A Distance Measure for Perspective Observability and Observability of Riccati Systems

Richard Seeber and Nicolaos Dourdoumas, *Life Member, IEEE*

***Abstract*—Systems governed by Riccati differential equations arise in several areas of control system theory. In combination with a linear fractional output, observability of such systems is relevant in the context of robotics and computer vision, for example, when studying the reconstruction of point locations from their perspective projections. The so-called perspective observability criteria exist to verify this observability property algebraically, but they provide only a binary answer. The present contribution studies the assessment of perspective and Riccati observability in a quantitative way, in terms of the distance to the closest nonobservable system. For this purpose, a distance measure is proposed. An optimization problem for determining it is derived, which features a quadratic cost function and an orthogonality constraint. The solution of this optimization problem by means of a descent algorithm is discussed and demonstrated in the course of a practically motivated numerical example.**

***Index Terms*—Computer vision, nonlinear systems, observability measures, optimization, perspective projection.**

I. Introduction

Riccati differential equations play an important role in the field of control theory. They occur in a variety of different contexts such as robust and optimal control [1] or observability analysis and observer design for time-varying systems [2], for example. This article studies systems of Riccati differential equations with fractional outputs, which are called Riccati systems in the following. As discussed in [3], [4], these are of particular importance in the context of computer vision in the form of perspective systems. Such systems describe the motion of points and curves observed via their perspective projection by means of a camera.

Extensive literature is available on observers for perspective systems; see, e.g., [5]–[9]. Contributions studying the observability of such systems or of Riccati systems in general are more scarce. An important breakthrough was achieved in [10], where a generalization of the well-known Popov–Belevitch–Hautus rank criterion is shown to be connected to the observability of Riccati systems. This connection is shown via the closely connected concept of *perspective observability* of linear time-invariant systems, which was studied before in [11]–[14] and more recently in [15], [16]. Observability of a subset of the Riccati systems considered here is also investigated in [17].

While algebraic criteria may be used to check if a system is observable in theory, ensuring the reliable operation of observers in practice requires a *quantitative* assessment of this property. For linear time-invariant systems, measures for observability (or for the dual property, controllability) exist for this purpose, see, e.g., [18]–[21], and they are used, e.g., for parameter design and sensor placement [22]. For Riccati systems, no such measures have been studied yet.

This article considers the problem of quantitatively assessing perspective observability and observability of Riccati systems in terms of their distance to unobservability. For this purpose, a distance measure for perspective observability is proposed. It is shown that this measure yields the distance to unobservability for both observability notions, and hence may be used to assess their robustness with respect to parameter perturbations in practice.

The article is structured as follows. Section II introduces Riccati systems and Section III discusses some existing results regarding their observability. The problem statement is then given in Section IV. Section V discusses the distance to unobservability for Riccati systems and proposes the distance measure for perspective observability as its solution. Section VI studies properties of the distance measure and proposes a nonlinear optimization problem for obtaining it. The numerical solution of this optimization problem is discussed in Section VII. In Section VIII, the proposed measure is computed numerically for an exemplary system that describes the rotation of a point observed by a perspective projection. Section IX gives a conclusion and a brief outlook.

Throughout the article, the following notation is used. The identity matrix is denoted by $\mathbf{I}$, and $\mathbf{M}^{\mathrm{T}}$ and $\mathbf{M}^{\mathrm{H}}$ are written for the transpose and complex conjugate (Hermitian) transpose of a matrix $\mathbf{M}$. Its trace, determinant, minimum singular value, and maximum singular value are denoted by $\mathrm{tr}(\mathbf{M})$, $\det(\mathbf{M})$, $\sigma_{\min}(\mathbf{M})$, and $\sigma_{\max}(\mathbf{M})$, respectively, and $\|\mathbf{M}\|$ is written for its Frobenius norm, i.e., $\|\mathbf{M}\| = \sqrt{\mathrm{tr}(\mathbf{M}^{\mathrm{H}}\mathbf{M})}$. For a vector $\mathbf{v}$, $\|\mathbf{v}\|$ means its Euclidian norm, i.e., $\|\mathbf{v}\| = \sqrt{\mathbf{v}^{\mathrm{H}}\mathbf{v}}$.

II. Riccati Systems

Consider the Riccati differential equation

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = \mathbf{A}_{11}\mathbf{P} - \mathbf{P}\mathbf{A}_{22} - \mathbf{P}\mathbf{A}_{21}\mathbf{P} + \mathbf{A}_{12}, \tag{1a}$$

with the state matrix $\mathbf{P}$ of size $(n-d) \times d$, and an output matrix $\mathbf{Q}$ of size $(m-d) \times d$, given by the relation

$$\mathbf{Q} = (\mathbf{C}_{11}\mathbf{P} + \mathbf{C}_{12})(\mathbf{C}_{21}\mathbf{P} + \mathbf{C}_{22})^{-1}. \tag{1b}$$

Manuscript received 15 April 2020; revised 19 February 2021; accepted 21 January 2022. Date of publication 4 February 2022; date of current version 30 January 2023. This work was supported in part by the Christian Doppler Research Association, in part by the Austrian Federal Ministry for Digital and Economic Affairs, and in part by the National Foundation for Research, Technology and Development. Recommended by Associate Editor K. Camlibel. *(Corresponding author: Richard Seeber.)*

The authors are with the Christian Doppler Laboratory for Model Based Control of Complex Test Bed Systems, Institute of Automation and Control, Graz University of Technology, 8010 Graz, Austria (e-mail: richard.seeber@tugraz.at; nicolaos.dourdoumas@tugraz.at).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2022.3148381.

Digital Object Identifier 10.1109/TAC.2022.3148381

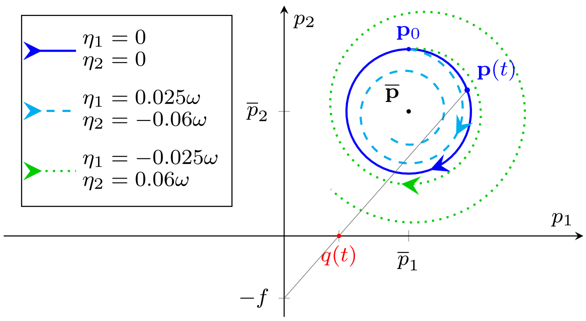

Fig. 1. Exemplary trajectories of a point $\mathbf{p}$ governed by (2b) and its camera image q projected onto the p_1-axis with focal length f according to (2a).

Therein, $\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21}, \mathbf{A}_{22}$ and $\mathbf{C}_{11}, \mathbf{C}_{12}, \mathbf{C}_{21}, \mathbf{C}_{22}$ are constant matrices of appropriate[1] sizes with $\operatorname{rank}[\mathbf{C}_{21}\ \mathbf{C}_{22}] = d$. These are real-valued typically, but systems with complex-valued parameters, state, and output are permitted here as well. The initial state is denoted by $\mathbf{P}_0 := \mathbf{P}(0)$.

