2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Inverse Function Theorem for Polynomial Equations using Semidefinite Programming

Morteza Ashraphijuo, Ramtin Madani, and Javad Lavaei

Abstract— **This paper is concerned with obtaining the inverse of polynomial functions using semidefinite programming (SDP). Given a polynomial function and a nominal point at which the Jacobian of the function is invertible, the inverse function theorem states that the inverse of the polynomial function exists at a neighborhood of the nominal point. In this work, we show that this inverse function can be found locally using convex optimization. More precisely, we propose infinitely many SDPs, each of which finds the inverse function at a neighborhood of the nominal point. We also design a convex optimization to check the existence of an SDP problem that finds the inverse of the polynomial function at multiple nominal points and a neighborhood around each point. This makes it possible to identify an SDP problem (if any) that finds the inverse function over a large region. As an application, any system of polynomial equations can be solved by means of the proposed SDP problem whenever an approximate solution is available. The method developed in this work is numerically compared with Newton's method and the nuclear-norm technique.**

I. Introduction

Consider the feasibility problem

$$\text{find} \qquad \mathbf{x} \in \mathbb{R}^m \tag{1a}$$

$$\text{subject to} \qquad \mathbf{P}(\mathbf{x}) = \mathbf{z}, \tag{1b}$$

for a given vector $\mathbf{z} \in \mathbb{R}^q$, where

$$\mathbf{P}(\mathbf{x}) \triangleq [P_1(\mathbf{x}), P_2(\mathbf{x}), \ldots, P_q(\mathbf{x})]^T \tag{2}$$

and $P_i : \mathbb{R}^m \to \mathbb{R}$ is a multivariate polynomial for $i = 1, \ldots, q$. The vector of variables $\mathbf{x} \in \mathbb{R}^m$ can be interpreted as the state of a system, for which the measurements or specifications $z_1, ..., z_q$ are known. Consider an arbitrary pair $(\mathbf{x}_0, \mathbf{z}_0)$ such that $\mathbf{P}(\mathbf{x}_0) = \mathbf{z}_0$. The vector $\mathbf{x}_0$ is referred to as a **nominal point** throughout this paper. If $m = q$ and the Jacobian of the function $\mathbf{P}(\cdot)$ is invertible at the point $\mathbf{x}_0$, then the inverse function $\mathbf{P}^{-1}(\mathbf{z})$ exists in a neighborhood of $\mathbf{z}_0$, in light of the inverse function theorem. A question arises as to whether this inverse function can be obtained via a convex optimization problem. This paper aims to address the above problem using a convex relaxation technique.

Semidefinite programming (SDP) is a subfield of convex optimization, which has received a considerable amount of attention in the past two decades [1], [2]. More recently, SDP has been used to convexify hard non-convex optimization problems in various areas, including graph theory, communication networks, and power systems [3]–[8]. For instance, the maximum likelihood problem for multi-input multi-output systems in information theory can be cast as an SDP problem [9]. SDP relaxations are powerful in solving both polynomial feasibility and polynomial optimization problems. SDP relaxations naturally arise in the method of moments and sum-of-squares techniques for finding a global minimum of a polynomial optimization or checking the emptiness of a semi-algebraic set [10]–[12].

This paper aims to exploit SDP relaxations to find the inverse of a polynomial function around a nominal point. The problem under study includes finding feasible solutions for polynomial equations as a special case. It is well known that checking the feasibility of a system of polynomial equations is NP-hard in general. Some classical approaches for obtaining a feasible point (if any) are Newton's method, Grobner bases, and homotopy. In many applications, an initial guess is available, which could be utilized as a starting point for finding an exact solution of (1). Newton-based methods benefit from a local convergence property, meaning that if the initial guess is close enough to a solution of (1), then these algorithms are guaranteed to converge after a modest number of iterations. However, the basin of attraction (the set of all such initial states) could be fractal, which makes the analysis of these methods hard [13], [14].

As a classic result, every system of polynomial equations can be reformulated as a system of quadratic equations by a change of variables [15]. Due to this equivalence, every polynomial optimization problem can be transformed into a quadratically-constrained quadratic program (QCQP), whose complexity is extensively studied in the literature [16], [17]. This transformation can be performed using a lifting technique (by introducing additional variables and constraints). In this work, we transform the polynomial feasibility problem (1) into a quadratic feasibility problem in a way that the invertibility of the Jacobian is preserved through this transformation (as needed by both the inverse function theorem and our approach).

By working on the quadratic formulation of the problem, we show that there are infinitely many SDP relaxations that have the same local property as Newton's method, and moreover their regions of attractions can all be explicitly characterized via nonlinear matrix inequalities. More precisely, for a given nominal pair of vectors $\mathbf{x}_0 \in \mathbb{R}^m$ and $\mathbf{z}_0 \in \mathbb{R}^q$ satisfying the relation $\mathbf{P}(\mathbf{x}_0) = \mathbf{z}_0$, we present a

Morteza Ashraphijuo and Javad Lavaei are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley. Ramtin Madani is with the Department of Electrical Engineering, Columbia University (Email: ashraphijuo@berkeley.edu, madani@ee.columbia.edu, and lavaei@berkeley.edu).

This work was supported by the ONR YIP Award, DARPA Young Faculty Award, NSF CAREER Award 1351279, and NSF EECS Award 1406865.

family of SDP relaxations that solve the feasibility problem (1) precisely as long as the solution belongs to a recoverable region. It is shown that this region contains $\mathbf{x}_0$ and a ball around it. As a result, the solution of the SDP relaxation, which depends on its input $\mathbf{z}$, is automatically the inverse function $\mathbf{P}^{-1}(\mathbf{z})$ over the recoverable region. Associated with each SDP in the proposed class of SDP problems, we characterize the recoverable region. We also study the problem of identifying an SDP relaxation whose recoverable region is relatively large and cast it as a convex optimization.

Our approach to finding the inverse function $\mathbf{P}^{-1}(\mathbf{z})$ locally is based on four steps: (i) transforming (1) into a quadratic problem, (ii) converting the quadratic formulation to a rank-constrained matrix feasibility problem, (iii) dropping the rank constraint, (iv) changing the matrix feasibility problem to an SDP optimization problem by incorporating a linear objective. The literature of compressed sensing considers the trace of the matrix variable as the objective function [18]–[20]. As opposed to the trace function that works only under strong assumptions (such as certain randomness), we consider a general linear function and design it in such a way that the SDP relaxation finds the inverse function $\mathbf{P}^{-1}(\mathbf{z})$ at least locally.

In this work, we precisely characterize the region of solutions that can be found using an SDP relaxation of (1). Roughly speaking, this region is provably larger for over-specified problems, i.e., whenever $q > m$. In particular, if q is sufficiently larger than m, the recoverable region would be the entire space (because in the extreme case the feasible set of the SDP relaxation becomes either empty or a single point). Over-specified systems of equations have applications in various problems, such as state estimation for wireless sensor networks [21], communication systems [22] and electric power systems [23]. We will demonstrate the efficacy of the proposed method for over-specified systems in a numerical example.

