2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Multivariate Polynomial Optimization in Complex Variables Is a (Rectangular) Multiparameter Eigenvalue Problem

Christof Vermeersch, Sibren Lagauw, and Bart De Moor, *Fellow*, *IEEE & IFAC & SIAM*

***Abstract*— We extend the relation between univariate polynomial optimization in one complex variable and the polynomial eigenvalue problem to the multivariate case. The first-order necessary conditions for optimality of the multivariate polynomial optimization problem, which are computed using Wirtinger derivatives, constitute a system of multivariate polynomial equations in the complex variables and their complex conjugates. Wirtinger calculus provides an elegant way to differentiate real-valued (cost) functions in complex variables. An elimination of the complex conjugate variables, via the Macaulay matrix, results in a (rectangular) multiparameter eigenvalue problem, (some of) the eigentuples of which correspond to the stationary points of the original real-valued cost function. We illustrate our novel globally optimal optimization approach with several (didactical) examples.**

I. Introduction

Complex-valued signals arise in many areas of science and engineering, like communications, systems theory, oceanography, geophysics, optics, and electromagnetics [1], [2]. Especially in signal processing, one often encounters (nonlinear) functions in complex variables [3], for example, transfer functions of linear time-invariant models. An important issue when working with complex-valued signals and complex variables is related to (nonlinear) optimization. Most of the optimization literature deals with real variables only, seemingly suggesting that complex variables are not encountered in practice. However, optimization problems in complex variables appear in various applications, some of which very relevant to the systems and control community, e.g., filter design [2], [3], [4], [5], system identification [6], blind source separation [2], tensor decomposition [1], [7], [8], parameter estimation [2], and nonlinear electrical circuit simulation [1].

Cost functions of optimization problems in complex variables are real-valued: it makes no sense to optimize a complex-valued cost function, because the field of complex numbers is not (totally) ordered. So, from an application point of view, these real-valued functions are exactly the kind of cost functions that we expect to encounter. However, real-valued cost functions in complex variables are necessarily non-holomorphic (i.e., the complex generalization of non-analytic) [1]. They have no complex derivatives. An optimization problem in complex variables is typically tackled by reformulating the cost function as a function of the real and imaginary parts of the complex variables, so that standard real optimization techniques can be used. Wirtinger calculus provides a more elegant solution by relaxing the definition of differentiability and defining a general framework that includes holomorphic functions as a special case [2], [3], [4]. The development of Wirtinger calculus[1] by the Austrian mathematician Wilhelm Wirtinger dates back to 1927 [10]. It was rediscovered in 1983, without any reference to Wirtinger, by David Brandwood [5]. The advantage of Wirtinger calculus is that the expressions do not become unnecessarily complicated and the derivations are rather similar to the real situation.

Sorber et al. [7], [8] have highlighted an interesting relation between univariate polynomial optimization in one complex variable and the polynomial eigenvalue problem. The first-order necessary conditions for optimality of the real-valued univariate polynomial cost function obtained via Wirtinger calculus yield a system of two polyanalytic polynomials in the complex variable and its complex conjugate. An elimination of this complex conjugate variable, via the Sylvester matrix, results in a polynomial eigenvalue problem that can be solved with standard techniques from numerical linear algebra. In this paper, we extend this relation to the multivariate case: we show that multivariate polynomial optimization in multiple complex variables is a (rectangular[2]) multiparameter eigenvalue problem, of which at least one of the eigentuples corresponds to the global minimizer. This paper has not the ambition to provide a competitive alternative with respect to the current state-of-the-art. Rather it serves a tutorial purpose, presenting a novel optimization approach in a didactical way. It highlights several interesting research avenues initiated by this reformulation, while omitting technical derivations.

Note that reformulating a multivariate polynomial opti-

Christof Vermeersch (corresponding author), Sibren Lagauw, and Bart De Moor are with the Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS), Dept. of Electrical Engineering (ESAT), KU Leuven. Kasteelpark Arenberg 10, 3001 Leuven, Belgium. E-mail addresses: {christof.vermeersch, sibren.lagauw, bart.demoor}@esat.kuleuven.be.

This work was supported in part by the KU Leuven: Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117, and C3I-21-00316), Industrial Research Fund (fellowships 13-0260, IOFm/16/004, and IOFm/20/002), and LRD bilateral industrial projects; in part by Flemish Government agencies: FWO (EOS project G0F6718N (SeLMA), SBO project S005319, infrastructure project I013218N, and TBM project T001919N), EWI (Flanders AI Research Program), and VLAIO (Baekeland PhD mandate HBC.2019.2204 and HBC.2021.0076); and in part by the European Commission (ERC Advanced Grant under grant 885682). The work of Christof Vermeersch and Sibren Lagauw was supported by FWO fellowships (under grants SB/1SA1319N and FR/11K5623N, respectively).

[1]The idea of using Wirtinger derivatives can be traced back to at least 1899, with Henri Poincaré [1], [9]. The name *Wirtinger calculus* is especially present in the German literature, while in other sources one often reads $\mathbb{CR}$-*calculus*, referring to the fields $\mathbb{C}$ and $\mathbb{R}$ [4].

[2]Since we only consider the rectangular multiparameter eigenvalue problem, we no longer mention the qualification *rectangular* in the remainder of this paper.

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

mization problem in real variables as an (one-parameter) eigenvalue problem is a well-established methodology and there exist several techniques in the optimization literature to obtain such an eigenvalue problem. For example, in the scope of the moment hierarchy approach, the extraction of global minimizers reduces to eigenvalue computations [11], [12], [13], [14]. Furthermore, the use of the first-order necessary conditions for optimality (i.e., the gradient ideal) for polynomial optimization has thoroughly been investigated in, among others, [15], [16], [17], and can be combined with a multivariate polynomial system solving approach that resorts on eigenvalue computations [18], [19]. A reformulation of the cost function as a function of the real and imaginary parts of the complex variables makes it possible to also use these (efficient) numerical optimization techniques in a complex setting. In this paper, however, we stick to an optimization approach over the complex variables, which avoids doubling the number of variables at the cost of solving a more difficult multiparameter eigenvalue problem.

Notation and preliminaries: We denote scalars by lowercase letters, e.g., a, and tuples/vectors by boldface lowercase letters, e.g., $\boldsymbol{a}$. Matrices are characterized by boldface uppercase letters, e.g., $\boldsymbol{A}$. When a matrix contains one or more parameters, we use a bold calligraphic font, e.g., $\boldsymbol{\mathcal{A}}(a)$ with parameter a. We use j to denote the imaginary unit $\sqrt{-1}$. The complex conjugate, transpose, and Hermitian transpose of $\boldsymbol{a}$ are indicated by $\bar{\boldsymbol{a}}$, $\boldsymbol{a}^{\mathrm{T}}$, and $\boldsymbol{a}^{\mathrm{H}}$, respectively. $\|\cdot\|_2$ is the 2-norm and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius-norm.

