0.5018), \quad y^* = (2.4143, 1.1463, 0.0000, 0.0000).$$

It took about 4.8 s.

Example 6.8 Let $X = Y = \mathbb{R}^3$ be the entire space, i.e., there are no constraints. There are no needs for Lagrange multiplier expressions. Consider the function

$$F(x, y) = \sum_{i=1}^3 (x_i^4 - y_i^4 + x_i + y_i) + \sum_{i \neq j} x_i^3 y_j^3.$$

It is neither convex in x nor concave in y . After 1 iteration by Algorithm 3.1, we got the saddle point

$$x^* = -(0.6981, 0.6981, 0.6981), \quad y^* = (0.4979, 0.4979, 0.4979).$$

It took about 113 s.

Example 6.9 Consider the sets and the function

$$\begin{aligned} X &:= \{x \in \mathbb{R}^3 : x_1 \geq 0, x_1x_2 \geq 1, x_2x_3 \geq 1\}, \\ Y &:= \{y \in \mathbb{R}^3 : y_1 \geq 0, y_1y_2 \geq 1, y_2y_3 \geq 1\}, \\ F(x, y) &:= x_1^3y_1 + x_2^3y_2 + x_3^3y_3 - 3x_1x_2x_3 - y_1^2 - 2y_2^2 - 3y_3^2. \end{aligned}$$

The function $F(x, y)$ is not convex in x but is concave in y . The Lagrange multipliers can be expressed as

$$\lambda_1 = (1 - x_1x_2)F_{x_1}, \quad \lambda_2 = x_1F_{x_1}, \quad \lambda_3 = -x_1F_{x_1} + x_2F_{x_2}.$$

The same expressions are for $\mu_j(x, y)$. After 9 iterations by Algorithm 3.1, we get the saddle point:

$$x^* = (1.2599, 1.2181, 1.3032), \quad y^* = (1.0000, 1.1067, 0.9036).$$

It took about 64 s.

6.2 Some Application Problems

Example 6.10 We consider the saddle point problem arising from zero-sum games with two players. Suppose $x \in \mathbb{R}^n$ is the strategy for the first player and $y \in \mathbb{R}^m$ is the strategy for the second one. The usual constraints for strategies are given by simplices, which represent probability measures on finite sets. So we consider feasible sets $X = \Delta_n, Y = \Delta_m$. Suppose the profit function of the first player is

$$f_1(x, y) = x^T A_1 x + y^T A_2 y + x^T B y,$$

for matrices $A_1 \in \mathbb{R}^{n \times n}$, $A_2 \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times m}$. For the zero-sum game, the profit function for the second player is $f_2(x, y) := -f_1(x, y)$. Each player wants to maximize the profit, for the given strategy of the other player. The Nash equilibrium is a point (x^*, y^*) such that the maximum of $f_1(x, y^*)$ over Δ_n is achieved at x^* , while the maximum of $f_2(x^*, y)$ over Δ_m is achieved at y^* . This is equivalent to that (x^*, y^*) is a saddle point of the function $F := -f_1(x, y)$ over X, Y . For instance, we

consider the matrices

$$A_1 = \begin{pmatrix} -4 & 4 & 0 & 3 & -4 \\ 3 & 4 & 3 & -4 & -5 \\ -3 & 0 & -2 & 0 & 4 \\ -4 & -4 & -1 & 3 & -5 \\ 4 & 1 & -3 & 0 & -5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -4 & 4 & 1 & 0 & 1 \\ -2 & -4 & 2 & -3 & 1 \\ -3 & 1 & 1 & 4 & 4 \\ 3 & -4 & 0 & 1 & -2 \\ -1 & -3 & -1 & 3 & -2 \end{pmatrix},$$

$$B = \begin{pmatrix} -2 & -4 & -2 & -5 & 3 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & -4 & -1 & -5 & 3 \\ 1 & -3 & -4 & 0 & -3 \\ 3 & -1 & -5 & 4 & -4 \end{pmatrix}.$$

The resulting saddle point problem is of the non-convex–concave type. After 2 iterations by Algorithm 3.1, we get two Nash equilibria

$$x^* = (0, 1, 0, 0, 0), \quad y^* = (1, 0, 0, 0, 0),$$

$$x^* = (0, 1, 0, 0, 0), \quad y^* = (0, 1, 0, 0, 0).$$

It took about 7 s.