A. Notations

The symbols $\mathbb{R}$ and $\mathbb{S}^n$ denote the sets of real numbers and $n \times n$ real symmetric matrices, respectively. $\mathrm{rank}\{\cdot\}$, $\mathrm{trace}\{\cdot\}$, and $\det\{\cdot\}$ denote the rank, trace, and determinant of a given scalar/matrix. $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. Matrices are shown by capital and bold letters. The symbol $(\cdot)^T$ denotes the transpose operator. The notation $\langle \mathbf{A}, \mathbf{B} \rangle$ represents $\mathrm{trace}\{\mathbf{A}^{\mathrm{T}}\mathbf{B}\}$, which is the inner product of $\mathbf{A}$ and $\mathbf{B}$. The notation $\mathbf{W} \succeq 0$ means that $\mathbf{W}$ is a symmetric and positive semidefinite matrix. Also, $\mathbf{W} \succ 0$ means that it is symmetric and positive definite. The (i,j) entry of $\mathbf{W}$ is denoted as W_{ij}. The interior of a set $\mathcal{D} \in \mathbb{R}^n$ is denoted by $\mathrm{int}\{\mathcal{D}\}$. The notation $\mathrm{null}(\cdot)$ denotes the null space of a matrix. $\mathbf{0}_n$ and $\mathbf{1}_n$ denote the $n \times 1$ vectors of zeros and ones, respectively.

II. Preliminaries

Consider the arbitrary system of polynomial equations (1). This feasibility problem admits infinitely many quadratic formulations as

$$\begin{aligned} &\text{find} && \mathbf{v} \in \mathbb{R}^n && && \text{(3a)} \\ &\text{subject to} && F_i(\mathbf{v}) = y_i, && i = 1, 2, \ldots, l, && \text{(3b)} \end{aligned}$$

where $F_1, F_2, \ldots, F_l : \mathbb{R}^n \to \mathbb{R}$ are quadratic and homogeneous functions. For every $i \in \{1, ..., l\}$, $F_i(\mathbf{v})$ can be expressed as $\mathbf{v}^T\mathbf{M}_i\mathbf{v}$ for some symmetric matrix $\mathbf{M}_i$. Define

$$\mathbf{F}(\mathbf{v}) \triangleq [F_1(\mathbf{v}), F_2(\mathbf{v}), \ldots, F_l(\mathbf{v})]^T. \tag{4}$$

To elaborate on the procedure of obtaining the above quadratic form, a simple illustrative example will be provided below.

Illustrative example: Consider the system of polynomial equations

$$P_1(\mathbf{x}) \triangleq 3x_1^3x_2 - x_2^2 + 1 = 0, \tag{5a}$$
$$P_2(\mathbf{x}) \triangleq 2x_1 + x_2^4 - 4 = 0. \tag{5b}$$

Define

$$\mathbf{v}(\mathbf{x}) \triangleq [1 \;\; x_1 \;\; x_1^2 \;\; x_2 \;\; x_2^2 \;\; x_1x_2]^T. \tag{6}$$

Let $v_i(\mathbf{x})$ denote the i-th component of $\mathbf{v}(\mathbf{x})$ for $i = 1, \ldots, 6$. The system of polynomial equations in (5) can be reformulated in terms of the vector $\mathbf{v}(\mathbf{x})$:

$$3v_3(\mathbf{x})v_6(\mathbf{x}) - v_4^2(\mathbf{x}) + v_1^2(\mathbf{x}) = 0, \tag{7a}$$
$$2v_2(\mathbf{x})v_1(\mathbf{x}) + v_5^2(\mathbf{x}) - 4v_1^2(\mathbf{x}) = 0, \tag{7b}$$
$$v_3(\mathbf{x})v_1(\mathbf{x}) - v_2^2(\mathbf{x}) = 0, \tag{7c}$$
$$v_5(\mathbf{x})v_1(\mathbf{x}) - v_4^2(\mathbf{x}) = 0, \tag{7d}$$
$$v_6(\mathbf{x})v_1(\mathbf{x}) - v_2(\mathbf{x})v_4(\mathbf{x}) = 0, \tag{7e}$$
$$v_1^2(\mathbf{x}) = 1, \tag{7f}$$

The four additional equations (7c), (7d), (7e) and (7f) capture the structure of the vector $\mathbf{v}(\mathbf{x})$ and are added to preserve the equivalence of the two formulations. For notational simplicity, the short notation $\mathbf{v}$ is used for the variable $\mathbf{v}(\mathbf{x})$ henceforth.

A. Invariance of Jacobian

Given an arbitrary function $\mathbf{G} : \mathbb{R}^{m'} \to \mathbb{R}^{q'}$, we denote its Jacobian at point $\mathbf{x} \in \mathbb{R}^{m'}$ as

$$\nabla \mathbf{G}(\mathbf{x}) = \left[\frac{\partial G_j(\mathbf{x})}{\partial x_i}\right]_{i=1,\ldots,m';\, j=1,\ldots,q'}. \tag{8}$$

Note that some sources define the Jacobian as the transpose of the $m' \times q'$ matrix given in (8). The method to be proposed for finding the inverse function $\mathbf{P}^{-1}(\mathbf{x})$ at a neighborhood of $\mathbf{x}_0$ requires the Jacobian $\nabla\mathbf{P}(\mathbf{x}_0)$ to have full rank at the nominal point $\mathbf{x}_0$. Assume that this Jacobian matrix has full rank. Consider the equivalent quadratic formulation (3) and let $\mathbf{v}_0$ denote the nominal point for the reformulated problem associated with the point $\mathbf{x}_0$ for the original problem. A question arises as to whether the full-rank assumption of the Jacobian matrix is preserved through the transformation from

(1) to (3). It will be shown in the appendix that the quadratic reformulation can be done in a way that the relation

$$\nabla \mathbf{P}(\mathbf{x}_0) = \text{full rank} \quad \Longleftrightarrow \quad \nabla \mathbf{F}(\mathbf{v}_0) = \text{full rank} \tag{9}$$

is satisfied. The quadratic reformulation can be obtained in line with the illustrative example provided earlier. Note that the Jacobian of $\mathbf{F}(\mathbf{v})$ can be obtained as

$$\nabla \mathbf{F}(\mathbf{v}) = 2[\mathbf{M}_1\mathbf{v} \;\; \mathbf{M}_2\mathbf{v} \;\ldots\; \mathbf{M}_l\mathbf{v}]. \tag{10}$$

Definition 1: Define $\mathcal{I}_{\mathbf{F}}$ as the set of all vectors $\mathbf{v} \in \mathbb{R}^n$ for which $\nabla \mathbf{F}(\mathbf{v})$ has full row rank.

B. SDP relaxation

Observe that the quadratic constraints in (3b) can be expressed linearly in terms of the matrix $\mathbf{v}\mathbf{v}^T \in \mathbb{S}^n$, i.e.,

$$\mathbf{v}^T\mathbf{M}_i\mathbf{v} = \langle \mathbf{M}_i, \mathbf{v}\mathbf{v}^T \rangle. \tag{11}$$

Therefore, problem (3) can be cast in terms of a matrix variable $\mathbf{W} \in \mathbb{S}^n$ that replaces $\mathbf{v}\mathbf{v}^T$:

$$\begin{aligned} &\text{find} && \mathbf{W} \in \mathbb{S}^n && \text{(12a)}\\ &\text{subject to} && \langle \mathbf{M}_i, \mathbf{W}\rangle = y_i, \quad i = 1, 2, \ldots, l, && \text{(12b)}\\ &&& \mathbf{W} \succeq 0, && \text{(12c)}\\ &&& \operatorname{rank}(\mathbf{W}) = 1. && \text{(12d)} \end{aligned}$$

By dropping the constraint (12d) from the above non-convex optimization, and by penalizing its effect through minimizing an objective function, we obtain the following SDP problem.