Outline and contribution: The remainder of this paper is organized as follows: In Section II, we define the multivariate polynomial optimization problem. Next, in Section III, we look at the implications of working with real-valued cost functions in complex variables and give a brief introduction to Wirtinger calculus. The reformulation of the multivariate polynomial minimization problem in complex variables as a multiparameter eigenvalue problem is the **main contribution of this paper** and can be found in Section IV. Section V deals with so-called *ghost solutions*. Since this paper serves mainly a tutorial purpose, it also identifies some open research questions; a non-exhaustive overview of which is given in Section VI. Finally, we conclude this paper and point to ideas for future work in Section VII.

II. Problem definition

In this paper, we deal with real-valued (multivariate) polynomial cost functions $f(\boldsymbol{z}, \bar{\boldsymbol{z}})$ in n complex (decision) variables $\boldsymbol{z} \in \mathbb{C}^n$ and their complex conjugates $\bar{\boldsymbol{z}} \in \mathbb{C}^n$,

$$f : \mathbb{C}^n \rightarrow \mathbb{R} : \boldsymbol{z} \mapsto f(\boldsymbol{z}, \bar{\boldsymbol{z}}), \tag{1}$$

where we express the dependency of the cost function on the complex variables $\boldsymbol{z}$ and their complex conjugates $\bar{\boldsymbol{z}}$ explicitly to highlight that the polynomial is real-valued (Section III). We consider, primarily, the unconstrained minimization problem, i.e.,

$$\min_{\boldsymbol{z}} f(\boldsymbol{z}, \bar{\boldsymbol{z}}), \tag{2}$$

but adaptations to maximization or constrained optimization via the Lagrangian are straightforward (Example 1). A prototypical problem with a real-valued polynomial cost function is the minimization of the squared Frobenius-norm of a matrix polynomial $\boldsymbol{\mathcal{F}}(\boldsymbol{z}, \bar{\boldsymbol{z}})$,

$$\boldsymbol{\mathcal{F}} : \mathbb{C}^n \rightarrow \mathbb{C}^{m_1 \times m_2} : \boldsymbol{z} \mapsto \boldsymbol{\mathcal{F}}(\boldsymbol{z}, \bar{\boldsymbol{z}}),$$

that maps n complex variables $\boldsymbol{z}$ and their complex conjugates $\bar{\boldsymbol{z}}$ onto $m_1 m_2$ function values, i.e.,

$$\min_{\boldsymbol{z}} \|\boldsymbol{\mathcal{F}}(\boldsymbol{z}, \bar{\boldsymbol{z}})\|_{\mathrm{F}}^2, \tag{3}$$

which is also known as the complex nonlinear least-squares optimization problem. Because of the imposed norm, the cost function in (3) is a real-valued polynomial in $\boldsymbol{z}$ and $\bar{\boldsymbol{z}}$.

III. Wirtinger derivatives

Before we tackle (2), we need to take a closer look at the implications of differentiation in the complex domain. Consider a multivariate complex-valued function $f(\boldsymbol{z})$ in n complex variables $\boldsymbol{z} \in \mathbb{C}^n$,

$$f : \mathbb{C}^n \rightarrow \mathbb{C} : \boldsymbol{z} = \boldsymbol{x} + \mathrm{j}\boldsymbol{y} \mapsto f(\boldsymbol{z}) = u(\boldsymbol{x}, \boldsymbol{y}) + \mathrm{j}v(\boldsymbol{x}, \boldsymbol{y}),$$

where $u(\boldsymbol{x}, \boldsymbol{y})$ and $v(\boldsymbol{x}, \boldsymbol{y})$ are ordinary real-valued functions in $2n$ multivariate real variables $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^n$. The transformation from $(\boldsymbol{z}, \bar{\boldsymbol{z}})$ to $(\boldsymbol{x}, \boldsymbol{y})$ is a simple change of variables for two independent vector variables,

$$\boldsymbol{x} = \frac{\boldsymbol{z} + \bar{\boldsymbol{z}}}{2}, \quad \boldsymbol{y} = \frac{\boldsymbol{z} - \bar{\boldsymbol{z}}}{2\mathrm{j}}, \tag{4}$$

and, vice versa,

$$\boldsymbol{z} = \boldsymbol{x} + \mathrm{j}\boldsymbol{y}, \quad \bar{\boldsymbol{z}} = \boldsymbol{x} - \mathrm{j}\boldsymbol{y}. \tag{5}$$

The complex-valued function is said to be *differentiable* at a point $\boldsymbol{z}_0 \in \mathbb{C}^n$ if the complex-valued limit operation

$$\lim_{\Delta \boldsymbol{z} \rightarrow \boldsymbol{0}} \frac{f(\boldsymbol{z}_0 + \Delta \boldsymbol{z}) - f(\boldsymbol{z}_0)}{\Delta \boldsymbol{z}} \tag{6}$$

exists, i.e., when the limit value is independent of the direction in which $\Delta \boldsymbol{z}$ approaches zero. For example, the result of the limit should be the same when $\Delta \boldsymbol{z}$ approaches zero on the real axis ($\Delta \boldsymbol{x} \rightarrow \boldsymbol{0}$) or on the imaginary axis ($\Delta \boldsymbol{y} \rightarrow \boldsymbol{0}$). This requirement is formalized in the *Cauchy–Riemann conditions* [2], [3] for differentiability at $\boldsymbol{z}_0 = \boldsymbol{x}_0 + \mathrm{j}\boldsymbol{y}_0$:

$$\begin{aligned} \frac{\partial u(\boldsymbol{x}_0, \boldsymbol{y}_0)}{\partial \boldsymbol{x}} &= \frac{\partial v(\boldsymbol{x}_0, \boldsymbol{y}_0)}{\partial \boldsymbol{y}}, \\ -\frac{\partial v(\boldsymbol{x}_0, \boldsymbol{y}_0)}{\partial \boldsymbol{x}} &= \frac{\partial u(\boldsymbol{x}_0, \boldsymbol{y}_0)}{\partial \boldsymbol{y}}. \end{aligned} \tag{7}$$

The Cauchy–Riemann conditions (7) are necessary and sufficient conditions[3] for the existence of the limit defining the complex differentiation operation in (6). A multivariate function in complex variables is said to be *holomorphic* in a domain (i.e., the complex generalization of analytic), if the function is differentiable for all points in that domain.

[3]The Cauchy–Riemann conditions are necessary and sufficient only for continuous functions $u(\boldsymbol{x}, \boldsymbol{y})$ and $v(\boldsymbol{x}, \boldsymbol{y})$, see [2] for more information.