Example 6.11 Consider the portfolio optimization problem [21,61]

$$\min_{x \in X} \quad -\mu^T x + x^T Q x,$$

where Q is a covariance matrix and μ is the estimation of some parameters. There often exists a perturbation $(\delta\mu, \delta Q)$ for (μ, Q) . This results in two types of robust optimization problems

$$\begin{aligned} \min_{x \in X} \quad & \max_{(\delta\mu, \delta Q) \in Y} \quad -(\mu + \delta\mu)^T x + x^T (Q + \delta Q)x, \\ \max_{(\delta\mu, \delta Q) \in Y} \quad & \min_{x \in X} \quad -(\mu + \delta\mu)^T x + x^T (Q + \delta Q)x. \end{aligned}$$

We look for x^* and $(\delta\mu^*, \delta Q^*)$ that can solve the above two robust optimization problems simultaneously. This is equivalent to the saddle point problem with $F = -(\mu + \delta\mu)^T x + x^T (Q + \delta Q)x$. For instance, consider the case that

$$Q = \begin{pmatrix} 5 & -4 & -2 \\ -4 & 13 & 10 \\ -2 & 10 & 8 \end{pmatrix}, \quad \mu = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix},$$

with the feasible sets

$$X := \{x \in \mathbb{R}^3 \mid -0.5 \leq x_i \leq 0.5, i = 1, \dots, n\},$$

$$Y := \left\{ (\delta\mu, \delta Q) \in \mathbb{R}^3 \times \mathcal{S}\mathbb{R}^{3 \times 3} \mid \begin{array}{l} -0.1 \leq (\delta\mu)_k, (\delta Q)_{ij} \leq 0.1, \\ 1 \leq k \leq 3, 1 \leq i, j \leq 3 \end{array} \right\}.$$

In the above, $S\mathbb{R}^{3 \times 3}$ denotes the space of real symmetric 3-by-3 matrices. The Lagrange multipliers can be similarly expressed as in (6.3). After 1 iteration by Algorithm 3.1, we got the saddle point

$$x^* = \begin{pmatrix} -0.1289 \\ -0.4506 \\ 0.5000 \end{pmatrix}, \delta Q^* = \begin{pmatrix} 0.1 & 0.1 & -0.1 \\ 0.1 & 0.1 & -0.1 \\ -0.1 & -0.1 & 0.1 \end{pmatrix}, \delta \mu^* = \begin{pmatrix} 0.1 \\ 0.1 \\ -0.1 \end{pmatrix}.$$

It took about 32 s. The above two min–max and max–min optimization problems are solved simultaneously by them.

6.3 Some Comparisons with Other Methods

Upon the request by referees, we give some comparisons between Algorithm 3.1 and other methods. The saddle point problems can also be solved by the straightforward approach of enumerating all KKT points. When all KKT points are to be computed, the numerical homotopy method such as Bertini [3] can be used. Saddle points can also be computed by quantifier elimination (QE) methods. Note that the definition (1.1) automatically gives a quantifier formula for saddle points. In the following, we give a comparison of the performance of these methods and Algorithm 3.1.

For computing all KKT points, the Maple function `Solve` is used. After they are obtained, we select saddle points from them by checking the definition. For the QE approach, the Maple function `QuantifierElimination` is used to solve the quantifier elimination formulae. For the numerical homotopy approach, the software Bertini is used to solve the KKT system for getting all KKT points first and then we select saddle points from them. When there are infinitely many KKT points, the function `Solve` and the software Bertini experience difficulty to get saddle points by enumerating KKT points. This happens for Examples 6.1(i), 6.2(i), 6.4(ii). The computational time (in seconds or hours) for these methods is reported in Table 1. We would like to remark that the Maple function `QuantifierElimination` cannot solve any example question (it does not terminate within 6 h for each one). So we also try the Mathematica function `Resolve` to implement the QE method. It can solve Example 6.1(iii) in about 1 second, but it cannot solve any other example question (it does not terminate within 6 h). The software Bertini can solve Examples 6.1(ii), 6.2(ii), 6.3(i), 6.4(i), 6.5, 6.7, 6.8, 6.9. For other example questions, it cannot finish within 6 hours. For these examples, Algorithm 3.1 took much less computational time, except Example 6.1(iii). We also like to remark that the symbolic methods like QE can obtain saddle points exactly in symbolic operations, while numerical methods can only obtain saddle points correctly up to round off errors.