Primal SDP:

$$\begin{aligned} &\underset{\mathbf{W} \in \mathbb{S}^n}{\text{minimize}} && \langle \mathbf{M}, \mathbf{W}\rangle && \text{(13a)}\\ &\text{subject to} && \langle \mathbf{M}_i, \mathbf{W}\rangle = y_i, \quad i = 1, \ldots, l && \text{(13b)}\\ &&& \mathbf{W} \succeq 0. && \text{(13c)} \end{aligned}$$

We intend to design the objective of the above SDP problem, namely the matrix $\mathbf{M}$, to guarantee the existence of a unique and rank-1 solution.

Definition 2: For a given positive semidefinite matrix $\mathbf{M} \in \mathbb{S}^n$, define $\mathcal{R}_{\mathbf{F}}(\mathbf{M})$ as the region of all vectors $\mathbf{v} \in \mathbb{R}^n$ for which $\mathbf{v}\mathbf{v}^T$ is the unique optimal solution of the SDP problem (13) for some vector $\mathbf{y} = [y_1, \ldots, y_l]^T$.

Given an arbitrary input vector $\mathbf{y}$, a solution $\mathbf{v}$ of the equation $\mathbf{F}(\mathbf{v}) = \mathbf{y}$ can be obtained from the primal SDP problem if and only if $\mathbf{v} \in \mathcal{R}_{\mathbf{F}}(\mathbf{M})$. As shown earlier in an illustrative example and in particular equation (6), the variable $\mathbf{v}$ of the quadratic formulation (3) is a function of the variable $\mathbf{x}$ of the original problem (1). In other words, $\mathbf{v}$ should technically be written as $\mathbf{v}(\mathbf{x})$. This fact will be used in the next definition.

Definition 3: For a given positive semidefinite matrix $\mathbf{M} \in \mathbb{S}^n$, define $\mathcal{R}_{\mathbf{P}}(\mathbf{M})$ as the region of all vectors $\mathbf{x} \in \mathbb{R}^m$ for which the corresponding vector $\mathbf{v}(\mathbf{x})$ belongs to $\mathcal{R}_{\mathbf{F}}(\mathbf{M})$.

Given an arbitrary input vector $\mathbf{z}$, a solution $\mathbf{x}$ of the equation $\mathbf{P}(\mathbf{x}) = \mathbf{z}$ can be obtained from the primal SDP problem if and only if $\mathbf{x} \in \mathcal{R}_{\mathbf{P}}(\mathbf{M})$. In the next section, we will show that $\mathcal{R}_{\mathbf{P}}(\mathbf{M})$ contains the nominal point $\mathbf{x}_0$ and a ball around this point in the case $m = q$. This implies that the inverse function $\mathbf{x} = \mathbf{P}^{-1}(\mathbf{z})$ exists in a neighborhood of $\mathbf{z}_0$ and can be obtained from an eigenvalue decomposition of the unique solution of the primal SDP problem (note that $\mathbf{F}^{-1}(\mathbf{y})\mathbf{F}^{-1}(\mathbf{y})^T$ becomes the "argmin" of the SDP problem over that region).

III. MAIN RESULTS

It is useful to consider the dual of the problem (13), which is stated below.

Dual SDP:

$$\begin{aligned} &\underset{\mathbf{u} \in \mathbb{R}^l}{\text{minimize}} && \mathbf{y}^{\mathrm{T}}\mathbf{u} && \text{(14a)}\\ &\text{subject to} && \mathbf{B}_{\mathbf{F}}(\mathbf{M}, \mathbf{u}) \succeq 0 && \text{(14b)} \end{aligned}$$

where $\mathbf{u} \in \mathbb{R}^l$ is the vector of dual variables and $\mathbf{B}_{\mathbf{F}} : \mathbb{S}^n \times \mathbb{R}^l \to \mathbb{S}^n$ is defined as

$$\mathbf{B}_{\mathbf{F}}(\mathbf{M}, \mathbf{u}) \triangleq \mathbf{M} + \sum_{i=1}^{l} u_i \mathbf{M}_i. \tag{15}$$

Definition 4: Given a nonnegative number ε, define $\mathcal{P}^n_\varepsilon$ as the set of all $n \times n$ positive semidefinite symmetric matrices with the sum of the two smallest eigenvalues greater than ε.

The following lemma provides a sufficient condition for strong duality.

Lemma 1: Suppose that $\mathbf{M} \in \mathcal{P}^n_0$ and $\operatorname{null}(\mathbf{M}) \subseteq \mathcal{I}_{\mathbf{F}}$. Then, strong duality holds between the primal SDP (13) and the dual SDP (14).

Proof: In order to show the strong duality, it suffices to build a strictly feasible point $\mathbf{u}$ for the dual problem. If $\mathbf{M} \succ 0$, then $\mathbf{u} = 0$ is a candidate. Now, assume that $\mathbf{M}$ has a zero eigenvalue. This eigenvalue must be simple due to the assumption $\mathbf{M} \in \mathcal{P}^n_0$. Let $\mathbf{h} \in \mathbb{R}^n$ be a nonzero eigenvector of $\mathbf{M}$ corresponding to its eigenvalue 0. The assumption $\operatorname{null}(\mathbf{M}) \subseteq \mathcal{I}_{\mathbf{F}}$ implies that

$$\mathbf{h}^T[\mathbf{M}_1\mathbf{h} \;\; \mathbf{M}_2\mathbf{h} \;\ldots\; \mathbf{M}_l\mathbf{h}] \neq 0.$$

Therefore, the relation $\mathbf{h}^T\mathbf{M}_k\mathbf{h} \neq 0$ holds for at least one index $k \in \{1, \ldots, l\}$. Let $e_1, \ldots, e_l$ be the standard basis vectors for $\mathbb{R}^l$. Set $\mathbf{u} = c \times e_k$, where c is a nonzero number with an arbitrarily small absolute value such that $c\,\mathbf{h}^T\mathbf{M}_k\mathbf{h} > 0$. Then, one can write

$$\mathbf{B}_{\mathbf{F}}(\mathbf{M}, \mathbf{u}) = \mathbf{M} + c\,\mathbf{M}_k \succ 0$$

if c is sufficiently small. ■

Lemma 2: Suppose that $\mathbf{M} \in \mathcal{P}^n_0$ and $\operatorname{null}(\mathbf{M}) \subseteq \mathcal{I}_{\mathbf{F}}$. Let $\mathbf{v} \in \mathcal{I}_{\mathbf{F}}$ be a feasible solution of problem (3) and $\mathbf{u} \in \mathbb{R}^l$ be a feasible point for the dual SDP (14). The following two statements are equivalent:

i) $(\mathbf{v}\mathbf{v}^T, \mathbf{u})$ is a pair of primal and dual optimal solutions for the primal SDP (13) and the dual SDP (14),

ii) $\mathbf{v} \in \operatorname{null}(\mathbf{B}_{\mathbf{F}}(\mathbf{M}, \mathbf{u}))$.