To see the relevance of such systems in the context of computer vision, consider, for example, the movement of a single point observed by means of a camera. For simplicity and ease of illustration, considerations are restricted to a two-dimensional space (the three-dimensional extension is straightforward). Suppose that the point is observed by a camera with focal length f and a horizontal (one-dimensional) image plane located at the origin. Denoting the coordinates of the point coordinates by the vector $\mathbf{p} := [p_1 \quad p_2]^{\mathrm{T}}$, the coordinate of the camera image q is then given by

$$q = \frac{f p_1}{p_2 + f} = \left(\begin{bmatrix} f & 0 \end{bmatrix} \mathbf{p}\right)\left(\begin{bmatrix} 0 & 1 \end{bmatrix} \mathbf{p} + f\right)^{-1} \tag{2a}$$

i.e., by a perspective projection of the point onto the abscissa, as depicted in Fig. 1. A Riccati system is then obtained, if the point movement is governed by a Riccati differential equation, such as, for example

$$\begin{aligned}\dot{\mathbf{p}} &= \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix}(\mathbf{p} - \overline{\mathbf{p}}) - (\eta_1 p_1 + \eta_2 p_2)(\mathbf{p} - \overline{\mathbf{p}}) \\ &= \begin{bmatrix} \eta_1 \overline{p}_1 & \omega + \eta_2 \overline{p}_1 \\ -\omega + \eta_1 \overline{p}_2 & \eta_2 \overline{p}_2 \end{bmatrix} \mathbf{p} - \mathbf{p}\begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \mathbf{p} + \begin{bmatrix} -\omega \overline{p}_2 \\ \omega \overline{p}_1 \end{bmatrix}\end{aligned} \tag{2b}$$

with parameters $\overline{\mathbf{p}} = [\overline{p}_1 \quad \overline{p}_2]^{\mathrm{T}}$, η_1, η_2, and ω. Fig. 1 depicts some exemplary trajectories. For $\eta_1 = 0$, $\eta_2 = 0$, the differential equation (2b) describes a rigid rotation of the point with angular velocity ω and center $\overline{\mathbf{p}}$; with nonzero η_1, η_2, nonrigid motions may be modeled as well; see [3].

III. Observability Properties

This section defines and discusses the observability of Riccati systems, the related concept of perspective observability, and their connection.

[1]The matrices $\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21}, \mathbf{A}_{22}$ and $\mathbf{C}_{11}, \mathbf{C}_{12}, \mathbf{C}_{21}, \mathbf{C}_{22}$ have sizes $(n-d) \times (n-d)$, $(n-d) \times d$, $d \times (n-d)$, $d \times d$ and $(m-d) \times (n-d)$, $(m-d) \times d$, $d \times (n-d)$, $d \times d$, respectively.

A. Observability of Riccati Systems

The observability of system (1) may be defined in terms of its solutions. Unlike with linear systems, a solution $\mathbf{P}(t)$ of (1) need not exist for all $t \geq 0$, however. Furthermore, even if $\mathbf{P}(t)$ exists for some values of t, the output $\mathbf{Q}(t)$ need not be defined for all those values.

Let $\mathcal{I}(\mathbf{P}_0)$ denote the maximal (possibly empty or unbounded) time interval of the form $[0, T)$ such that the solution $\mathbf{P}(t)$ of (1a) with initial value $\mathbf{P}_0$ exists and satisfies

$$\det\left[\mathbf{C}_{21}\mathbf{P}(t) + \mathbf{C}_{22}\right] \neq 0 \quad \text{for all } t \in \mathcal{I}(\mathbf{P}_0). \tag{3}$$

System (1) is said to have a solution for a given initial value $\mathbf{P}_0$, if $\mathcal{I}(\mathbf{P}_0)$ is nonempty. The observability of the Riccati system (1) is then defined using essentially the same idea as in [23, Def. 2.1] by requiring that two distinct initial states yield different outputs.

Definition 1: System (1) is called observable, if for all pairs of initial values $\mathbf{P}_0^{(1)}, \mathbf{P}_0^{(2)}$, for which (1) has a solution, and for corresponding outputs $\mathbf{Q}^{(1)}(t), \mathbf{Q}^{(2)}(t)$, the following implication holds:

$$\mathbf{Q}^{(1)}(t) = \mathbf{Q}^{(2)}(t) \ \ \forall t \in \mathcal{I}\big(\mathbf{P}_0^{(1)}\big) \cap \mathcal{I}\big(\mathbf{P}_0^{(2)}\big) \Rightarrow \mathbf{P}_0^{(1)} = \mathbf{P}_0^{(2)}. \tag{4}$$

The parameters of the Riccati system (1) may concisely be specified in terms of a pair $(\mathbf{A}, \mathbf{C})$ with matrices $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{m \times n}$ given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}. \tag{5}$$

Note that the representation of (1) as such a pair is not unique. In particular, $(\mathbf{A} - \delta\mathbf{I}, \nu\mathbf{C})$ with any constant δ and non-zero constant ν corresponds to the same system. Thus, it is often useful to impose restrictions such as

$$\operatorname{tr}\mathbf{A}_{22} = 0, \qquad \|\mathbf{C}_{21}\|^2 + \|\mathbf{C}_{22}\|^2 = 1 \tag{6}$$

for example, to reduce ambiguity in that regard.[2]

In the following, the Riccati system (1) is commonly referred to by its associated pair $(\mathbf{A}, \mathbf{C})$, where the integer d, which determines the sizes of $\mathbf{P}$ and $\mathbf{Q}$ and thus the partitioning (5), is assumed to be given. To differentiate between the classical notion of observability of the pair $(\mathbf{A}, \mathbf{C})$ meaning observability of a linear system with dynamic matrix $\mathbf{A}$ and output matrix $\mathbf{C}$, and observability of the associated Riccati system (1), the following notion of *Riccati observability* is introduced.

Definition 2 (Riccati Observability): For a given integer d, the pair $(\mathbf{A}, \mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{m \times n}$ is called *Riccati observable* if the associated Riccati system (1) with $\mathbf{P} \in \mathbb{C}^{(n-d) \times d}$, $\mathbf{Q} \in \mathbb{C}^{(m-d) \times d}$ is observable.