Real-valued functions are, however, non-holomorphic. It is easy to see that the Cauchy–Riemann conditions (7) do not hold, except for the constant real-valued polynomial, because $v(\boldsymbol{x},\boldsymbol{y}) \equiv 0$. In other words, there exists no Taylor series in $\boldsymbol{z}$ of $f(\boldsymbol{z},\bar{\boldsymbol{z}})$ at $\boldsymbol{z}_0$ so that the series converges to $f(\boldsymbol{z},\bar{\boldsymbol{z}})$ in a neighborhood of $\boldsymbol{z}_0$ [1].

Wirtinger calculus provides a general framework for differentiating non-holomorphic functions; it is general in the sense that it includes holomorphic functions as a special case. It only requires that $f(\boldsymbol{z},\bar{\boldsymbol{z}})$ or $f(\boldsymbol{z})$ is *real differentiable*: if $u(\boldsymbol{x},\boldsymbol{y})$ and $v(\boldsymbol{x},\boldsymbol{y})$ have continuous partial derivatives with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$, then the function is real differentiable [2]. The idea in Wirtinger calculus is to differentiate functions of the form $f(\boldsymbol{z},\bar{\boldsymbol{z}})$ by considering the partial derivatives with respect to the complex variables $\boldsymbol{z}$ and their complex conjugates $\bar{\boldsymbol{z}}$, which can be formally written as

$$\begin{aligned}\frac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \boldsymbol{z}} &= \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{z}} + \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{y}}\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{z}} \\ &= \frac{1}{2}\left(\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{x}} - \mathrm{j}\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{y}}\right), \\ \frac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \bar{\boldsymbol{z}}} &= \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \bar{\boldsymbol{z}}} + \frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{y}}\frac{\partial \boldsymbol{y}}{\partial \bar{\boldsymbol{z}}} \\ &= \frac{1}{2}\left(\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{x}} + \mathrm{j}\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{y}}\right).\end{aligned}$$

We call $\frac{\partial(\cdot)}{\partial \boldsymbol{z}}$ and $\frac{\partial(\cdot)}{\partial \bar{\boldsymbol{z}}}$ the *cogradient operator* and *conjugate cogradient operator*, respectively. They act as a partial derivative with respect to $\boldsymbol{z}$ (or $\bar{\boldsymbol{z}}$), while treating $\bar{\boldsymbol{z}}$ (or $\boldsymbol{z}$) as a constant vector. Note that, for a complex-valued cost function that satisfies the Cauchy–Riemann conditions (7), $\frac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \bar{\boldsymbol{z}}}$ is equal to zero [3]. Hence, differentiability in a complex domain requires the function $f(\boldsymbol{z},\bar{\boldsymbol{z}})$ to be solely a function of $\boldsymbol{z}$ and not exhibit any dependency on $\bar{\boldsymbol{z}}$. This is also the reason why we explicitly write real-valued functions in terms of $\boldsymbol{z}$ and $\bar{\boldsymbol{z}}$. For a real-valued function $f(\boldsymbol{z},\bar{\boldsymbol{z}})$, it holds that

$$\overline{\left(\frac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \boldsymbol{z}}\right)} = \frac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \bar{\boldsymbol{z}}}. \tag{8}$$

Although their definitions allow the cogradient and conjugate cogradient to be expressed elegantly in terms of $\boldsymbol{z}$ and $\bar{\boldsymbol{z}}$, neither contains enough information by itself to express the change in a function with respect to a change in $\boldsymbol{z}$ or $\bar{\boldsymbol{z}}$ as independent variables. Therefore, we define the *complex gradient operator* $\nabla(\cdot)$ as

$$\nabla(\cdot) = \left(\frac{\partial(\cdot)}{\partial \boldsymbol{z}}, \frac{\partial(\cdot)}{\partial \bar{\boldsymbol{z}}}\right).$$

Relation (8) between both cogradients, however, allows us to only compute one cogradient and obtain the other one by simply taking the complex conjugate of that expression.

For the real-valued multivariate polynomial cost functions in complex variables in (1), a complex derivative does not exist, but Wirtinger calculus provides an elegant alternative framework to compute the first-order necessary conditions for optimality:

$$\begin{cases} p_i(\boldsymbol{z},\bar{\boldsymbol{z}}) = \dfrac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial z_i} = 0, & \text{for } i = 1,\ldots,n, \\ p_i(\boldsymbol{z},\bar{\boldsymbol{z}}) = \dfrac{\partial f(\boldsymbol{z},\bar{\boldsymbol{z}})}{\partial \bar{z}_{i-n}} = 0, & \text{for } i = n+1,\ldots,2n. \end{cases} \tag{9}$$

The common roots $(\boldsymbol{z}_0,\bar{\boldsymbol{z}}_0)$ of this square system of $2n$ multivariate polynomial equations in the complex variables $\boldsymbol{z}$ and their complex conjugates $\bar{\boldsymbol{z}}$ correspond to the stationary points of (1):

$$\mathcal{V}_{\mathbb{C}} = \{\boldsymbol{z}_0 \in \mathbb{C}^n : p_i(\boldsymbol{z}_0,\bar{\boldsymbol{z}}_0) = 0, \forall i = 1,\ldots,2n\}. \tag{10}$$

Notice that the polynomials in (9) are not necessarily real-valued, only the original cost function is. We can illustrate all the above with a didactical example.

Example 1: Let us consider the optimization problem:

$$\min_{z} -\mathrm{j}\big(z^3 + z^2\bar{z} - z\bar{z}^2 - \bar{z}^3\big)$$
$$\text{subject to } \|z\|_2^2 - 3 = 0,$$

which amounts to minimizing the real-valued polynomial cost function $f(z,\bar{z}) = -\mathrm{j}\big(z^3 + z^2\bar{z} - z\bar{z}^2 - \bar{z}^3\big) = 8x^2y$, where $x = \Re(z)$ and $y = \Im(z)$, on a circle with radius $\sqrt{3}$. We could approach this constrained optimization problem from the traditional point of view, via (4), and consider x and y as two independent real variables:

$$\min_{x,y} 8x^2y$$
$$\text{subject to } x^2 + y^2 - 3 = 0.$$

However, since the cost function is real-valued, we can alternatively use Wirtinger derivatives, such that we keep using complex variables. The Lagrangian that corresponds to this optimization problem is

$$\mathcal{L}(z,\bar{z},\lambda) = -\mathrm{j}\big(z^3 + z^2\bar{z} - z\bar{z}^2 - \bar{z}^3\big) + \lambda(z\bar{z} - 3)$$

and its first-order necessary conditions for optimality are

$$\begin{cases} \dfrac{\partial \mathcal{L}(z,\bar{z},\lambda)}{\partial z} = -\mathrm{j}\big(3z^2 + 2z\bar{z} - \bar{z}^2\big) + \lambda\bar{z} = 0, \\ \dfrac{\partial \mathcal{L}(z,\bar{z},\lambda)}{\partial \bar{z}} = -\mathrm{j}\big(z^2 - 2z\bar{z} - 3\bar{z}^2\big) + \lambda z = 0, \\ \dfrac{\partial \mathcal{L}(z,\bar{z},\lambda)}{\partial \lambda} = z\bar{z} - 3 = 0. \end{cases} \tag{11}$$

When we solve this system of multivariate polynomial equations, we obtain six stationary points: two global maximizers $\pm\sqrt{2} + 1\mathrm{j}$, two global minimizers $\pm\sqrt{2} - 1\mathrm{j}$, one local minimizer $\sqrt{3}\mathrm{j}$, and one local maximizer $-\sqrt{3}j$ (subject to the constraints). This agrees with the visually identified stationary points in Fig. 1. Notice that the first and second equation in (11) are complex conjugates of each other and that they are clearly not real-valued.