Proof: $(i) \Rightarrow (ii)$: According to Lemma 1, strong duality holds. Due to the complementary slackness, one can write

$$\begin{aligned} 0 &= \langle \mathbf{v}\mathbf{v}^T, \mathbf{B_F}(\mathbf{M}, \mathbf{u})\rangle \\ &= \operatorname{trace}\left\{\mathbf{v}\mathbf{v}^T \mathbf{B_F}(\mathbf{M}, \mathbf{u})\right\} \\ &= \mathbf{v}^T \mathbf{B_F}(\mathbf{M}, \mathbf{u})\mathbf{v}. \end{aligned} \tag{16}$$

On the other hand, it follows from the dual feasibility that

$$\mathbf{B_F}(\mathbf{M}, \mathbf{u}) \succeq 0,$$

which together with (16) concludes that $\mathbf{B_F}(\mathbf{M}, \mathbf{u})\mathbf{v} = 0$.

$(ii) \Rightarrow (i)$: Since $\mathbf{v} \in \mathcal{I}_\mathbf{F}$ is a feasible solution of (3), the matrix $\mathbf{v}\mathbf{v}^T$ is a feasible point for (13). On the other hand, since $\mathbf{v} \in \operatorname{null}(\mathbf{B_F}(\mathbf{M}, \mathbf{u}))$, we have

$$\langle \mathbf{v}\mathbf{v}^T, \mathbf{B_F}(\mathbf{M}, \mathbf{u})\rangle = 0,$$

which certifies the optimality of the pair $(\mathbf{v}\mathbf{v}^T, \mathbf{u})$. ■

Lemma 2 is particularly interesting in the special case $n = l$ (or equivalently $m = q$). In the sequel, we first study the case where the numbers of equations and parameters are the same, and then generalize the results to the case where the number of equations exceeds the number of unknown parameters. The latter scenario is referred to as an *over-specified* problem.

A. Region of Recoverable Solutions

In this subsection, we assume that the number of equations is equal to the number of unknowns, i.e., $l = n$. Given a positive semidefinite matrix $\mathbf{M} \in \mathcal{P}_0^n$, we intend to find the region $\mathcal{R}_\mathbf{F}(\mathbf{M})$, i.e., the set of all vectors that can be recovered using the convex problem (13).

Definition 5: Define the function of Lagrange multipliers $\mathbf{\Lambda} : \mathbb{S}^n \times \mathcal{I}_\mathbf{F} \to \mathbb{R}^l$ and the matrix function $\mathbf{A} : \mathbb{S}^n \times \mathcal{I}_\mathbf{F} \to \mathbb{S}^n$ as follows:

$$\begin{aligned} \mathbf{\Lambda_F}(\mathbf{M}, \mathbf{v}) &\triangleq -2(\nabla \mathbf{F}(\mathbf{v}))^{-1}\mathbf{M}\mathbf{v} \\ \mathbf{A_F}(\mathbf{M}, \mathbf{v}) &\triangleq \mathbf{B_F}(\mathbf{M}, \mathbf{\Lambda_F}(\mathbf{M}, \mathbf{v})). \end{aligned}$$

Lemma 3: Suppose that $n = l$ and let $\mathbf{v} \in \mathcal{I}_\mathbf{F}$. Then, we have $\mathbf{v} \in \operatorname{null}(\mathbf{B_F}(\mathbf{M}, \mathbf{u}))$ if and only if

$$\mathbf{u} = \mathbf{\Lambda_F}(\mathbf{M}, \mathbf{v}) \tag{17}$$

Proof: The equation $\mathbf{B_F}(\mathbf{M}, \mathbf{u})\mathbf{v} = 0$ can be rearranged as

$$[\mathbf{M}_1\mathbf{v} \;\; \mathbf{M}_2\mathbf{v} \;\; \ldots \;\; \mathbf{M}_n\mathbf{v}]\mathbf{u} = -\mathbf{M}\,\mathbf{v}.$$

Now, the proof follows immediately from the invertibility of $\nabla \mathbf{F}(\mathbf{v})$. ■

Whenever the SDP relaxation is exact, Lemmas 2 and 3 offer a closed-form relationship between a feasible solution of the problem (3) and an optimal solution of the dual problem (13), through the equation (17).

Since $\nabla \mathbf{F}(\mathbf{v_0})$ is invertible due to (9), the region $\mathbb{R}^n \backslash \mathcal{I}_\mathbf{F}$ is a set of measure zero in $\mathbb{R}^n$. In what follows, we characterize the interior of the region $\mathcal{R}_\mathbf{F}(\mathbf{M})$ restricted to $\mathcal{I}_\mathbf{F}$. It will be later shown that this region has dimension n. As a result, the next theorem is able to characterize $\mathcal{R}_\mathbf{F}(\mathbf{M})$ after excluding a subset of measure zero.

Theorem 1: Consider the case $n = l$. For every matrix $\mathbf{M} \in \mathcal{P}_0^n$ with the property $\operatorname{null}(\mathbf{M}) \subseteq \mathcal{I}_\mathbf{F}$, the equality

$$\mathcal{I}_\mathbf{F} \cap \operatorname{int}(R(\mathbf{M})) = \{\mathbf{v} \in \mathcal{I}_\mathbf{F} \,|\, \mathbf{A_F}(\mathbf{M}, \mathbf{v}) \in \mathcal{P}_0^n\}$$

holds.

Proof: We first need to show that $\{\mathbf{v} \in \mathcal{I}_\mathbf{F} \mid \mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n\}$ is an open set. Consider a vector $\mathbf{v}$ such that $\mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n$ and let δ denote the second smallest eigenvalue of $\mathbf{A_F}(\mathbf{v}, \mathbf{M})$. Due to the continuity of $\det\{\nabla \mathbf{F}(\cdot)\}$ and $\mathbf{A_F}(\cdot, \mathbf{M})$, there exists a neighborhood $\mathcal{B} \in \mathbb{R}^n$ around $\mathbf{v}$ such that for every $\mathbf{v}'$ within this neighborhood, $\mathbf{A_F}(\mathbf{v}', \mathbf{M})$ is well defined (i.e., $\mathbf{v}' \in \mathcal{I}_\mathbf{F}$) and

$$\|\mathbf{A_F}(\mathbf{v}', \mathbf{M}) - \mathbf{A_F}(\mathbf{v}, \mathbf{M})\|_F < \sqrt{\delta}. \tag{18}$$

It follows from an eigenvalue perturbation analysis that $\mathbf{A_F}(\mathbf{v}', \mathbf{M}) \in \mathcal{P}_0^n$ for every $\mathbf{v}' \in \mathcal{B}$. This proves that $\{\mathbf{v} \in \mathcal{I}_\mathbf{F} \mid \mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n\}$ is an open set. Now, consider a vector $\mathbf{v} \in \mathcal{I}_\mathbf{F}$ such that $\mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n$. The objective is to show that $\mathbf{v} \in \operatorname{int}\{\mathcal{R}_\mathbf{F}(\mathbf{M})\}$. Notice that since $\mathbf{A_F}(\mathbf{v}, \mathbf{M})$ is assumed to be in the set $\mathcal{P}_0^n$, the vector $\mathbf{\Lambda_F}(\mathbf{v}, \mathbf{M})$ is a feasible point for the dual problem (14). Therefore, it follows from Lemmas 2 and 3 that the matrix $\mathbf{v}\mathbf{v}^T$ is an optimal solution for the primal problem (13). In addition, every solution $\mathbf{W}$ must satisfy

$$\langle \mathbf{A_F}(\mathbf{v}, \mathbf{M}), \mathbf{W}\rangle = 0. \tag{19}$$