B. Perspective Observability

In [10], it is shown that observability of the Riccati system (1) is related to the so-called *perspective observability* of the pair $(\mathbf{A}, \mathbf{C})$. The integer d, which determines the dimensions of the state $\mathbf{P}$ and output $\mathbf{Q}$ of the Riccati system, is called the *dimension loss* in this context. Originally, this form of observability was introduced in [12], where the pair $(\mathbf{A}, \mathbf{C})$ is called perspectively observable with dimension loss d, if for any solution of a linear time-invariant system with dynamic matrix $\mathbf{A}$ and output matrix $\mathbf{C}$, knowledge of the output up to d dimensions permits reconstructing the initial state up to d dimensions. The definition used here follows [15], where for $d > 0$, the concept is defined more formally in terms of the observability of linear vectorspaces. Here,

[2]Note, however, that in some special cases, the pair may not be determined uniquely even with (6), e.g., for $\mathbf{C}_{11} = \mathbf{C}_{21} = \mathbf{0}$, $\mathbf{C}_{12} = \mathbf{C}_{22}$, i.e., $\mathbf{Q} = \mathbf{I}$.

the definition proposed therein is extended to affine rather than only linear vectorspaces, which permits to include also the case $d = 0$, while yielding an equivalent notion for $d > 0$.

Definition 3 (Perspective Observability): For a given integer d, the pair $(\mathbf{A}, \mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{C} \in \mathbb{C}^{m\times n}$ is called *perspectively observable*, if for any two complex-valued d-dimensional affine subspaces $\mathcal{X}^{(1)}(t), \mathcal{X}^{(2)}(t)$ of $\mathbb{C}^n$ and corresponding output spaces $\mathcal{Y}^{(1)}(t)$, $\mathcal{Y}^{(2)}(t)$ of $\mathbb{C}^m$ that satisfy[3]

$$\frac{\mathrm{d}\mathcal{X}}{\mathrm{d}t} = \mathbf{A}\mathcal{X}, \qquad \mathcal{Y} = \mathbf{C}\mathcal{X} \tag{7}$$

the following implication holds:

$$\mathcal{Y}^{(1)}(t) = \mathcal{Y}^{(2)}(t) \text{ for all t} \Rightarrow \mathcal{X}^{(1)}(0) = \mathcal{X}^{(2)}(0). \tag{8}$$

Note that for $d = 0$, the affine subspaces $\mathcal{X}$ and $\mathcal{Y}$ collapse to vectors $\mathbf{x}$ and $\mathbf{y}$; in this case, perspective observability is thus equivalent to classical observability of the linear time-invariant system $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}$, $\mathbf{y} = \mathbf{C}\mathbf{x}$.

In [13], it is shown that perspective observability may be checked by means of the following generalization of the well-known Popov–Belevitch–Hautus observability criterion.

Theorem 1 ([13, Th. 2]): The pair $(\mathbf{A}, \mathbf{C})$ with matrices $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{C} \in \mathbb{C}^{m\times n}$ is perspectively observable with dimension loss d if and only if

$$\operatorname{rank} \begin{bmatrix} \prod_{i=1}^{d+1} (\lambda_i \mathbf{I} - \mathbf{A}) \\ \mathbf{C} \end{bmatrix} = n \tag{9}$$

holds for all $\lambda_1, \ldots, \lambda_{d+1} \in \mathbb{C}$.

Note that the matrix $\lambda_i \mathbf{I} - \mathbf{A}$ has rank n unless λ_i is an eigenvalue of $\mathbf{A}$. For applying the criterion, it is thus sufficient to check the rank condition for $\lambda_1, \ldots, \lambda_{d+1}$ taken from the set of eigenvalues of $\mathbf{A}$.

In [10], a connection between Riccati observability and perspective observability is shown in the form of the following theorem. Note that its converse only holds subject to some additional conditions, which are discussed later in Section VI-D.

Theorem 2 ([10, Th. 2.1]): If the pair $(\mathbf{A}, \mathbf{C})$ is perspectively observable, then it is Riccati observable.

IV. Problem Statement

If the Riccati system (1) is not observable, then there are initial states that cannot be reconstructed using the output. Unlike in the case of linear systems, however, the reconstruction need not be impossible for *all* initial states. To see this, consider the pair $(\mathbf{A}, \mathbf{C})$ with $d = 1$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}. \tag{10}$$

The corresponding Riccati system is

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = p_1 - 1, \qquad \frac{\mathrm{d}p_2}{\mathrm{d}t} = p_2, \qquad q = \frac{p_1 - 1}{p_2 + 1} \tag{11}$$

with state $[p_1 \quad p_2]^{\mathrm{T}}$ and output q. If the time derivative $\dot{q}$ of the output, which is given by

$$\dot{q} := \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{(p_2 + 1)(p_1 - 1) - (p_1 - 1)p_2}{(p_2 + 1)^2} = \frac{p_1 - 1}{(p_2 + 1)^2}, \tag{12}$$

is non-zero, one can compute p_1, p_2 from (11), (12) using

$$p_1 = \frac{q^2}{\dot{q}} + 1, \qquad p_2 = \frac{q}{\dot{q}} - 1. \tag{13}$$

The system is not observable, however, because $q(t) = 0$ holds for all t regardless of p_2, when the initial value of p_1 is one.

Due to this fact, an observer constructed for a nonobservable Riccati system may still work most of the time, but will fail for certain initial conditions. Thus, checking observability before constructing an observer is of crucial importance. In practice, this is not enough, however, because reconstructing the state even for an observable system, while always possible in theory, may be ill-conditioned. Therefore, rather than just verifying whether a given system is observable or not, it is important to quantify observability numerically.

To address this issue, the distance to unobservability may be considered. Such a distance was first considered for linear time-invariant systems in [19], [24]. In addition to quantifying observability, it allows to assess robustness of observability in the presence of parameter uncertainties. The present contribution considers the following problem of determining the *distance to Riccati unobservability* or *to perspective unobservability:* For a given pair $(\mathbf{A}, \mathbf{C})$, what is the distance, in terms of the norm of the change in parameters, to the closest Riccati unobservable or perspectively unobservable system?

V. Distance to Riccati Unobservability

To put the considered problem into context, the problem of quantitatively *measuring* observability in general is first discussed. For the classical, dual properties observability and controllability of *linear time-invariant systems*, several measures have been proposed in the literature for this purpose; see, e.g., [19]–[21]. In the following, properties desired from such a measure for any observability property are introduced similar to [25], [26].

Definition 4 (Observability Measure): A real-valued function $\mu(\mathbf{A}, \mathbf{C})$ is called an observability measure for the pair $(\mathbf{A}, \mathbf{C})$, if it satisfies the conditions

1) *non-negativity:* $\mu(\mathbf{A}, \mathbf{C}) \geq 0$ for all $\mathbf{A}, \mathbf{C}$;
2) *continuity:* $\mu(\mathbf{A}, \mathbf{C})$ is a continuous function of $\mathbf{A}, \mathbf{C}$; and
3) *consistency:* $\mu(\mathbf{A}, \mathbf{C}) \neq 0 \Leftrightarrow (\mathbf{A}, \mathbf{C})$ observable.