Remark 1: In the real case ($\boldsymbol{z} = \boldsymbol{x}$), (9) corresponds to the well-known real gradient set equal to zero. Suppose that we are only interested in the real stationary points $\boldsymbol{x}_0$ of the real-valued cost function $f(\boldsymbol{z},\bar{\boldsymbol{z}})$, then we need to consider only the real gradient [7], given by

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\frac{\partial f(\boldsymbol{z})}{\partial \boldsymbol{z}}\bigg|_{\boldsymbol{z}=\boldsymbol{x}} = 2\frac{\partial f(\boldsymbol{z})}{\partial \bar{\boldsymbol{z}}}\bigg|_{\boldsymbol{z}=\boldsymbol{x}}.$$

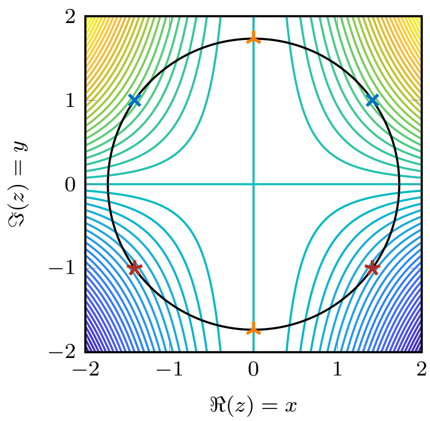

Fig. 1. Contour lines of the real-valued polynomial cost function $f(z, \bar{z})$ in Example 1: the optimization problem has two global minimizers (★), two global maximizers (×), and two local optima (⅄), subject to (—).

IV. Multiparameter eigenvalue problem

The fact that both the complex (decision) variables $\boldsymbol{z}$ and their complex conjugates $\bar{\boldsymbol{z}}$ are present in (9) clearly creates redundancy. After all, solving a system of multivariate polynomial equations is not an easy task at hand [20]. In this section, we show that the stationary points of (1) correspond to (some of) the eigentuples of a multiparameter eigenvalue problem (MEP), by eliminating $\bar{\boldsymbol{z}}$ via the Macaulay matrix.

Firstly, we rewrite every polynomial in (9) in terms of the different complex conjugate monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha}}$:

$$p_i(\boldsymbol{z}, \bar{\boldsymbol{z}}) = \sum_{\{\boldsymbol{\alpha}\}} p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z}) \bar{\boldsymbol{z}}^{\boldsymbol{\alpha}},$$

for $i = 1, \ldots, 2n$, where the summation runs over all multi-indices $\boldsymbol{\alpha}$. The multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ labels the powers of the conjugate variables $\bar{\boldsymbol{z}}$ in the monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha}} = \prod_{k=1}^{n} \bar{z}_k^{\alpha_k} = \bar{z}_1^{\alpha_1} \cdots \bar{z}_n^{\alpha_n}$. The total degree of a monomial with respect to $\bar{\boldsymbol{z}}$ is equal to the sum of its powers, denoted by $|\boldsymbol{\alpha}| = \sum_{k=1}^{n} \alpha_k$, and the highest total degree with respect to $\bar{\boldsymbol{z}}$ among all the monomials of $p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$ defines the degree d_i in $\bar{\boldsymbol{z}}$ of that polynomial. For example,

$$\begin{aligned} p(\boldsymbol{z}, \bar{\boldsymbol{z}}) &= 2 + z_2 + 3 z_1 z_2 \bar{z}_1 + z_1^2 \bar{z}_2 \\ &= p^{(00)}(\boldsymbol{z}) + p^{(10)}(\boldsymbol{z}) \bar{z}_1 + p^{(01)}(\boldsymbol{z}) \bar{z}_2 \end{aligned} \tag{12}$$

has a degree in $\bar{\boldsymbol{z}}$ equal to 1 and the corresponding polynomial coefficients $p^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ are

$$\begin{aligned} p^{(00)}(\boldsymbol{z}) &= 2 + z_2, \\ p^{(10)}(\boldsymbol{z}) &= 3 z_1 z_2, \\ p^{(01)}(\boldsymbol{z}) &= z_1^2. \end{aligned}$$

When we multiply a polynomial $p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$ by an arbitrary monomial $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i}$, we obtain a "new" polynomial,

$$\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i} p_i(\boldsymbol{z}, \bar{\boldsymbol{z}}) = \sum_{\{\boldsymbol{\alpha}\}} p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z}) \bar{\boldsymbol{z}}^{\boldsymbol{\alpha} + \boldsymbol{\delta}_i}, \tag{13}$$

which is similar to re-assigning every $p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ to a monomial of higher total degree. Note that these "new" polynomials do not alter the solution set $\mathcal{V}_{\mathbb{C}}$ in (10) when we add them, after equating to zero, to (9). Therefore, we define the Macaulay matrix with respect to the conjugate variables $\bar{\boldsymbol{z}}$. This matrix corresponds to the Macaulay matrix from elimination theory when treating the complex variables $\boldsymbol{z}$ as a constant vector [21], [20].

Definition 1: Given the polynomials $p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$, each with total degree d_i in $\bar{\boldsymbol{z}}$, the *Macaulay matrix with respect to the complex conjugate variables* of degree d in $\bar{\boldsymbol{z}}$, $\boldsymbol{\mathcal{M}}(\boldsymbol{z}) \in \mathbb{C}^{k \times l}$, contains the polynomial coefficients $p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ of the polynomials $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i} p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$ with all monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i}$ so that $|\boldsymbol{\delta}_i| = 0, \ldots, d - d_i$, for $i = 1, \ldots, 2n$. Every row of $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ contains one polynomial $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i} p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$, while every column is associated with one monomial $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha} + \boldsymbol{\delta}_i}$, the highest total degree of the monomials is equal to d.