According to Lemma 3, $\mathbf{v}$ is an eigenvector of $\mathbf{A_F}(\mathbf{v}, \mathbf{M})$ corresponding to the eigenvalue 0. Therefore, since $\mathbf{A_F}(\mathbf{V}, \mathbf{M}) \succeq 0$ and $\operatorname{rank}\{\mathbf{A_F}(\mathbf{v}, \mathbf{M})\} = n - 1$, every positive semidefinite matrix $\mathbf{W}$ satisfying (19) is equal to $c\mathbf{v}\mathbf{v}^T$ for a nonnegative constant c. This concludes that $\mathbf{v}\mathbf{v}^T$ is the unique solution to (13), and therefore $\mathbf{v}$ belongs to $\mathcal{R}_\mathbf{F}(\mathbf{M})$. Since $\{\mathbf{v} \in \mathcal{I}_\mathbf{F} \mid \mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n\}$ is shown to be an open set, the above result can be translated as

$$\{\mathbf{v} \in \mathcal{I}_\mathbf{F} \mid \mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n\} \subseteq \operatorname{int}\{\mathcal{R}_\mathbf{F}(\mathbf{M})\} \cap \mathcal{I}_\mathbf{F}. \tag{20}$$

In order to complete the proof, it is requited to show that $\operatorname{int}\{\mathcal{R}_\mathbf{F}(\mathbf{M})\} \cap \mathcal{I}_\mathbf{F}$ is a subset of $\{\mathbf{v} \in \mathcal{I}_\mathbf{F} \mid \mathbf{A_F}(\mathbf{v}, \mathbf{M}) \in \mathcal{P}_0^n\}$. To this end, consider a vector $\mathbf{v} \in \operatorname{int}\{\mathcal{R}_\mathbf{F}(\mathbf{M})\} \cap \mathcal{I}_\mathbf{F}$. This means that $\mathbf{v}\mathbf{v}^T$ is a solution to (13), and therefore $\mathbf{A_F}(\mathbf{v}, \mathbf{M}) \succeq 0$, due to Lemma 2. To prove the aforementioned inclusion by contradiction, suppose that $\mathbf{A_F}(\mathbf{v}, \mathbf{M}) \notin \mathcal{P}_0^n$, implying that 0 is an eigenvalue of $\mathbf{A_F}(\mathbf{v}, \mathbf{M})$ with multiplicity at least 2. Let $\hat{\mathbf{v}}$ denote a second eigenvector corresponding to the eigenvalue 0 such that $\mathbf{v}^T\hat{\mathbf{v}} = 0$. Since $\mathbf{v} \in \mathcal{I}_\mathbf{F}$, in light of the inverse function theorem, there exists a constant $\varepsilon_0 > 0$ with the property that for every $\varepsilon \in [0, \varepsilon_0]$, there is a vector $\mathbf{w}_\varepsilon \in \mathbb{R}^n$ satisfying the relation

$$\mathbf{F}(\mathbf{w}_\varepsilon) = \mathbf{F}(\mathbf{v}) + \varepsilon\,\mathbf{F}(\hat{\mathbf{v}}). \tag{21}$$

This means that the rank-2 matrix

$$\mathbf{W} = \mathbf{v}\mathbf{v}^T + \varepsilon\hat{\mathbf{v}}\hat{\mathbf{v}}^T \tag{22}$$

is a solution to the problem (13) associated with the dual certificate $\mathbf{A_F}(\mathbf{v},\mathbf{M})$, and therefore $\mathbf{w}_\varepsilon \notin \mathcal{R}_\mathbf{F}(\mathbf{M})$. This contradicts the previous assumption that $\mathbf{v} \in \mathrm{int}\{\mathcal{R}(\mathbf{M})\}$. Therefore, we have $\mathbf{A_F}(\mathbf{v},\mathbf{M}) \in \mathcal{P}_0^n$, which completes the proof. ■

The next theorem states that the region $\mathcal{R}_\mathbf{F}(\mathbf{M})$ has dimension n around the nominal point $\mathbf{v}_0$.

Theorem 2: Consider the case $l = n$ and the nominal point $\mathbf{v}_0$. Let $\mathbf{M}$ be a matrix such that $\mathbf{M} \in \mathcal{P}_0^n$ and $\mathbf{v}_0 \in \mathrm{null}(\mathbf{M})$. Then, we have $\mathbf{v}_0 \in \mathrm{int}(\mathcal{R}_\mathbf{F}(\mathbf{M}))$.

Proof: The proof is omitted due to its similarity to the proof of Theorem 4. ■

Remark 1: Consider an arbitrary matrix $\mathbf{M} \in \mathcal{P}_0^n$ such that $\mathbf{v}_0 \in \mathrm{null}(\mathbf{M})$. Theorem 2 states that $\mathbf{v}_0 \in \mathrm{int}(\mathcal{R}_\mathbf{F}(\mathbf{M}))$. This concludes that $\mathbf{x}_0 \in \mathrm{int}(\mathcal{R}_\mathbf{P}(\mathbf{M}))$, meaning that $\mathcal{R}_\mathbf{P}(\mathbf{M})$ contains the nominal point $\mathbf{x}_0$ and a ball around this point. This implies that the inverse function $\mathbf{x} = \mathbf{P}^{-1}(\mathbf{z})$ exists in a neighborhood of $\mathbf{z}_0$ and can be obtained from an eigenvalue decomposition of the unique solution of the primal SDP problem.

Since there are infinitely many $\mathbf{M}$'s satisfying the conditions of Theorem 2, it is desirable to find one whose corresponding region $\mathcal{R}_\mathbf{F}(\mathbf{M})$ is large (if there exists any such matrix). To address this problem, consider an arbitrary set of vectors $\mathbf{v}_1,\ldots,\mathbf{v}_r \in \mathcal{I}_\mathbf{F}$. The next theorem explains that the problem of finding a matrix $\mathbf{M}$ such that

$$\mathbf{v}_1,\ldots,\mathbf{v}_r \in \mathrm{int}(\mathcal{R}_\mathbf{F}(\mathbf{M})) \tag{23}$$

or certifying the non-existence of such a matrix can be cast as a convex optimization.

Theorem 3: Consider the case $n = l$. Given r arbitrary points $\mathbf{v}_1,\ldots,\mathbf{v}_r \in \mathcal{I}_\mathbf{F}$, consider the problem

$$\begin{aligned} &\text{find} && \mathbf{M} \in \mathbb{S}^n && \text{(24a)}\\ &\text{subject to} && \mathbf{A_F}(\mathbf{v}_r,\mathbf{M}) \in \mathcal{P}_\varepsilon^n, \quad l = 1,2,\ldots,r && \text{(24b)}\\ &&& \mathbf{M} \in \mathcal{P}_\varepsilon^n && \text{(24c)}\\ &&& \mathbf{M}\mathbf{v}_0 = 0 && \text{(24d)} \end{aligned}$$

where $\varepsilon > 0$ is an arbitrary constant. The following statements hold:

i) The feasibility problem (24) is convex.
ii) There exists a matrix $\mathbf{M}$ satisfying the conditions given in Theorem 2 whose associated recoverable set $\mathcal{R}_\mathbf{F}(\mathbf{M})$ contains $\mathbf{v}_0,\mathbf{v}_1,\ldots,\mathbf{v}_r$ and a ball around each of these points if and only if the convex problem (24) has a solution $\mathbf{M}$.