While such a formal definition is typically not considered in the literature, it is clear that these properties are desirable: Consistency guarantees that a loss of observability causes the measure to vanish, and continuity ensures that this cannot happen abruptly when changing the parameters.

Existence of such a measure requires, however, that the set of observable systems is open. In the following, this is shown to be the case only for perspective observability. Motivated by this result, a distance measure for perspective observability is then introduced, which is finally shown to yield also the distance to Riccati unobservability.

A. Nonexistence of Measures for Riccati Observability

The following counterexample illustrates that the set of Riccati observable systems in general is not open, and hence measures in the sense of Definition 4 do not exist for Riccati observability.

Example 1: Consider the pairs $(\mathbf{A}_\varepsilon, \mathbf{C})$ with a non-negative parameter ε and matrices $\mathbf{A}_\varepsilon, \mathbf{C}$ given by

$$\mathbf{A}_\varepsilon = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ -\varepsilon & \varepsilon & 0 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{14}$$

[3]For a given initial value $\mathcal{X}(0)$, the solution of system (7) is understood as $\mathcal{X}(t) = e^{\mathbf{A}t}\mathcal{X}(0) = \{e^{\mathbf{A}t}\mathbf{x} : \mathbf{x} \in \mathcal{X}(0)\}$.

For $d = 1$, the corresponding Riccati systems (1) are

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = -p_1 + \varepsilon(p_1^2 - p_1 p_2) \tag{15a}$$

$$\frac{\mathrm{d}p_2}{\mathrm{d}t} = p_2 + \varepsilon(p_1 p_2 - p_2^2) \tag{15b}$$

$$q = p_1 + p_2 \tag{15c}$$

with state $[p_1 \quad p_2]^{\mathrm{T}}$ and output q. For $\varepsilon = 0$, this is an observable linear system; thus, $(\mathbf{A}_0, \mathbf{C})$ is Riccati observable. Otherwise, two different solutions that yield the same output $q(t)$ are given by $p_1(t) = \varepsilon^{-1}$, $p_2(t) = 0$ and $p_1(t) = 0$, $p_2(t) = \varepsilon^{-1}$. Thus, $(\mathbf{A}_\varepsilon, \mathbf{C})$ is not Riccati observable for $\varepsilon > 0$. By consistency and continuity, one then has for any Riccati observability measure μ the contradiction

$$0 \neq \mu(\mathbf{A}_0, \mathbf{C}) = \lim_{\varepsilon \to 0} \mu(\mathbf{A}_\varepsilon, \mathbf{C}) = \lim_{\varepsilon \to 0} 0 = 0. \tag{16}$$

Thus, no function μ satisfying Definition 4 exists for the case of Riccati observability.

The following proposition formalizes the insight of this example by extending it to the general case.

Proposition 1: For positive integers d, n, m, let $\mathcal{R}_d^{(n,m)}$ be the set of all pairs $(\mathbf{A}, \mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{m \times n}$ that correspond to observable Riccati systems. If $n - d \geq 2$, then the set $\mathcal{R}_d^{(n,m)}$ is not open.

Proof: The proof is given in the Appendix. ■

B. Distance Measure for Perspective Observability

Unlike the set of Riccati observable systems, the set of perspectively observable systems is open, because it is characterized by the algebraic rank condition (9). Hence, motivated by the considered distance problem for Riccati observability, the following *distance measure* for perspective observability is introduced, cf. [27]. Denoting by $\overline{\mathcal{P}}_d$ the set of pairs that are not perspectively observable for a given dimension loss d, it is defined as

$$\mu_d(\mathbf{A}, \mathbf{C}) := \inf \left\{ \left\| \begin{bmatrix} \tilde{\mathbf{A}} \\ \tilde{\mathbf{C}} \end{bmatrix} \right\| : (\mathbf{A} + \tilde{\mathbf{A}}, \mathbf{C} + \tilde{\mathbf{C}}) \in \overline{\mathcal{P}}_d \right\}. \tag{17}$$

It is straightforward to verify that (17) satisfies Definition 4; in particular, item 3 holds because the set of perspectively observable systems is open: If $(\mathbf{A}, \mathbf{C})$ is perspectively observable, then it is contained in a neighborhood of perspectively observable systems, and hence $\mu_d(\mathbf{A}, \mathbf{C}) > 0$.

In the following section, an optimization problem for this measure is formulated and its properties are discussed. Furthermore, the following connection to the distance to Riccati unobservability is shown.

Theorem 3: Suppose that $(\mathbf{A}, \mathbf{C})$ is not perspectively observable. Then, for every $\varepsilon > 0$, there exists $\tilde{\mathbf{A}}$ such that $\|\tilde{\mathbf{A}}\| \leq \varepsilon$ holds and the pair $(\mathbf{A} + \tilde{\mathbf{A}}, \mathbf{C})$ is not Riccati observable.

Proof: The proof is given in Section VI-D. ■

From Theorems 2 and 3, the following corollary is obvious; it shows that the proposed distance measure indeed solves the considered distance problem to Riccati unobservability.

Corollary 1: For the pair $(\mathbf{A}, \mathbf{C})$, the distance to Riccati unobservability is given by the distance measure for perspective observability $\mu_d(\mathbf{A}, \mathbf{C})$.

VI. Distance Measure

To study the proposed distance measure as defined in (17), an algebraic characterization of perspective observability is first introduced. Then, an optimization problem for the distance measure is derived, its properties are studied, and its connection to Riccati observability is shown.

A. Algebraic Characterization of Perspective Observability

Using the criterion for perspective observability given in Theorem 1, the following technical lemma may be obtained.

Lemma 1: The following statements are equivalent:

1) The pair $(\mathbf{A}, \mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{m \times n}$ is not perspectively observable.
2) For every vector $\mathbf{u} \in \mathbb{C}^{d+1}$ with $\|\mathbf{u}\| = 1$, there exist matrices $\mathbf{V} \in \mathbb{C}^{n \times (d+1)}$, $\mathbf{S} \in \mathbb{C}^{(d+1) \times (d+1)}$ such that
$$\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}, \quad \mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{S}, \quad \mathbf{C}\mathbf{V}\mathbf{u} = \mathbf{0} \tag{18}$$
holds.
3) There exists a nonzero vector $\mathbf{u} \in \mathbb{C}^{d+1}$ as well as two matrices $\mathbf{V} \in \mathbb{C}^{n \times (d+1)}$, $\mathbf{S} \in \mathbb{C}^{(d+1) \times (d+1)}$ such that (18) holds.