The Macaulay matrix $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ with respect to the conjugate variables is clearly a polynomial matrix in the complex variable $\boldsymbol{z}$ that gathers the polynomial coefficients $p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ according to a certain pre-defined monomial ordering [20], [22]. The number of rows and columns of $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ depend on the degree d in $\bar{\boldsymbol{z}}$:

$$k = \sum_{i=1}^{2n} \binom{d - d_i + n}{d - d_i} \quad \text{and} \quad l = \binom{d + n}{d}.$$

Definition 1 can be used to rewrite (9) and additional equations (13) as a matrix-vector product,

$$\boldsymbol{\mathcal{M}}(\boldsymbol{z}) \boldsymbol{q} = \boldsymbol{0}, \tag{14}$$

where the vector $\boldsymbol{q}$ contains the different complex conjugate monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha} + \boldsymbol{\delta}_i}$ ordered the same as the columns of $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$. This Macaulay matrix eliminates the explicit dependency on the complex conjugate variables $\bar{\boldsymbol{z}}$. For example, we can multiply $p(\boldsymbol{z}, \bar{\boldsymbol{z}})$ in (12) with all monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}}$ for which $|\boldsymbol{\delta}| = 1$, i.e., $\bar{z}_1 p(\boldsymbol{z}, \bar{\boldsymbol{z}})$ and $\bar{z}_2 p(\boldsymbol{z}, \bar{\boldsymbol{z}})$, or, construct the Macaulay matrix with respect to $\bar{\boldsymbol{z}}$ of degree $d = 2$ in $\bar{\boldsymbol{z}}$, to obtain a matrix-vector product as in (14):

$$\underbrace{\begin{bmatrix} 2 + z_2 & 3 z_1 z_2 & z_1^2 & 0 & 0 & 0 \\ 0 & 2 + z_2 & 0 & 3 z_1 z_2 & z_1^2 & 0 \\ 0 & 0 & 2 + z_2 & 0 & 3 z_1 z_2 & z_1^2 \end{bmatrix}}_{\boldsymbol{\mathcal{M}}(\boldsymbol{z})} \underbrace{\begin{bmatrix} 1 \\ \bar{z}_1 \\ \bar{z}_2 \\ \bar{z}_1^2 \\ \bar{z}_1 \bar{z}_2 \\ \bar{z}_2^2 \end{bmatrix}}_{\boldsymbol{q}} = \boldsymbol{0}.$$

Finally, notice that (14) is an MEP when expanding the multivariate matrix polynomial $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ in terms of the different complex monomials $\boldsymbol{z}^{\boldsymbol{\beta}}$,

$$\boldsymbol{\mathcal{M}}(\boldsymbol{z}) \boldsymbol{q} = \left(\sum_{\{\boldsymbol{\beta}\}} \boldsymbol{M}_{\boldsymbol{\beta}} \boldsymbol{z}^{\boldsymbol{\beta}} \right) \boldsymbol{q} = \boldsymbol{0}, \tag{15}$$

where the summation runs over all the multi-indices $\boldsymbol{\beta}$. The minimal required degree d of $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ is such that $k \geq l + n - 1$, which is a necessary condition for the MEP to have a zero-dimensional solution set [22]. The coefficient matrices $\boldsymbol{M}_{\boldsymbol{\beta}} \in \mathbb{C}^{k \times l}$ of the MEP impose the structure of $\boldsymbol{q}$ and contain the coefficients of the polynomial coefficients

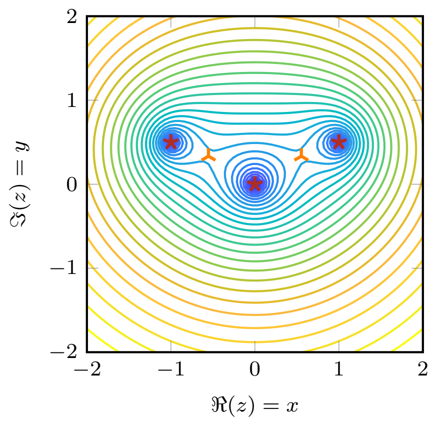

Fig. 2. Contour lines of the real-valued polynomial cost function $f(z, \bar{z})$ in Example 2: the optimization problem has three minimizers (★) and two saddle points (⅄).

$p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ associated with $\boldsymbol{z}^{\boldsymbol{\beta}}$. For the polynomial $p(\boldsymbol{z}, \bar{\boldsymbol{z}})$ in (12), the polynomial coefficient $p^{(00)}(\boldsymbol{z}) = 2 + z_2$ creates coefficients 2 and 1 in $\boldsymbol{M}_{00}$ and $\boldsymbol{M}_{01}$, respectively. For a more rigorous definition of MEPs, we refer the interested reader to [22]. One approach to solve[4] MEPs is via the block Macaulay matrix approach [22], [23].

Remark 2: In the univariate case, (9) only consists out of two bivariate equations in on complex variable z and its complex conjugate $\bar{z}$. An elimination of $\bar{z}$ results in a polynomial eigenvalue problem (PEP) instead of the MEP in (15). The Macaulay matrix from Definition 1 reduces to the well-known Sylvester matrix for $n = 1$. Note that the univariate case of our proposed optimization approach yields a similar PEP as in [7], [8].

Example 2: Consider the univariate problem

$$\min_{z} \left\| z(z - 0.5j)^2 - z \right\|_2^2.$$

The corresponding system of Wirtinger derivatives for one complex variable z and its complex conjugate $\bar{z}$ is

$$\begin{cases} p_1(z, \bar{z}) = 3z^2\bar{z}^3 + 3\mathrm{j}z^2\bar{z}^2 - 3.75z^2\bar{z} - 2\mathrm{j}z\bar{z}^3 \\ \qquad + 2z\bar{z}^2 + 2.5\mathrm{j}z\bar{z} - 1.25\bar{z}^3 - 1.25\mathrm{j}\bar{z}^2 \\ \qquad + 1.5625\bar{z} = 0, \\ p_2(z, \bar{z}) = 3z^3\bar{z}^2 - 3\mathrm{j}z^2\bar{z}^2 - 3.75z\bar{z}^2 + 2\mathrm{j}z^3\bar{z} \\ \qquad + 2z^2\bar{z} - 2.5\mathrm{j}z\bar{z} - 1.25z^3 + 1.25\mathrm{j}z^2 \\ \qquad + 1.5625z = 0. \end{cases} \tag{16}$$

We can construct the corresponding Sylvester matrix via Definition 1 (cf., Remark 2),

$$\boldsymbol{\mathcal{M}}(z) = \left[\begin{array}{ccccc} p_1^{(0)}(z) & p_1^{(1)}(z) & p_1^{(2)}(z) & p_1^{(3)}(z) & 0 \\ 0 & p_1^{(0)}(z) & p_1^{(1)}(z) & p_1^{(2)}(z) & p_1^{(3)}(z) \\ \hline p_2^{(0)}(z) & p_2^{(1)}(z) & p_2^{(2)}(z) & 0 & 0 \\ 0 & p_2^{(0)}(z) & p_2^{(1)}(z) & p_2^{(2)}(z) & 0 \\ 0 & 0 & p_2^{(0)}(z) & p_2^{(1)}(z) & p_2^{(2)}(z) \end{array}\right],$$

[4]We do not elaborate further on the solution methods, but we mention that we use the block Macaulay matrix methods available at www.macaulaylab.net to obtain all our numerical results.