7 Conclusions and Discussions

This paper discusses how to solve the saddle point problem of polynomials. We propose an algorithm (i.e., Algorithm 3.1) for computing saddle points. The Lasserre-type semidefinite relaxations are used to solve the polynomial optimization problems.

Table 1 Comparisons with other types of methods

Exemp.	Algorithm 3.1	KKT (Solve)	QE	Bertini
6.1(i)	2 s	∞ KKT points	> 6 h	∞ KKT points
6.1(ii)	7.5 s	> 6 h	> 6 h	380 s
6.1(iii)	99 s	190 s	1 s	> 6 h
6.1(iv)	32 s	> 6 h	> 6 h	> 6 h
6.2(i)	3.7 s	∞ KKT points	> 6 h	∞ KKT points
6.2(ii)	12.8 s	> 6 h	> 6 h	780 s
6.3(i)	75 s	30 s	> 6 h	1100 s
6.3(ii)	6 s	13011 s	> 6 h	> 6 h
6.4(i)	64 s	13921 s	> 6 h	441 s
6.4(ii)	127 s	∞ KKT points	> 6 h	∞ KKT points
6.5	3.3 s	> 6 h	> 6 h	8701 s
6.6	37.3 s	> 6 h	> 6 h	> 6 h
6.7	4.8 s	> 6 h	> 6 h	78 s
6.8	113 s	> 6 h	> 6 h	102 s
6.9	64 s	> 6 h	> 6 h	6293 s
6.10	7 s	> 6 h	> 6 h	> 6 h
6.11	32 s	> 6 h	> 6 h	> 6 h

Under some genericity assumptions, the proposed algorithm can compute a saddle point if there exists one. If there does not exist a saddle point, the algorithm can detect the nonexistence. However, we would like to remark that Algorithm 3.1 can always be applied, no matter whether the defining polynomials are generic or not. The algorithm needs to solve semidefinite programs for Lasserre-type relaxations. Since semidefinite programs are usually solved numerically (e.g., by SeDuMi), the computed solutions by Algorithm 3.1 are correct up to numerical errors. If the computed solutions are not accurate enough, classical nonlinear optimization methods can be applied to improve the accuracy. The method given in this paper can be used to solve saddle point problems from broad applications, such as zero-sum games, min–max optimization and robust optimization.

If the polynomials are such that the set \mathcal{K}_0 is infinite, then the convergence of Algorithm 3.1 is not theoretically guaranteed. For future work, the following questions are important and interesting.

Question 7.1 *When F, g, h are generic, what is an accurate (or sharp) upper bound for the number of iterations required by Algorithm 3.1 to terminate? What is the complexity of Algorithm 3.1 for generic F, g, h ?*

Theorem 3.4 gives an upper bound for the number of iterations. However, the bound given in (3.7) is certainly not sharp. Beyond the number of iterations, the complexity of solving the polynomial optimization problems (3.3), (3.4) and (3.5) is another concern. For efficient computational performance, Algorithms 4.1, 4.5 and 4.8 are applied to solve them. However, their complexities are mostly open, to the best of

the authors’ knowledge. We remark that Algorithms 4.1, 4.5 and 4.8 are based on the tight relaxation method in [49], instead of the classical Lasserre relaxation method in [31]. For the method in [31], there is no complexity result for the worst cases, i.e., there exist instances of polynomial optimization such that the method in [31] does not terminate within finitely many steps. The method in [49] always terminates within finitely many steps, under nonsingularity assumptions on constraints, while its complexity is currently unknown.