Proof: Part (i) is implied by the fact that the sum of the two smallest eigenvalues of a matrix is a concave function and that $\mathbf{A_F}(\mathbf{v}_r,\mathbf{M})$ is a linear function with respect to $\mathbf{M}$. Part (ii) follows immediately from Theorem 1. ■

B. Over-specified Systems

The results presented in the preceding subsection can all be generalized to the case $l > n$. We will present one of these extensions below.

Theorem 4: Consider the case $l > n$. Let $\mathbf{M}$ be a matrix such that $\mathbf{M} \in \mathcal{P}_0^n$ and $\mathbf{v}_0 \in \mathrm{null}(\mathbf{M})$. The relation $\mathbf{v}_0 \in \mathrm{int}(\mathcal{R}_\mathbf{F}(\mathbf{M}))$ holds.

Proof: Let $h_1,\ldots,h_n \in \{1,\ldots,l\}$ correspond to a set of n linearly independent columns of $\nabla\mathbf{F}(\mathbf{v}_0)$. Define the function $\mathbf{H}:\mathbb{R}^n \to \mathbb{R}^n$ as

$$\mathbf{H}(\mathbf{v}) \triangleq [F_{h_1}(\mathbf{v}),\ldots,F_{h_n}(\mathbf{v})]^T \tag{25}$$

Observe that $\mathcal{R}_\mathbf{H}(\mathbf{M}) \subseteq \mathcal{R}_\mathbf{F}(\mathbf{M})$. On the other hand, since $\mathbf{M}\mathbf{v}_0 = 0$, we have

$$\mathbf{\Lambda_H}(\mathbf{v}_0,\mathbf{M}) = \mathbf{0}_n, \tag{26}$$

which concludes that

$$\mathbf{A_H}(\mathbf{v}_0,\mathbf{M}) = \mathbf{M} \in \mathcal{P}_0^n. \tag{27}$$

Therefore, it follows from Theorem 1 that $\mathbf{v}_0 \in \mathrm{int}\{\mathcal{R}_\mathbf{H}(\mathbf{M})\}$ and therefore $\mathbf{v}_0 \in \mathrm{int}\{\mathcal{R}_\mathbf{F}(\mathbf{M})\}$ ■

IV. Numerical Examples

In this section, we will provide two examples to illustrate the results of this work.

Example 1: Consider the polynomial function $\mathbf{P}(\mathbf{x}) : \mathbb{R}^4 \to \mathbb{R}^4$ defined as

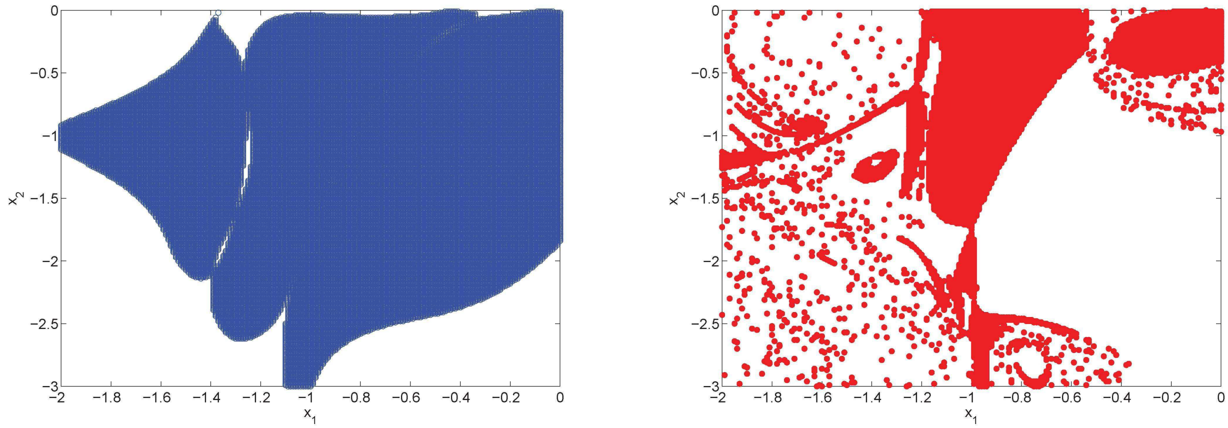
$$P_1(\mathbf{x}) = 3x_1^5x_3x_4^4 - x_2^2x_3^3 + 3x_1x_2x_3x_4^2 \tag{28}$$
$$P_2(\mathbf{x}) = 2x_1x_2x_3x_4 - 3x_1^2x_3x_4^4 + x_3^3x_4^2 \tag{29}$$
$$P_3(\mathbf{x}) = x_1^4x_2^2x_3^3 - 3x_2^2x_3^2x_4 \tag{30}$$
$$P_4(\mathbf{x}) = -2x_1^4x_4^2 + 6x_1x_2x_3^2 - x_3^2x_4^2. \tag{31}$$

The objective is to solve the feasibility problem $\mathbf{P}(\mathbf{x}) = \mathbf{z}$ for an input vector $\mathbf{z}$. Let $\mathbf{x}_0 = [-1 \ \ -1 \ \ -2 \ \ 1]^T$ be the nominal point (a guess for the solution). To solve the equation $\mathbf{P}(\mathbf{x}) = \mathbf{z}$, we use the primal SDP problem, where the matrix $\mathbf{M}$ in its objective is designed based on Theorem 2 (by solving a simple convex optimization problem to pick one matrix out of infinitely many candidates). Consider the region $\mathcal{R}_\mathbf{P}(\mathbf{M})$, which is the set of all points $\mathbf{x}$ that can be recovered by the primal SDP problem. To be able to visualize this recoverable region in a 2-dimensional space, consider the restriction of $\mathcal{R}_\mathbf{P}(\mathbf{M})$ to the subspace $\{\mathbf{x} \,|x_3 = -2, x_4 = 1\}$ (these numbers are borrowed from the third and fourth entries of the nominal point $\mathbf{x}_0$). The region $\mathcal{R}_\mathbf{P}(\mathbf{M})$ after the above restriction can be described as

$$\left\{(x_1,x_2) \;\middle|\; [x_1 \ x_2 \ -2 \ 1]^T \in \mathcal{R}_\mathbf{P}(\mathbf{M})\right\} \tag{32}$$