Proof: The proof is given in the Appendix. ■

B. Optimization Problem for the Distance Measure

By rewriting the constraint in (17) using the algebraic characterization of perspective observability in Lemma 1, the following optimization problem is obtained.

Theorem 4: Consider the pair $(\mathbf{A}, \mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{m \times n}$ and let the dimension loss $d \leq n - 1$ and a vector $\mathbf{u} \in \mathbb{C}^{d+1}$ with $\|\mathbf{u}\| = 1$ be given. Then, the distance measure μ_d is given by $\mu_d(\mathbf{A}, \mathbf{C}) = \sqrt{\kappa}$, where κ is the solution of the nonlinear optimization problem

$$\kappa := \min_{\mathbf{V}, \mathbf{S}} \|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\|^2 + \|\mathbf{C}\mathbf{V}\mathbf{u}\|^2 \tag{19a}$$

subject to

$$\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}, \quad \mathbf{V} \in \mathbb{C}^{n \times (d+1)}, \quad \mathbf{S} \in \mathbb{C}^{(d+1) \times (d+1)}. \tag{19b}$$

Remark 1: For $d = 0$, the well-known optimization problem

$$\mu_0(\mathbf{A}, \mathbf{C}) = \min_{s \in \mathbb{C}} \min_{\mathbf{v}^{\mathrm{H}}\mathbf{v} = 1} \sqrt{\|(\mathbf{A} - s\mathbf{I})\mathbf{v}\|^2 + \|\mathbf{C}\mathbf{v}\|^2}$$
$$= \min_{s \in \mathbb{C}} \sigma_{\min} \begin{bmatrix} \mathbf{A} - s\mathbf{I} \\ \mathbf{C} \end{bmatrix} \tag{20}$$

for the distance measure of observability is obtained [19], [28].

Proof: It is first shown that $\mu_d(\mathbf{A}, \mathbf{C}) \leq \sqrt{\kappa}$ holds. To that end, denote by $\mathbf{V}, \mathbf{S}$ the optimal solution of (19). Consider the pair $(\hat{\mathbf{A}}, \hat{\mathbf{C}})$ with

$$\hat{\mathbf{A}} = \mathbf{A} - (\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S})\mathbf{V}^{\mathrm{H}}, \qquad \hat{\mathbf{C}} = \mathbf{C} - \mathbf{C}\mathbf{V}\mathbf{u}\mathbf{u}^{\mathrm{H}}\mathbf{V}^{\mathrm{H}}. \tag{21}$$

Using $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$ and $\mathbf{u}^{\mathrm{H}}\mathbf{u} = \|\mathbf{u}\|^2 = 1$, one verifies that $\hat{\mathbf{A}}\mathbf{V} = \mathbf{V}\mathbf{S}$, $\hat{\mathbf{C}}\mathbf{V}\mathbf{u} = \mathbf{0}$ holds; thus, $(\hat{\mathbf{A}}, \hat{\mathbf{C}})$ is not perspectively observable according to Lemma 1, item 3. Therefore, one has

$$\mu_d(\mathbf{A}, \mathbf{C}) \leq \left\| \begin{bmatrix} \mathbf{A} - \hat{\mathbf{A}} \\ \mathbf{C} - \hat{\mathbf{C}} \end{bmatrix} \right\| = \left\| \begin{bmatrix} (\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S})\mathbf{V}^{\mathrm{H}} \\ \mathbf{C}\mathbf{V}\mathbf{u}\mathbf{u}^{\mathrm{H}}\mathbf{V}^{\mathrm{H}} \end{bmatrix} \right\|$$
$$= \sqrt{\|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\|^2 + \|\mathbf{C}\mathbf{V}\mathbf{u}\|^2} = \sqrt{\kappa}. \tag{22}$$

To show $\mu_d(\mathbf{A}, \mathbf{C}) \geq \sqrt{\kappa}$, let the matrices $\tilde{\mathbf{A}}, \tilde{\mathbf{C}}$ attain the minimum in (17), i.e., let $(\mathbf{A} + \tilde{\mathbf{A}}, \mathbf{C} + \tilde{\mathbf{C}})$ be the pair closest to $(\mathbf{A}, \mathbf{C})$ that is not perspectively observable. By Lemma 1, item 2, there exist matrices $\mathbf{V}, \mathbf{S}$ with $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$ such that

$$(\mathbf{A} + \tilde{\mathbf{A}})\mathbf{V} = \mathbf{V}\mathbf{S}, \qquad (\mathbf{C} + \tilde{\mathbf{C}})\mathbf{V}\mathbf{u} = \mathbf{0}. \tag{23}$$

One then has

$$\mu_d(\mathbf{A},\mathbf{C}) = \left\| \begin{bmatrix} \tilde{\mathbf{A}} \\ \tilde{\mathbf{C}} \end{bmatrix} \right\| \geq \sqrt{\left\|\tilde{\mathbf{A}}\mathbf{V}\right\|^2 + \left\|\tilde{\mathbf{C}}\mathbf{V}\mathbf{u}\right\|^2}$$
$$= \sqrt{\|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\|^2 + \|\mathbf{C}\mathbf{V}\mathbf{u}\|^2} \geq \sqrt{\kappa} \tag{24}$$

which concludes the proof. ■

While the structure of the optimization problem (19) reflects the algebraic structure visible in Lemma 1, the number of optimization variables can be reduced by further minimizing the objective function. This is shown in the following corollary.

Corollary 2: Consider the pair $(\mathbf{A},\mathbf{C})$ with $\mathbf{A} \in \mathbb{C}^{n\times n}, \mathbf{C} \in \mathbb{C}^{m\times n}$ and let the dimension loss d be given. The distance measure μ_d is given by $\mu_d(\mathbf{A},\mathbf{C}) = \sqrt{\kappa}$, where κ is the solution of the nonlinear optimization problem

$$\kappa := \min_{\mathbf{V}} \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\|^2 + \sigma_{\min}(\mathbf{C}\mathbf{V})^2 \tag{25a}$$

subject to

$$\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}, \qquad \mathbf{V} \in \mathbb{C}^{n\times(d+1)}. \tag{25b}$$

Remark 2: Note that the objective function in (25a) depends only on the space spanned by the columns of $\mathbf{V}$: It is invariant with respect to replacing $\mathbf{V}$ by $\mathbf{V}\mathbf{Q}$ with a matrix $\mathbf{Q} \in \mathbb{C}^{(d+1)\times(d+1)}$ that is unitary, i.e., $\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{\mathrm{H}} = \mathbf{I}$. Therefore, the optimization problem always has multiple global optima.