The finite convergence of Algorithm 3.1 is guaranteed if the set $\mathcal{K}_0 \setminus \mathcal{S}_a$ is finite. If it is infinite, it may or may not have finite convergence. If it does not, how can we get a saddle point? The following question is mostly open for the authors.

Question 7.2 *For polynomials F, g, h such that the set $\mathcal{K}_0 \setminus \mathcal{S}_a$ is not finite, how can we compute a saddle point if it exists? Or how can we detect its nonexistence if it does not exist?*

When X, Y are nonempty compact convex sets and the function F is convex–concave, there always exists a saddle point [5, §2.6]. However, if one of X, Y is nonconvex or if F is not convex–concave, a saddle point may, or may not, exist. This is the case even if F is a polynomial and X, Y are semialgebraic sets. The existence and nonexistence of saddle points for SPPPs are shown in various examples in Sect. 6. However, there is a new interesting property for SPPPs. We can write the objective polynomial $F(x, y)$ as

$$F(x, y) = [x]_{d_1}^T G [y]_{d_2},$$

for degrees $d_1, d_2 > 0$ and a matrix G (see Sect. 2 for the notation $[x]_{d_1}$ and $[y]_{d_2}$). Consider new variables $u := [x]_{d_1}, v := [y]_{d_2}$ and the new sets

$$\mathcal{X} = \{[x]_{d_1} : x \in X\}, \quad \mathcal{Y} = \{[y]_{d_2} : y \in Y\}.$$

A convex moment relaxation for the SPPP is to find $(u^*, v^*) \in \text{conv}(\mathcal{X}) \times \text{conv}(\mathcal{Y})$ such that

$$(u^*)^T G v \leq (u^*)^T G (v^*) \leq u^T G v^*$$

for all $u \in \text{conv}(\mathcal{X}), v \in \text{conv}(\mathcal{Y})$. (The notation $\text{conv}(T)$ denotes the convex hull of T .) When X, Y are nonempty compact sets, the above (u^*, v^*) always exists, because $u^T G v$ is bilinear in (u, v) and $\text{conv}(\mathcal{X}), \text{conv}(\mathcal{Y})$ are compact convex sets. In particular, if such u^* is an extreme point of $\text{conv}(\mathcal{X})$ and v^* is an extreme point of $\text{conv}(\mathcal{Y})$, say, $u^* = [a]_{d_1}$ and $v^* = [b]_{d_2}$ for $a \in X, b \in Y$, then (a, b) must be a saddle point of the original SPPP. If there is no saddle point (u^*, v^*) such that u^*, v^* are both extreme, the original SPPP does not have saddle points. We refer to [30] for related work about this kind of problems.