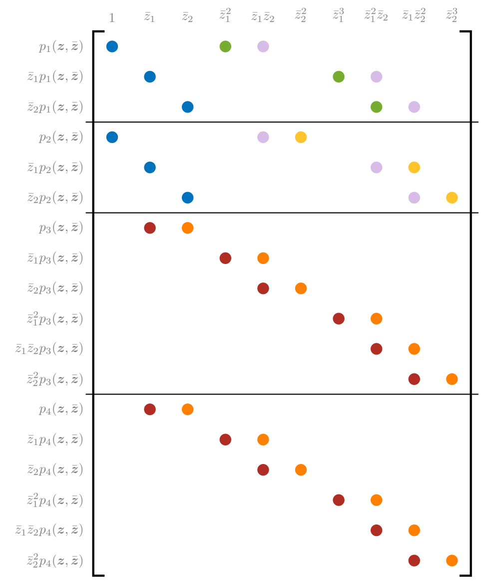

Fig. 3. Visualization of the Macaulay matrix that generates the coefficient matrices of the MEP in (18). The row-labels denote the shifted polynomials $\bar{\boldsymbol{z}}^{\boldsymbol{\delta}_i} p_i(\boldsymbol{z}, \bar{\boldsymbol{z}})$, while the column-labels denote the associated monomials $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha}}$. Every colored dot corresponds to one of the (non-zero) polynomial coefficients $p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ of the polynomials. For example, the green dot (●) corresponds to $p_1^{(20)}(\boldsymbol{z})$, which is shifted throughout the Macaulay matrix after multiplying $p_1(\boldsymbol{z}, \bar{\boldsymbol{z}})$ by 1 (i.e., the original polynomial), $\bar{z}_1$, and $\bar{z}_2$.

where $p_1^{(\alpha)}(z)$ contains the polynomial that is associated with $\bar{z}^{\alpha}$ in (16). For example, $p_1^{(2)}(z) = 3\mathrm{j}z^2 + 2z - 1.25\mathrm{j}$, because that are the monomials of $p_1(z, \bar{z})$ that are associated with $\bar{z}^2$. Subsequently, we create the coefficient matrices of the PEP from the Sylvester matrix by extracting the coefficients that belong to a power of z^{β}:

$$\left(\boldsymbol{M}_0 + \boldsymbol{M}_1 z + \boldsymbol{M}_2 z^2\right)\boldsymbol{q} = \boldsymbol{0}. \tag{17}$$

Taking again $p_1^{(2)}(z) = 3\mathrm{j}z^2 + 2z - 1.25\mathrm{j}$, this leads to the coefficients $-1.25\mathrm{j}$ in $\boldsymbol{M}_0$, 2 in $\boldsymbol{M}_1$, and 3j in $\boldsymbol{M}_2$ at the positions of $p_1^{(2)}(z)$ in $\boldsymbol{\mathcal{M}}(z)$. For clarity, we show $\boldsymbol{M}_2$:

$$\boldsymbol{M}_2 = \left[\begin{array}{ccccc} 0 & -3.75 & 3\mathrm{j} & 3 & 0 \\ 0 & 0 & -3.75 & 3\mathrm{j} & 3 \\ \hline 1.25\mathrm{j} & 2 & -3\mathrm{j} & 0 & 0 \\ 0 & 1.25\mathrm{j} & 2 & -3\mathrm{j} & 0 \\ 0 & 0 & 1.25\mathrm{j} & 2 & -3\mathrm{j} \end{array}\right].$$

Solving the resulting PEP, or the system (16) directly, yields 13 affine solutions: 3 minimizers, 2 saddle points, and 8 ghost solutions. Fig. 2 visualizes the minimizers and saddle points on top of the contour lines of the real-valued polynomial cost function. We discuss the reason for the emergence of these ghost solutions in Section V.

Example 3: We consider the problem where we try to fit a rank-1 matrix to a given complex 2×2 matrix $\boldsymbol{A} \in \mathbb{C}^{2\times 2}$:

$$\min_{\boldsymbol{z}} \left\| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} z_1^2 & z_1 z_2 \\ z_1 z_2 & z_2^2 \end{bmatrix} \right\|_{\mathrm{F}}^2,$$

which is an example of a nonlinear least-squares optimization problem (3). The corresponding system of first-order necessary conditions for optimality is

$$\begin{cases} p_1(\boldsymbol{z},\bar{\boldsymbol{z}}) = -2\bar{a}_{11}z_1 + 2z_1\bar{z}_1^2 - (\bar{a}_{12}+\bar{a}_{21})z_2 \\ \qquad + 2z_2\bar{z}_1\bar{z}_2 = 0, \\ p_2(\boldsymbol{z},\bar{\boldsymbol{z}}) = -2\bar{a}_{22}z_2 + 2z_2\bar{z}_2^2 - (\bar{a}_{12}+\bar{a}_{21})z_1 \\ \qquad + 2z_1\bar{z}_1\bar{z}_2 = 0, \\ p_3(\boldsymbol{z},\bar{\boldsymbol{z}}) = -2a_{11}\bar{z}_1 + 2z_1^2\bar{z}_1 - (a_{12}+a_{21})\bar{z}_2 \\ \qquad + 2z_1z_2\bar{z}_2 = 0, \\ p_4(\boldsymbol{z},\bar{\boldsymbol{z}}) = -2a_{22}\bar{z}_2 + 2z_2^2\bar{z}_2 - (a_{12}+a_{21})\bar{z}_1 \\ \qquad + 2z_1z_2\bar{z}_1 = 0, \end{cases}$$

with complex variables $\boldsymbol{z} = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^{\mathrm{T}}$ and their complex conjugates $\bar{\boldsymbol{z}} = \begin{bmatrix} \bar{z}_1 & \bar{z}_2 \end{bmatrix}^{\mathrm{T}}$. Fig. 3 visualizes the Macaulay matrix $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ with respect to the complex conjugate variables of degree $d = 3$ in $\bar{\boldsymbol{z}}$ for these polynomials. Each coefficient of $\boldsymbol{\mathcal{M}}(\boldsymbol{z})$ is a polynomial coefficient $p_i^{(\boldsymbol{\alpha})}(\boldsymbol{z})$ associated with a complex conjugate monomial $\bar{\boldsymbol{z}}^{\boldsymbol{\alpha}}$. For example, the green dot (●) corresponds to $p_1^{(20)}(\boldsymbol{z}) = 2z_1$ and is associated with $\bar{z}_1^2$. This leads to a quadratic two-parameter eigenvalue problem,