Proof: In Theorem 4, the unit-length vector $\mathbf{u}$ may be chosen arbitrarily. Thus, minimizing with respect to $\mathbf{u}$ does not change κ and (19) may equivalently be written as

$$\kappa = \min_{\mathbf{u}} \min_{\mathbf{V},\mathbf{S}} \|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\|^2 + \|\mathbf{C}\mathbf{V}\mathbf{u}\|^2 \tag{26}$$

subject to the constraints (19b) and $\|\mathbf{u}\| = 1$. Taking $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$ into account, one has

$$\|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\|^2 = \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\|^2 + \left\|\mathbf{S} - \mathbf{V}^{\mathrm{H}}\mathbf{A}\mathbf{V}\right\|^2 \tag{27}$$

and minimization of terms in (26) with respect to $\mathbf{S}$, $\mathbf{u}$ yields

$$\min_{\mathbf{S}} \|\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{S}\| = \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\| \tag{28a}$$

$$\min_{\|\mathbf{u}\|=1} \|\mathbf{C}\mathbf{V}\mathbf{u}\| = \sigma_{\min}(\mathbf{C}\mathbf{V}). \tag{28b}$$

Thus, κ is given by

$$\kappa = \min_{\mathbf{V}} \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\|^2 + \sigma_{\min}(\mathbf{C}\mathbf{V})^2 \tag{29}$$

subject to $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$, which completes the proof. ■

C. Bounds of the Distance Measure

When using the distance measure in practice, it is useful to know the range of values that the measure can take. The following proposition gives lower and upper bounds for this purpose.

Proposition 2: The distance measure $\mu_d(\mathbf{A},\mathbf{C})$ satisfies the inequalities $\sqrt{\sigma_{\min}(\mathbf{C}^{\mathrm{H}}\mathbf{C})} \leq \mu_d(\mathbf{A},\mathbf{C}) \leq \sqrt{\sigma_{\max}(\mathbf{C}^{\mathrm{H}}\mathbf{C})}$ for all matrices $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{C} \in \mathbb{C}^{m\times n}$.

Remark 3: Note that the given lower bound is nonzero if and only if $\operatorname{rank}\mathbf{C} = n$.

Proof: The inequality may be obtained from optimization problem (25). Denoting its optimal solution by $\mathbf{V}$, one has

$$\mu_d(\mathbf{A},\mathbf{C}) \geq \sigma_{\min}(\mathbf{C}\mathbf{V}) = \sqrt{\sigma_{\min}(\mathbf{V}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}\mathbf{C}\mathbf{V})}$$
$$\geq \sqrt{\sigma_{\min}(\mathbf{C}^{\mathrm{H}}\mathbf{C})}. \tag{30}$$

Furthermore, using any matrices $\mathbf{V}$, $\mathbf{S}$ that satisfy $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$ and $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{S}$, one obtains

$$\mu_d(\mathbf{A},\mathbf{C}) \leq \sqrt{\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\|^2 + \sigma_{\min}(\mathbf{C}\mathbf{V})^2}$$
$$= \sqrt{\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{V}\mathbf{S}\|^2 + \sigma_{\min}(\mathbf{C}\mathbf{V})^2}$$
$$= \sigma_{\min}(\mathbf{C}\mathbf{V}) \leq \sqrt{\sigma_{\max}(\mathbf{C}^{\mathrm{H}}\mathbf{C})} \tag{31}$$

which completes the proof. ■

D. Connection to Riccati Observability

In order to draw the connection to Riccati observability and to prove Theorem 3, the following lemma is introduced. It is essentially a restatement of [10, Th. 2.2] in terms of the algebraic criterion from Section VI-A.

Lemma 2: Suppose that item 3 of Lemma 1 is fulfilled with $\operatorname{rank}\mathbf{V} = d+1$ and that the following two $d \times (d+1)$ matrices have maximal rank

$$\operatorname{rank}\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{V} = \operatorname{rank}\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{C}\mathbf{V} = d. \tag{32}$$

Then, the pair $(\mathbf{A},\mathbf{C})$ is not Riccati observable.

Proof: The proof is given in the Appendix. ■

Using this technical lemma, Theorem 3 can now be proven.

Proof of Theorem 3: Let matrices $\mathbf{V},\mathbf{S}$ and vector $\mathbf{u}$ be as in item 2 of Lemma 1. Furthermore, let matrices $\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2 \in \mathbb{C}^{n\times(d+1)}$ be such that $\tilde{\mathbf{V}}_1\mathbf{u} = \tilde{\mathbf{V}}_2\mathbf{u} = \mathbf{0}$ and that

$$\operatorname{rank}\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\tilde{\mathbf{V}}_1 = \operatorname{rank}\begin{bmatrix}\mathbf{0} & \mathbf{I}\end{bmatrix}\mathbf{C}\tilde{\mathbf{V}}_2 = d. \tag{33}$$

Existence of such matrices is verified by comparing the number of constraints with the degrees of freedom. Then, the matrix $\mathbf{V} + \tilde{\mathbf{V}}$ with $\tilde{\mathbf{V}} = \varepsilon_1\tilde{\mathbf{V}}_1 + \varepsilon_2\tilde{\mathbf{V}}_2$ satisfies (32) for almost all $\varepsilon_1, \varepsilon_2$, and moreover $\mathbf{C}(\mathbf{V} + \tilde{\mathbf{V}})\mathbf{u} = \mathbf{0}$ holds. Choosing

$$\tilde{\mathbf{A}} = (-\mathbf{A}\tilde{\mathbf{V}} + \tilde{\mathbf{V}}\mathbf{S})\left[(\mathbf{V} + \tilde{\mathbf{V}})^{\mathrm{H}}(\mathbf{V} + \tilde{\mathbf{V}})\right]^{-1}(\mathbf{V} + \tilde{\mathbf{V}})^{\mathrm{H}} \tag{34}$$

furthermore achieves $(\mathbf{A} + \tilde{\mathbf{A}})(\mathbf{V} + \tilde{\mathbf{V}}) = (\mathbf{V} + \tilde{\mathbf{V}})\mathbf{S}$ and applying Lemma 2 shows loss of Riccati observability for $(\mathbf{A} + \tilde{\mathbf{A}}, \mathbf{C})$. The proof is concluded by noting that $\|\tilde{\mathbf{A}}\|$ can be made arbitrarily small by making $\varepsilon_1, \varepsilon_2$ small enough. ■

VII. Computation of the Distance Measure

Computing the distance measure requires solving the optimization problem (25), which is nonconvex due to the orthogonality constraint (25b). In general, it cannot be solved analytically, and efficient methods to find its global optimum only exist for $d = 0$, i.e., for classical observability; see [28]. Nevertheless, local optima may be found numerically; this section discusses a procedure to do so.

The handling of orthogonality constraints is studied in the literature, e.g., in [29]–[31]. In [31], a gradient descent procedure based on the Cayley transformation is proposed. Here, a descent algorithm based on this idea is applied to the optimization problem (25) in Corollary 2 with cost function

$$J(\mathbf{V}) := \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\|^2 + \sigma_{\min}(\mathbf{C}\mathbf{V})^2 \tag{35}$$

in the following.