Acknowledgements Jiawang Nie and Zi Yang were partially supported by the NSF Grants DMS-1417985 and DMS-1619973.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. T. Başar and G. Olsder, *Dynamic noncooperative game theory*, SIAM, 1998
2. S. Basu, R. Pollack and M.-F. Roy, *Algorithms in Real Algebraic Geometry*, Algorithms and Computation in Mathematics, 10, Springer-Verlag, 2003.
3. D. Bates, J. Hauenstein, A. Sommese and C. Wampler, *Bertini: Software for Numerical Algebraic Geometry*, Available at bertini.nd.edu with permanent <https://doi.org/10.7274/R0H41PB5>.
4. D. Bertsekas, *Nonlinear programming*, second edition, Athena Scientific, 1995.
5. D. Bertsekas, A. Nedić and A. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, 2003.
6. D. Bertsimas, O. Nohadani and M. Teo, Robust optimization for unconstrained simulation-based problems, *Operations Research* 58(2010), no. 1, 161–178.
7. J. Bochnak, M. Coste and M.-F. Roy, *Real Algebraic Geometry*, Springer, 1998.
8. G. Blekherman, There are significantly more nonnegative polynomials than sums of squares, *Isr. J. Math.* 153(2006), 355–380.
9. G. Blekherman, P. Parrilo and R. Thomas (eds.), *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Series on Optimization, SIAM, 2013.
10. C. Brown and J. Davenport, The complexity of quantifier elimination and cylindrical algebraic decomposition, *Proceedings of the 2007 international symposium on Symbolic and algebraic computation*, 54–60, 2007.
11. B. Caviness and J. Johnson, *Quantifier elimination and cylindrical algebraic decomposition*, Springer Science & Business Media, 2012.
12. A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *Journal of Mathematical Imaging And Vision* 40(2011), no. 1, 120–145.
13. Y. Chen, G. Lan, and Y. Ouyang, Optimal primal-dual methods for a class of saddle point problems, *SIAM J. Optim.* 24(2014), no. 4, 1779–1814.
14. B. Cox, A. Juditsky, and A. Nemirovski, Decomposition techniques for bilinear saddle point problems and variational inequalities with affine monotone operators, *J. Optim. Theory Appl.* 172(2017), no. 2, 402–435.
15. R. Curto and L. Fialkow, Truncated K-moment problems in several variables, *J. Operator Theory* 54(2005), 189–226.
16. Y. Dauphin, R. Pascanu, C. Gülehre, K. Cho, S. Ganguli and Y. Bengio, Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, in: *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 2933–2941, Curran Associates, Inc., 2014.
17. J. Davenport and J. Hertrich, Real quantifier elimination is doubly exponential, *J. Symb. Comput.* 5(1988), no. 1-2, 29–35.
18. J. Demmel, J. Nie and V. Powers, Representations of positive polynomials on non-compact semialgebraic sets via KKT ideals, *J. Pure Appl. Algebra* 209(2007), no. 1, pp. 189–200.
19. E. Esser, X. Zhang, and T. F. Chan, A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science, *J. Imaging Sci.* 3(2010), 1015–1046.
20. L. Fialkow and J. Nie, The truncated moment problem via homogenization and flat extensions, *J. Funct. Anal.* 263(2012), no. 6, 1682–1700.
21. B. Halldórsson and R. Tütüncü, An interior-point method for a class of saddle-point problems, *J. Optim. Theory Appl.* 116(2003), no. 3, 559–590.
22. J. Harris, *Algebraic Geometry, A First Course*, Springer Verlag, 1992.