$$\big(\boldsymbol{M}_{00} + \boldsymbol{M}_{10}z_1 + \boldsymbol{M}_{01}z_2 + \boldsymbol{M}_{20}z_1^2 + \boldsymbol{M}_{11}z_1z_2 + \boldsymbol{M}_{02}z_2^2\big)\boldsymbol{q} = \boldsymbol{0}, \tag{18}$$

which we can solve, for example, via a block Macaulay matrix approach [22]. If we consider the given matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1\mathrm{j} \\ 1\mathrm{j} & -2 \end{bmatrix},$$

then we obtain nine stationary points after solving (18). The global minimizer is $(0.8507, 1.3764\mathrm{j})$, which also corresponds to the first triplet obtained via the complex singular value decomposition of $\boldsymbol{A}$.

V. About Ghost Solutions

In the context of complex optimization, *ghost solutions* (sometimes called *spurious solutions*) arise due to the fact that numerical optimization algorithms can not properly deal with complex conjugate variables. Ghost solutions can also arise in our proposed optimization approach, cf. Example 2. They emerge when solving the MEP in (15), or the system of multivariate polynomial equations (9) directly, via numerical linear algebra algorithms that can not impose that $\bar{\boldsymbol{z}}$ is the complex conjugate of $\boldsymbol{z}$. In that case, we essentially tackle the problem as if $\boldsymbol{z}$ and $\bar{\boldsymbol{z}}$ (let us call them $\boldsymbol{u}$ and $\boldsymbol{v}$ here) are independent variables, which results in the candidate solution set (instead of the desired solution set of (10))

$$\mathcal{V}_{\tilde{\mathbb{C}}} = \big\{(\boldsymbol{u}_0,\boldsymbol{v}_0) \in \mathbb{C}^{2n} : p_i(\boldsymbol{u}_0,\boldsymbol{v}_0) = 0, \forall i = 1,\ldots,2n\big\}.$$

Of course, we only want the subset for which $(\boldsymbol{u}_0,\boldsymbol{v}_0) = (\boldsymbol{z}_0,\bar{\boldsymbol{z}}_0)$, i.e., the true stationary points of (1), and need to remove these ghost solutions. Luckily, this is not a difficult task, even if we only compute the eigentuples $\boldsymbol{u}$ ($\boldsymbol{v}$ is then part of the eigenvector): we can (i) substitute the obtained eigentuples and their complex conjugates in (9) and check if $(\boldsymbol{u}_0,\boldsymbol{v}_0)$ is indeed a stationary point of (1) or (ii) construct an eigenvector $\boldsymbol{q}_0$ from the complex conjugate of $\boldsymbol{u}_0$ and check if $\boldsymbol{q}_0$ is indeed an eigenvector of $\boldsymbol{\mathcal{M}}(\boldsymbol{u}_0)$. An alternative heuristic technique to filter out ghost solutions (and to prune wrong solutions due to rounding errors) based on the well-known Newton–Raphson method was proposed in [7]. However, this technique is known to fail in some cases [7].

TABLE I

Numerical values of the candidate solutions (u_0, v_0) of (16). A classification for every candidate solution is given.

$\boldsymbol{u}_0$	$\boldsymbol{v}_0$	**classification**
$1.0000 + 0.5000\mathrm{j}$	$1.0000 - 0.5000\mathrm{j}$	minimizer
$-1.0000 + 0.5000\mathrm{j}$	$-1.0000 - 0.5000\mathrm{j}$	minimizer
$0.0000 + 0.0000\mathrm{j}$	$0.0000 + 0.0000\mathrm{j}$	minimizer
$0.5528 + 0.3333\mathrm{j}$	$0.5528 - 0.3333\mathrm{j}$	saddle point
$-0.5528 + 0.3333\mathrm{j}$	$-0.5528 - 0.3333\mathrm{j}$	saddle point
$1.0000 + 0.5000\mathrm{j}$	$-1.0000 - 0.5000\mathrm{j}$	ghost solution
$1.0000 + 0.5000\mathrm{j}$	$0.0000 + 0.0000\mathrm{j}$	ghost solution
$-1.0000 + 0.5000\mathrm{j}$	$1.0000 - 0.5000\mathrm{j}$	ghost solution
$-1.0000 + 0.5000\mathrm{j}$	$0.0000 + 0.0000\mathrm{j}$	ghost solution
$0.0000 + 0.0000\mathrm{j}$	$1.0000 - 0.5000\mathrm{j}$	ghost solution
$0.0000 + 0.0000\mathrm{j}$	$-1.0000 - 0.5000\mathrm{j}$	ghost solution
$-0.5528 + 0.3333\mathrm{j}$	$0.5528 - 0.3333\mathrm{j}$	ghost solution
$0.5528 + 0.3333\mathrm{j}$	$-0.5528 - 0.3333\mathrm{j}$	ghost solution

Remark 3: Note that using the standard approach for complex optimization, using derivatives with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$ and solving the resulting system of first-order necessary conditions for optimality, also can result in ghost solutions. In this approach, ghost solutions are candidate solutions that are complex-valued, while $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ have to be real-valued. These ghost solutions emerge because systems of multivariate polynomial equations and MEPs, without additional constraints, can also have complex solutions. When considering a specific problem with both approaches, it is possible to show that every ghost solution $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ corresponds to a ghost solution $(\boldsymbol{u}_0,\boldsymbol{v}_0)$, via (4), and vice versa, via (5).

Example 4: When solving the PEP in (17) or the system of multivariate polynomial equations in (16) with numerical linear algebra algorithms, we obtain 13 affine solutions (Table I): 3 minimizers, 2 saddle points, and 8 ghost solutions. The ghost solutions can be deflated from the candidate solution set by checking for every candidate solution u_0 if the candidate solution u_0 and its complex conjugate $\bar{u}_0$ are indeed a solution of (16) or by checking if the eigenvector $\boldsymbol{q}_0$ constructed from the complex conjugate $\bar{u}_0$ of u_0 is indeed an eigenvector of the PEP evaluated in u_0.

VI. Open Research Questions

Since this reformulation provides an alternative view on multivariate polynomial optimization in complex variables, it invokes new research challenges. We highlight some of these interesting open research questions below.

- A first question is about the efficiency of this new approach. "What is the computational complexity of this approach and how does it compare to current state-of-the-art polynomial optimization solvers?" Upper bounds on the computational complexity would be useful for this comparison.