Consider decision variables $\mathbf{V}_k \in \mathbb{C}^{n\times(d+1)}$ in iteration k of the algorithm's execution, which satisfy $\mathbf{V}_k^{\mathrm{H}}\mathbf{V}_k = \mathbf{I}$. An admissible matrix

$\mathbf{V}_{k+1}$ for the next iteration is then computed as $\mathbf{V}_{k+1} = \bar{\mathbf{V}}_k(\tau_k)$, where τ_k is a step-size, and the Cayley transformation-based function $\bar{\mathbf{V}}_k$ is given by

$$\bar{\mathbf{V}}_k(\tau) := \left(\mathbf{I} + \frac{\tau}{2}\mathbf{W}_k\right)^{-1}\left(\mathbf{I} - \frac{\tau}{2}\mathbf{W}_k\right)\mathbf{V}_k. \tag{36}$$

Therein, $\mathbf{W}_k = -\mathbf{W}_k^{\mathrm{H}}$ is a skew-symmetric matrix, whose choice is discussed later. This function satisfies $\bar{\mathbf{V}}_k(0) = \mathbf{V}_k$ as well as $\bar{\mathbf{V}}_k^{\mathrm{H}}(\tau)\bar{\mathbf{V}}_k(\tau) = \mathbf{V}_k^{\mathrm{H}}\mathbf{V}_k = \mathbf{I}$ for all τ, thus preserving the orthogonality constraint. Its derivative at $\tau = 0$ is given by

$$\left.\frac{\partial \bar{\mathbf{V}}_k}{\partial \tau}\right|_{\tau=0} = -\mathbf{W}_k\mathbf{V}_k. \tag{37}$$

By proper choice of the matrix $\mathbf{W}_k$, one may ensure that the inequality $J(\bar{\mathbf{V}}_k(\tau)) < J(\mathbf{V}_k)$ holds for small enough $\tau > 0$, provided that $\mathbf{V}_k$ is not a local optimum. An appropriate step-size τ_k to decrease J may thus be obtained by means of a line search.

In [31], the matrix $\mathbf{W}_k$ is obtained from the gradient of the objective function. The function J is not differentiable, however, if the multiplicity of the minimum singular value $\sigma_{\min}(\mathbf{CV})$ is greater than one. To see this, note, for example, that

$$\sigma_{\min}\begin{bmatrix}1 & 0\\ 0 & 1+\varepsilon\end{bmatrix} = \begin{cases}1 & \varepsilon \geq 0\\ 1+\varepsilon & \varepsilon < 0\end{cases} \tag{38}$$

holds for $|\varepsilon| \leq 1$. Thus, (38) is not differentiable at $\varepsilon = 0$. To nevertheless use the descent algorithm, the following proposition gives a descent direction of the objective function J defined in (35) regardless of its differentiability.

Proposition 3: Consider the function J in (35) for given matrices $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{C} \in \mathbb{C}^{m\times n}$. For a given matrix $\mathbf{V}_k \in \mathbb{C}^{n\times(d+1)}$ satisfying $\mathbf{V}_k^{\mathrm{H}}\mathbf{V}_k = \mathbf{I}$, let $\mathbf{u} \in \mathbb{C}^{d+1}$ be any unit-length right singular vector corresponding to the smallest singular value of the matrix $\mathbf{CV}_k$ and, using the abbreviation

$$\begin{aligned}\mathbf{G} :=\ & 2\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V}_k + 2\mathbf{C}^{\mathrm{H}}\mathbf{C}\mathbf{V}_k\mathbf{u}\mathbf{u}^{\mathrm{H}}\\ & - 2\mathbf{A}^{\mathrm{H}}\mathbf{V}_k\mathbf{V}_k^{\mathrm{H}}\mathbf{A}\mathbf{V}_k - 2\mathbf{A}\mathbf{V}_k\mathbf{V}_k^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{V}_k\end{aligned} \tag{39}$$

define $\mathbf{W}_k := \mathbf{G}\mathbf{V}_k^{\mathrm{H}} - \mathbf{V}_k\mathbf{G}^{\mathrm{H}}$. If the matrix $\mathbf{W}_k$ thus defined is nonzero, then $\bar{\tau} > 0$ exists such that with the function $\bar{\mathbf{V}}_k(\tau)$ defined in (36), the inequality $J(\bar{\mathbf{V}}_k(\tau)) < J(\mathbf{V}_k)$ holds for all $\tau \in (0, \bar{\tau}]$.

Proof: Define the function

$$\tilde{J}(\mathbf{V}) := \left\|(\mathbf{I} - \mathbf{V}\mathbf{V}^{\mathrm{H}})\mathbf{A}\mathbf{V}\right\|^2 + \|\mathbf{CVu}\|^2 \tag{40}$$

which satisfies $\tilde{J}(\mathbf{V}) \geq J(\mathbf{V})$ for all $\mathbf{V}$, because $\mathbf{u}^{\mathrm{H}}\mathbf{u} = 1$. This function $\tilde{J}$ is differentiable; with the abbreviations

$$\tilde{J}_1(\mathbf{V}) := \operatorname{tr}\mathbf{V}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V} \tag{41a}$$

$$\tilde{J}_2(\mathbf{V}) := \operatorname{tr}\mathbf{u}^{\mathrm{H}}\mathbf{V}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}\mathbf{CVu} = \operatorname{tr}\mathbf{V}^{\mathrm{H}}\mathbf{C}^{\mathrm{H}}\mathbf{CVu}\mathbf{u}^{\mathrm{H}} \tag{41b}$$

$$\tilde{J}_3(\mathbf{V}) := \operatorname{tr}\mathbf{V}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{V}\mathbf{V}^{\mathrm{H}}\mathbf{A}\mathbf{V} = \operatorname{tr}\mathbf{V}^{\mathrm{H}}\mathbf{A}\mathbf{V}\mathbf{V}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{V} \tag{41c}$$

it may be written as $\tilde{J} = \tilde{J}_1 + \tilde{J}_2 - \tilde{J}_3$ for matrices $\mathbf{V}$ satisfying $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$. Differentiation then yields

$$\frac{\partial \tilde{J}_1}{\partial \mathbf{V}^{\mathrm{H}}} = \mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V},\quad \frac{\partial \tilde{J}_2}{\partial \mathbf{V}^{\mathrm{H}}} = \mathbf{C}^{\mathrm{H}}\mathbf{CVu}\mathbf{u}^{\mathrm{H}} \tag{42a}$$

$$\frac{\partial \tilde{J}_3}{\partial \mathbf{V}^{\mathrm{H}}} = \mathbf{A}^{\mathrm{H}}\mathbf{V}\mathbf{V}^{\mathrm{H}}\mathbf{A}\mathbf{V} + \mathbf{A}\mathbf{V}\mathbf{V}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{V} \tag{42b}$$

see, e.g., [32], [33] for differentiation rules. Thus, one can see that $\mathbf{G}$ is the gradient of $\tilde{J}$ evaluated at $\mathbf{V}_k$, i.e.,

$$\mathbf{G} = \left[\frac{\partial \tilde{J}}{\partial \mathbf{V}^{\mathrm{H}}} + \left(\frac{\partial \tilde{J}}{\partial \mathbf{V}}\right)^{\mathrm{H}}\right]_{\mathbf{V}=\mathbf{V}_k} = 2\left.\frac{\partial \tilde{J}}{\partial \mathbf{V}^{\mathrm{H}}}\right|_{\mathbf{V}=\mathbf{V}_k}. \tag{43}$$