23. B. He and X. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective, *SIAM J. Imaging Sci.* 5(2012), no. 1, 119–149.
24. Y. He and R. Monteiro, Accelerating block-decomposition first-order methods for solving composite saddle-point and two-player nash equilibrium problems, *SIAM J. Optim.* 25(2015), no. 4, 2182–2211.
25. J.W. Helon and J. Nie, A semidefinite approach for truncated K-moment problems, *Found. Comput. Math.* 12(2012), no. 6, 851–881.
26. D. Henrion, J. Lasserre and J. Loeferberg, GloptiPoly 3: moments, optimization and semidefinite programming, *Optim. Methods Softw.* 24(2009), no. 4-5, 761–779.
27. D. Henrion and J. Lasserre, Detecting global optimality and extracting solutions in GloptiPoly, *Positive polynomials in control*, 293–310, *Lecture Notes in Control and Inform. Sci.*, 312, Springer, Berlin, 2005.
28. H. Hong and M. S. El Din, Variant quantifier elimination, *J. Symb. Comput.* 47(2012), 883–901.
29. G.M. Korpelevič, An extragradient method for finding saddle points and other problems, *Èkonom. i Mat. Metody* 12(1976), no. 4, 747–756.
30. R. Laraki and J. Lasserre, Semidefinite programming for min-max problems and games, *Math. Program.* 131(2012), 305–332.
31. J. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* 11(2011), no. 3, 796–817.
32. J. Lasserre, Min-max and robust polynomial optimization. *Journal of Global Optimization* 51(2011), no. 1, 1–10.
33. J. Lasserre, *Introduction to polynomial and semi-algebraic optimization*, Cambridge University Press, Cambridge, 2015.
34. J. Lasserre, The Moment-SOS Hierarchy, In *Proceedings of the International Congress of Mathematicians (ICM 2018)*, vol 4, B. Sirakov, P. Ney de Souza and M. Viana (Eds.), World Scientific, 2019, pp. 3773–3794.
35. M. Laurent, Revisiting two theorems of Curto and Fialkow on moment matrices, *Proceedings of the AMS* 133(2005), no. 10, 2965–2976.
36. M. Laurent, Sums of squares, moment matrices and optimization over polynomials, *Emerging Applications of Algebraic Geometry*, Vol. 149 of *IMA Volumes in Mathematics and its Applications* (Eds. M. Putinar and S. Sullivant), Springer, pp. 157–270, 2009.
37. M. Laurent, *Optimization over polynomials: Selected topics*, *Proceedings of the International Congress of Mathematicians*, 2014.
38. D. Maistrovskii, Gradient methods for finding saddle points, *Matekon* 13 (1977), 3–22.
39. R. Monteiro and B. Svaiter, Complexity of variants of Tseng’s modified F-B splitting and Korpelevich’s methods for hemivariational inequalities with applications to saddle-point and convex optimization problems, *SIAM J. Optim.* 21(2011), no. 4, 1688–1720.
40. J. Nash, Non-cooperative games, *Annals of Mathematics* (1951): 286–295
41. A. Nedić and A. Ozdaglar, Subgradient methods for saddle-point problems, *J. Optim. Theory Appl.* 142(2009), no. 1, 205–228.
42. A. Nemirovski, Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems, *SIAM J. Optim.* 15 (2004), no. 1, 229–251.
43. J. Nie and K. Ranestad, Algebraic degree of polynomial optimization, *SIAM J. Optim.* 20(2009), no. 1, 485–502.
44. J. Nie, Certifying convergence of Lasserre’s hierarchy via flat truncation, *Math. Program.* 142(2013), no. 1-2, 485–510.
45. J. Nie, Polynomial optimization with real varieties, *SIAM J. Optim.* 23(2013), no. 3, 1634–1646.
46. J. Nie, Optimality conditions and finite convergence of Lasserre’s hierarchy, *Math. Program.* 146(2014), no. 1-2, 97–121.
47. J. Nie, Linear optimization with cones of moments and nonnegative polynomials, *Math. Program.* 153(2015), no. 1, 247–274.
48. J. Nie, Generating polynomials and symmetric tensor decompositions, *Found. Comput. Math.* 17(2017), no. 2, 423–465.
49. J. Nie, Tight relaxations for polynomial optimization and lagrange multiplier expressions, *Math. Program.* 178(2019), no. 1-2, 1–37.
50. P. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Math. Program.* 96(2003), no.2, 293–320.

51. R. Pascanu, Y. Dauphin, S. Ganguli and Y. Bengio, On the saddle point problem for non-convex optimization, Preprint, 2014. [arXiv:1405.4604](https://arxiv.org/abs/1405.4604) [cs.LG]
52. M. Putinar, Positive polynomials on compact semi-algebraic sets, *Ind. Univ. Math. J.* 42(1993), 969–984.
53. L. Ratliff, S. Burden and S. Sastry, Characterization and computation of local Nash equilibria in continuous games, 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton), 917–924, 2013.
54. J. Renegar, On the computational complexity and geometry of the first-order theory of the reals, Parts I, II, III. *J. Symbolic Comput.* 13(1992), no. 3, 255–352.
55. P. J. Spaenlehauer, On the complexity of computing critical points with Gröbner bases, *SIAM J. Optim.* 24(2014), no. 3, 1382–1401.
56. P. Shah and P. Parrilo, Polynomial stochastic games via sum of squares optimization, Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, LA, USA, Dec. 12–14, 2007.
57. C. Scheiderer, Positivity and sums of squares: A guide to recent results, *Emerging Applications of Algebraic Geometry* (M. Putinar, S. Sullivant, eds.), IMA Volumes Math. Appl. 149, Springer, 2009, pp. 271–324.
58. J. Sturm, SeDuMi 1.02: A MATLAB toolbox for optimization over symmetric cones, *Optim. Methods Softw.* 11 & 12 (1999), 625–653.
59. M. Todd, Semidefinite optimization, *Acta Numerica* 10(2001), 515–560.
60. I. Zlotin, A subgradient method for finding a saddle point of a convex-concave function, *Issled. Prikl. Mat.* 15(1988), 6–12.
61. L. Zhu, T. Coleman and Y. Li, Min-max robust CVaR robust mean-variance portfolios, *Journal of Risk* 11(2009), no. 3, 55.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.