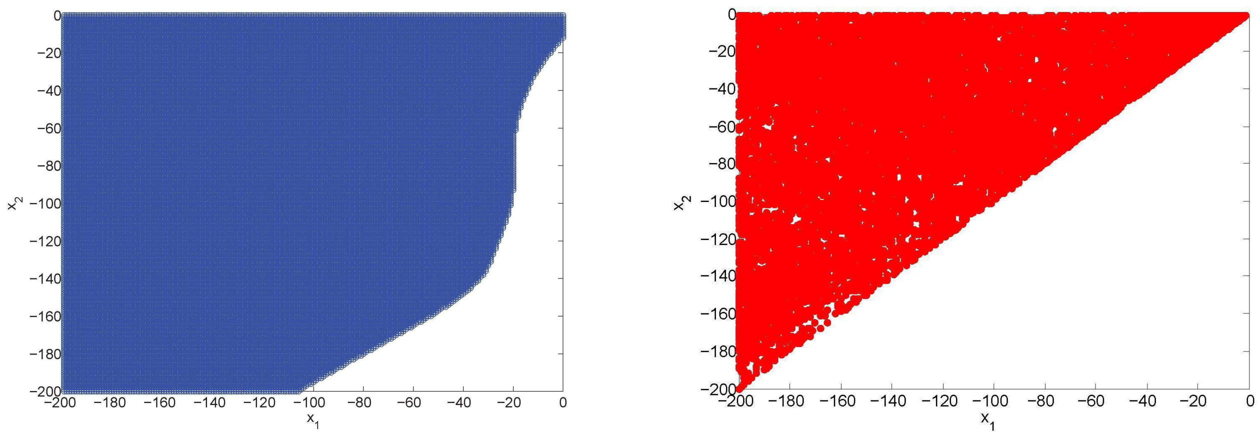
This set is drawn in Figure 1(a) over the box $[-2,\ 0] \times [-3,\ 0]$. It can be seen that $(-1,-1)$ is an interior point of this set, as expected from Theorem 2. Note that if $\mathbf{M}$ is chosen as I (identity matrix), as prescribed by the compressed sensing literature [19], its corresponding recoverable region restricted to the subspace $\{\mathbf{x} \mid x_3 = -2, x_4 = 1\}$ becomes empty (i.e., the SDP problem never works).

To compare the proposed technique with Newton's method, consider the feasibility problem $\mathbf{P}(\mathbf{x}) = \mathbf{z}$ for a given input vector $\mathbf{z}$. To find $\mathbf{x}$, we can use the nominal



(a) (b)

Fig. 1: Example 1: (a) recoverable region using the primal SDP problem, (b) recoverable region using Newton's method.



(a) (b)

Fig. 2: Example 2: (a) recoverable region using the primal SDP problem with $\mathbf{M}$ designed based on the reference point, (b) recoverable region using the primal SDP problem with $\mathbf{M} = I$.

point $\mathbf{x}_0$ for the initialization of Newton's method. Consider the set of all points $\mathbf{x}$ that can be found using Newton's method for some input vector $\mathbf{z}$. The restriction of this set to the subspace $\{\mathbf{x}\,|x_3 = -2, x_4 = 1\}$ is plotted in Figure 1(b). This region is complicated and has some fractal properties, which is a known feature of Newton's method [13], [14]. The superiority of the proposed SDP problem over Newton's method can be seen in Figures 1(a) and (b).

Example 2: The objective of this example is to compare the proposed technique for designing $\mathbf{M}$ with the commonly-used choice $\mathbf{M} = I$ in the context of the matrix completion problem [19], [24]. Consider a rank-1 positive semidefinite 20×20 matrix in the form of $\mathbf{x}\mathbf{x}^T$, where $\mathbf{x}$ is an unknown vector. Assume that the entries of this matrix are known at 30 fixed locations, and the aim is to recover the vector $\mathbf{x}$. Similarly to Example 1, we restrict the recoverable region to a 2-dimensional space to be able to visualize the set. To do so, assume that the last 18 entries of $\mathbf{x}$ are restricted to certain fixed numbers, while the first two entries can take any values in the finite grid $\{-200, -199, ..., 0\} \times \{-200, -199, ..., 0\}$. We design a matrix $\mathbf{M}$ using the reference point $(-100, -100)$ in the above grid (by solving a simple convex optimization problem based on Theorem 3). The region $\mathcal{R}_{\mathbf{P}}(\mathbf{M})$ after the aforementioned restriction is drawn in Figure 2(a). This means that the SDP problem finds the right solution of the matrix completion problem if the unknown solution belongs to this region. The corresponding region for $\mathbf{M} = I$ is depicted Figure 2(b), which is a subset of the region given in Figure 2(a).

V. CONCLUSIONS

Consider an arbitrary polynomial function together with a nominal point, and assume that the Jacobian of the function is invertible at this point. Due to the inverse function theorem, the polynomial function is locally invertible around the nominal point. This paper proposes a convex optimization method to find the inverse function. In particular, infinitely many semidefinite programs (SDPs) are proposed such that each SDP finds the inverse function over a neighborhood of the nominal point. The problem of designing an SDP that obtains the inverse function over a large region is also studied and cast as convex optimization. One main application of this work is in solving systems of polynomial equations. The benefits of the proposed approach over Newton's method and the nuclear-norm technique are numerically demonstrated.

REFERENCES

[1] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[2] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[3] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM (JACM)*, vol. 42, no. 6, pp. 1115–1145, 1995.

[4] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92–107, 2012.

[5] A. Kalbat, R. Madani, G. Fazelnia, and J. Lavaei, "Efficient convex relaxation for stochastic optimal distributed control problem," *Communication, Control, and Computing, IEEE 52nd Annual Allerton Conference on*, 2014.

[6] S. Sojoudi and J. Lavaei, "Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure," *SIAM Journal on Optimization*, vol. 24, no. 4, pp. 1746–1778, 2014.

[7] Y. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization Methods and Software*, vol. 9, pp. 141–160, 1998.

[8] S. Zhang and Y. Huang, "Complex quadratic optimization and semidefinite programming," *SIAM Journal on Optimization*, vol. 87, pp. 871–890, 2006.

[9] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[10] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, "New developments in sum of squares optimization and SOSTOOLS," in *Proceedings of the American Control Conference*, 2004, pp. 5606–5611.

[11] G. Chesi, "LMI techniques for optimization over polynomials in control: a survey," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2500–2510, 2010.

[12] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[13] H.-O. Peitgen, P. H. Richter, and H. Schultens, *The beauty of fractals: images of complex dynamical systems, Chapter 6*. Springer Berlin, 1986.

[14] A. Douady, "Julia sets and the Mandelbrot set," in *The Beauty of Fractals*. Springer Berlin, 1986, pp. 161–174.

[15] Y. Nesterov, A. Nemirovskii, and Y. Ye, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994, vol. 13.

[16] K. G. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Mathematical programming*, vol. 39, no. 2, pp. 117–129, 1987.

[17] P. M. Pardalos and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *Journal of Global Optimization*, vol. 1, no. 1, pp. 15–22, 1991.

[18] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[19] E. J. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, Aug. 2013.

[20] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 199–225, 2013.

[21] A. Boukerche, H. Oliveira, E. F. Nakamura, and A. A. Loureiro, "Localization systems for wireless sensor networks," *IEEE wireless Communications*, vol. 14, no. 6, pp. 6–12, 2007.

[22] J. Li, J. Conan, and S. Pierre, "Mobile terminal location for MIMO communication systems," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 8, pp. 2417–2420, 2007.

[23] S. D. Santos, P. J. Verveer, and P. I. Bastiaens, "Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate," *Nature cell biology*, vol. 9, no. 3, pp. 324–330, 2007.

[24] Y. Shechtman, A. Beck, and Y. C. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 928–938, 2014.

APPENDIX

To convert an arbitrary polynomial feasibility problem to a quadratic feasibility problem, we perform two operations as many times as necessary:

- Select an arbitrary monomial in the polynomial function.
- Replace all occurrences of that monomial with a new slack variable.
- Impose an additional constraint stating that the removed monomial is equal to the new slack variable.