- Closely related with the previous question is the fact that it is not know a priori whether the Macaulay matrix is of a sufficient degree to provide the minimizers. “Do there exist necessary and sufficient conditions on the degree of the Macaulay matrix?”
- Given that the coefficient matrices are constructed via the Macaulay matrix, they exhibit a lot of structure and sparsity: “Is it possible to exploit the structure and sparsity of the coefficient matrices in the MEP?”
- Furthermore, the solution set of the MEP is also linked to the specific coefficient matrices: “Can we enforce that the MEP has a zero-dimensional solution set?”
- An interesting property of the block Macaulay matrix algorithms to solve MEPs is that a user-defined shift polynomial can be used, as explained in [23]. “Could this property be exploited when the cost function is chosen as shift polynomial?”
- Finally, it is well known that a polynomial’s roots can be very sensitive to small changes in the polynomial’s coefficients [24]. “Does the MEP approach have better numerical properties than directly solving (9)?”

While Sorber et al. [7], [8] have only covered univariate problems, the multivariate extension in this paper provides the next step into exploring this alternative optimization approach.

VII. Conclusion and future work

In this paper, we extended the relation between univariate polynomial optimization in one complex variable and PEPs to the multivariate case. We showed that optimizing a real-valued multivariate polynomial cost function leads to an MEP, (some of) the eigentuples of which correspond to the stationary points of the optimization problem. Combining Wirtinger derivatives and the block Macaulay matrix provided a novel approach to solve multivariate polynomial optimization problems in complex variables. Furthermore, we also explained how to remove ghost solutions from the candidate solution set.

The perspective of this contribution was meant to be didactic; hence, we focussed on the implications of this alternative formulation and highlighted some interesting open research questions created by this novel optimization approach. We will tackle (some of) these research challenges in the (near) future. In particular, we want to look at the unavoidable structure in the coefficient matrices of the resulting MEP: since the Macaulay matrix consists of cogradients and conjugate cogradients, the coefficient matrices of the MEP are structured and exhibit significant sparsity, which could be exploited with improved solution algorithms. Our ambition is to use this method to tackle applications within the systems and control community, e.g., to identify the optimal complex poles of linear time-invariant models [6].

References

[1] L. Sorber, M. Van Barel, and L. De Lathauwer, “Unconstrained optimization of real functions in complex variables,” *SIAM Journal on Optimization*, vol. 22, no. 3, pp. 879–898, 2012.

[2] T. Adali and P. J. Schreier, “Optimization and estimation of complex-valued signals,” *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 112–128, 2014.

[3] Ç. Candan, “Properly handling complex differentiation in optimization and approximation problems,” *IEEE Signal Processing Magazine*, vol. 36, no. 2, pp. 117–124, 2019.

[4] K. Kreutz-Delgado, “The complex gradient operator and the $\mathbb{CR}$-calculus,” University of California, San Diego, CA, USA, Lecture notes, 2009.

[5] D. H. Brandwood, “A complex gradient operator and its application in adaptive array theory,” *IEE Proceedings H (Microwaves, Optics and Antennas)*, vol. 130, no. 1, pp. 11–16, 1983.

[6] B. De Moor, “Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem,” *Communications in Information and Systems*, vol. 20, no. 2, pp. 163–207, 2020.

[7] L. Sorber, M. Van Barel, and L. De Lathauwer, “Numerical solution of bivariate and polyanalytic polynomial systems,” *SIAM Journal on Numerical Analysis*, vol. 52, no. 4, pp. 1551–1572, 2014.

[8] L. Sorber, I. Domanov, M. Van Barel, and L. De Lathauwer, “Exact line and plane search for tensor optimization,” *Computational Optimization and Applications*, vol. 63, no. 1, pp. 121–142, 2016.

[9] H. Poincaré, “Sur les propriétés du potentiel et sur les fonctions Abéliennes,” *Acta Mathematica*, vol. 22, no. 1, pp. 89–178, 1899, [citation only].

[10] W. Wirtinger, “Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen,” *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927, [citation only].

[11] D. Henrion, M. Korda, and J.-B. Lasserre, *The Moment-SOS Hierarchy*, ser. Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs. London, United Kingdom: World Scientific, 2020.

[12] D. Henrion and J.-B. Lasserre, “Detecting global optimality and extracting solutions in GloptiPoly,” in *Positive Polynomials in Control*, ser. Lecture Notes in Control and Information Science, D. Henrion and A. Garulli, Eds. Berlin, Germany: Springer, 2005, vol. 312, pp. 293–310.

[13] M. Laurent, “Sums of squares, moment matrices and optimization over polynomials,” in *Emerging Applications of Algebraic Geometry*, ser. The IMA Volumes in mathematics and its Applications, M. Putinar and S. Sullivant, Eds. New York, NY, USA: Springer, 2008, vol. 149.

[14] J.-B. Lasserre, *An Introduction to Polynomial and Semi-Algebraic Optimization*, ser. Cambridge Texts in Applied Mathematics. Cambridge, UK: Cambridge University Press, 2015, vol. 52.

[15] J. Nie, J. Demmel, and B. Sturmfels, “Minimizing polynomials via sum of squares over the gradient ideal,” *Mathematical Programming*, vol. 106, pp. 587–606, 2006.

[16] P. Dreesen and B. De Moor, “Polynomial optimization problems are eigenvalue problems,” in *Model-Based Control*, P. M. J. Hof, C. Scherer, and P. S. C. Heuberger, Eds. Boston, MA, USA: Springer, 2009, pp. 49–68.

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., ser. Springer Series in Operations Research and Financial Engineering. New York, NY, USA: Springer, 2006.

[18] H. J. Stetter, *Numerical Polynomial Algebra*. Philadelphia, PA, USA: SIAM, 2004.

[19] C. Vermeersch and B. De Moor, “A column space based approach to solve systems of multivariate polynomial equations,” *IFAC-PapersOnLine*, vol. 54, no. 9, pp. 137–144, 2021, part of special issue: 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS).

[20] D. A. Cox, J. B. Little, and D. O’Shea, *Using Algebraic Geometry*, 2nd ed., ser. Graduate Texts in Mathematics. New York, NY, USA: Springer, 2004.

[21] F. S. Macaulay, “Some formulae in elimination,” *Proc. of the London Mathematical Society*, vol. 1, no. 1, pp. 3–27, 1902.

[22] C. Vermeersch and B. De Moor, “Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems,” *Linear Algebra and its Applications*, vol. 654, pp. 177–209, 2022.

[23] ——, “Two double recursive block Macaulay matrix algorithms to solve multiparameter eigenvalue problems,” *IEEE Control Systems Letters*, vol. 7, pp. 319–324, 2023.

[24] J. H. Wilkinson, “The perfidious polynomial,” in *Studies in Numerical Analysis*, G. H. Golub, Ed. Washington, D.C., USA: Mathematical Association of America, 1984, vol. 24, pp. 1–28.