The above operations preserve the full rank property of the Jacobian matrix. To prove this by induction, consider an arbitrary monomial and denote it as $P_{q+1}(\mathbf{x})$. Consider the following two operations:

i) Replace every occurrence of $P_{q+1}(\mathbf{x})$ in $P_1(\mathbf{x}),\dots,P_q(\mathbf{x})$ with x_{m+1}.

ii) Add a new slack variable x_{m+1} to the problem together with the constraint

$$P_{q+1}(\mathbf{x}) - x_{m+1} = 0,$$

Now, the problem (1) can be reformulated as

$$\begin{aligned} &\text{find} && \bar{\mathbf{x}} \in \mathbb{R}^{m+1} && \text{(33a)}\\ &\text{subject to} && Q_i(\bar{\mathbf{x}}) = 0, \qquad i = 1,\dots,q+1, && \text{(33b)} \end{aligned}$$

where the polynomials $Q_1,\dots,Q_{q+1} : \mathbb{R}^{m+1} \to \mathbb{R}$ satisfy the equations:

$$\begin{aligned} Q_i\left(\mathbf{x}, P_{q+1}(\mathbf{x})\right) &= P_i(\mathbf{x}) - z_i \quad \text{for} \quad i = 1,\dots,q\\ Q_{q+1}\left(\mathbf{x}, P_{q+1}(\mathbf{x})\right) &= P_{q+1}(\mathbf{x}) - x_{m+1}. \end{aligned}$$

Now, we have

$$\mathbf{P}(\mathbf{x}) = \begin{bmatrix} P_1(\mathbf{x}) \\ \vdots \\ P_q(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} Q_1(\mathbf{x}, P_{q+1}(\mathbf{x})) - z_1 \\ \vdots \\ Q_q(\mathbf{x}, P_{q+1}(\mathbf{x})) - z_q \end{bmatrix}.$$

Define $\mathbf{Q} : \mathbb{R}^{m+1} \to \mathbb{R}^{q+1}$ as

$$\mathbf{Q}(\bar{\mathbf{x}}) \triangleq [Q_1(\bar{\mathbf{x}}), Q_2(\bar{\mathbf{x}}),\dots,Q_{q+1}(\bar{\mathbf{x}})]^T. \tag{34}$$

Theorem 5: The following statements hold:

i) When the number of equations is equal to the number of unknowns (i.e., $q = m$), we have

$$|\det\{\nabla \mathbf{P}\}| = |\det\{\nabla \mathbf{Q}\}| \tag{35}$$

ii) In general, the matrix $\nabla \mathbf{P}$ has full row rank if and only if $\nabla \mathbf{Q}$ has full row rank.

Proof: To prove Part (i), define $s : \mathbb{R}^m \to \mathbb{R}^{m+1}$ as follows

$$s(\mathbf{x}) \triangleq (\mathbf{x}, P_{q+1}(\mathbf{x})). \tag{36}$$

Also, for simplicity, define the following operators:

$$\begin{aligned} \nabla^+ &\triangleq \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}, \frac{\partial}{\partial x_{m+1}}\right), \\ \nabla^- &\triangleq \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right). \end{aligned} \tag{37}$$

Using the chain rule, we get

$$\frac{\partial P_i(\mathbf{x})}{\partial \mathbf{x}_j} = \frac{\partial Q_i(\mathbf{s}(\mathbf{x}))}{\partial \mathbf{x}_j} + \frac{\partial Q_i(\mathbf{s}(\mathbf{x}))}{\partial \mathbf{x}_{m+1}} \frac{\partial P_{q+1}(\mathbf{x})}{\partial \mathbf{x}_j}, \tag{38}$$

which could be written as

$$\frac{\partial P_i(\mathbf{x})}{\partial \mathbf{x}_j} = \frac{\partial Q_i(\bar{\mathbf{x}})}{\partial \mathbf{x}_j} + \frac{\partial Q_i(\bar{\mathbf{x}})}{\partial \mathbf{x}_{m+1}} \frac{\partial P_{q+1}(\mathbf{x})}{\partial \mathbf{x}_j}. \tag{39}$$

Consequently,

$$\nabla^- P_i(\mathbf{x}) = \nabla^+ Q_i(\bar{\mathbf{x}}) \times \begin{bmatrix} I \\ \nabla^- P_{q+1}(\mathbf{x}) \end{bmatrix}. \tag{40}$$

Therefore,

$$(\nabla \mathbf{P})^T = \begin{bmatrix} \nabla^+ Q_1(\bar{\mathbf{x}}) \\ \vdots \\ \nabla^+ Q_q(\bar{\mathbf{x}}) \end{bmatrix} \begin{bmatrix} I \\ \nabla^- P_{q+1}(\mathbf{x}) \end{bmatrix}. \tag{41}$$

On the other hand,

$$\begin{bmatrix} \nabla^+ Q_1(\bar{\mathbf{x}}) \\ \vdots \\ \nabla^+ Q_q(\bar{\mathbf{x}}) \end{bmatrix} = \begin{bmatrix} \nabla^- Q_1(\bar{\mathbf{x}}) & \frac{\partial Q_1(\bar{\mathbf{x}})}{\partial x_{m+1}} \\ \vdots & \vdots \\ \nabla^- Q_q(\bar{\mathbf{x}}) & \frac{\partial Q_q(\bar{\mathbf{x}})}{\partial x_{m+1}} \end{bmatrix}, \tag{42}$$

Hence,

$$\begin{aligned} (\nabla \mathbf{P})^T &= \begin{bmatrix} \nabla^+ Q_1(\bar{\mathbf{x}}) \\ \vdots \\ \nabla^+ Q_q(\bar{\mathbf{x}}) \end{bmatrix} \begin{bmatrix} I \\ \nabla^- P_{q+1}(\mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} \nabla^- Q_1(\bar{\mathbf{x}}) \\ \vdots \\ \nabla^- Q_q(\bar{\mathbf{x}}) \end{bmatrix} + \\ &\quad \begin{bmatrix} \frac{\partial Q_1(\bar{\mathbf{x}})}{\partial x_{m+1}} \\ \vdots \\ \frac{\partial Q_q(\bar{\mathbf{x}})}{\partial x_{m+1}} \end{bmatrix} \nabla^- P_{q+1}(\mathbf{x}) \; . \end{aligned} \tag{43}$$

Now, consider the Jacobian of $\mathbf{Q}$:

$$\begin{aligned} (\nabla \mathbf{Q})^T &= \begin{bmatrix} \nabla^+ Q_1(\bar{\mathbf{x}}) \\ \vdots \\ \nabla^+ Q_q(\bar{\mathbf{x}}) \\ \nabla^+ Q_{q+1}(\bar{\mathbf{x}}) \end{bmatrix} \\ &= \left[\begin{array}{c|c} \nabla^- Q_1(\bar{\mathbf{x}}) & \frac{\partial Q_1(\bar{\mathbf{x}})}{\partial x_{m+1}} \\ \vdots & \vdots \\ \nabla^- Q_q(\bar{\mathbf{x}}) & \frac{\partial Q_q(\bar{\mathbf{x}})}{\partial x_{m+1}} \\ \hline \nabla^- P_{q+1}(\mathbf{x}) & -1 \end{array}\right]. \end{aligned} \tag{44}$$

According to (43) and (44) and using Schur's complement, we conclude that

$$|\det\{\nabla \mathbf{P}\}| = |\det\{\nabla \mathbf{Q}\}| \tag{45}$$

Part (ii) can be proven similarly. ■