Due to skew-symmetry of $\mathbf{W}_k$, and denoting the real part of a complex number z by $\operatorname{Re} z$, one then has

$$\begin{aligned}\operatorname{Re}\operatorname{tr}\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k &= \operatorname{Re}\operatorname{tr}(\mathbf{V}_k\mathbf{G}^{\mathrm{H}} - \mathbf{G}\mathbf{V}_k^{\mathrm{H}})\mathbf{W}_k\\ &= \operatorname{Re}\operatorname{tr}\mathbf{V}_k\mathbf{G}^{\mathrm{H}}\mathbf{W}_k + \operatorname{Re}\operatorname{tr}\mathbf{W}_k\mathbf{V}_k\mathbf{G}^{\mathrm{H}}\\ &= 2\operatorname{Re}\operatorname{tr}\mathbf{G}^{\mathrm{H}}\mathbf{W}_k\mathbf{V}_k.\end{aligned} \tag{44}$$

With (37), the derivative of $\tilde{J}(\bar{\mathbf{V}}_k(\tau))$ at $\tau = 0$ thus satisfies

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}\tilde{J}(\bar{\mathbf{V}}_k(\tau)) &= \operatorname{Re}\operatorname{tr}\mathbf{G}^{\mathrm{H}}\left.\frac{\mathrm{d}\bar{\mathbf{V}}_k}{\mathrm{d}\tau}\right|_{\tau=0} = -\operatorname{Re}\operatorname{tr}\mathbf{G}^{\mathrm{H}}\mathbf{W}_k\mathbf{V}_k\\ &= -\operatorname{Re}\operatorname{tr}\frac{1}{2}\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k = -\frac{1}{2}\|\mathbf{W}_k\|^2 < 0\end{aligned} \tag{45}$$

if $\mathbf{W}_k$ is non-zero. Therefore, a positive constant $\bar{\tau}$ exists, such that for all $\tau \in (0, \bar{\tau}]$, the relation

$$J(\bar{\mathbf{V}}_k(\tau)) \leq \tilde{J}(\bar{\mathbf{V}}_k(\tau)) < \tilde{J}(\bar{\mathbf{V}}_k(0)) = J(\mathbf{V}_k) \tag{46}$$

holds, which completes the proof. ■

Using the derivative in (45), the step size τ_k in each iteration k may be computed using backtracking line search based on the Armijo–Goldstein condition [34] according to $\tau_k = \alpha\beta^{j_k}$ with the abbreviation

$$j_k = \min\{j \in \mathbb{N} : J(\bar{\mathbf{V}}_k(\alpha\beta^j)) \leq J(\mathbf{V}_k) - \alpha\beta^j\frac{\gamma}{2}\|\mathbf{W}_k\|^2\} \tag{47}$$

and positive initial step-size α and parameters $\beta, \gamma \in (0, 1)$.

VIII. Numerical Example

Consider system (2) with $\eta_1 = \eta_2 = 0$, which describes the rotary motion of a point observed by a camera with focal length f as shown in Fig. 1. With angular velocity and center of rotation being given by ω and $(\overline{p}_1, \overline{p}_2)$, respectively, it is described by the pair $(\mathbf{A}, \mathbf{C})$ with

$$\mathbf{A} = \begin{bmatrix}0 & \omega & -\omega\overline{p}_2\\ -\omega & 0 & \omega\overline{p}_1\\ 0 & 0 & 0\end{bmatrix},\qquad \mathbf{C} = \begin{bmatrix}f & 0 & 0\\ 0 & 1 & f\end{bmatrix} \tag{48}$$

and $d = 1$. Note that this pair does not satisfy (6). Instead, $\mathbf{C}$ is scaled such that it depends affinely on the focal length f; this way, the perturbation of $\mathbf{C}$ considered in the distance measure's definition (17) is more meaningful from a practical point of view. According to Proposition 2, an upper bound for the distance measure is given by $\sqrt{\sigma_{\max}(\mathbf{C}^{\mathrm{H}}\mathbf{C})} = \sqrt{1+f^2}$.

Using the procedure outlined in Section VII, the distance measure $\mu_d(\mathbf{A}, \mathbf{C})$ was computed numerically for $\overline{p}_1 = 2$, $\overline{p}_2 = 2$. As initial values for the descent algorithm, 20 random matrices $\mathbf{V}_0 \in \mathbb{C}^{3\times 2}$ satisfying $\mathbf{V}_0^{\mathrm{H}}\mathbf{V}_0 = \mathbf{I}$ were used, and optimization was terminated once $|J_k - J_{k-1}| \leq 10^{-6}$. Table I lists the statistics of the number of cost function evaluations, including those required for evaluating (47), with $\beta = 0.5$ and different step-size parameters α, γ.

Fig. 2 depicts distance measures obtained for different values of f as well as local minima obtained for all algorithm runs with $f = 1$ as a function of ω. One can see that $\mu_d(\mathbf{A}, \mathbf{C}) = 0$ holds for $\omega = 0$. This is consistent with the fact that system (2) is not observable for $\omega = 0$, because the location of a stationary point cannot be reconstructed. For $\omega \neq 0$, the results allow to conclude robustness of Riccati observability

TABLE I
STATISTICS OF NUMBER OF COST FUNCTION EVALUATIONS FROM 4500 OPTIMIZATIONS WITH $f \in \{0.25, 0.5, 1\}$, $\omega \in [0, 5]$, AND RANDOM INITIAL CONDITION $\mathbf{V}_0$ FOR $\beta = 0.5$ AND DIFFERENT PARAMETERS α, γ

α	γ	1st quartile	2nd quartile	3rd quartile	maximum
0.5	0.5	33	50	80	453
1	0.1	43	67	102	321
1	0.5	38	60	92	285
1	0.9	43	66	96	286
2	0.5	46	72	108	385

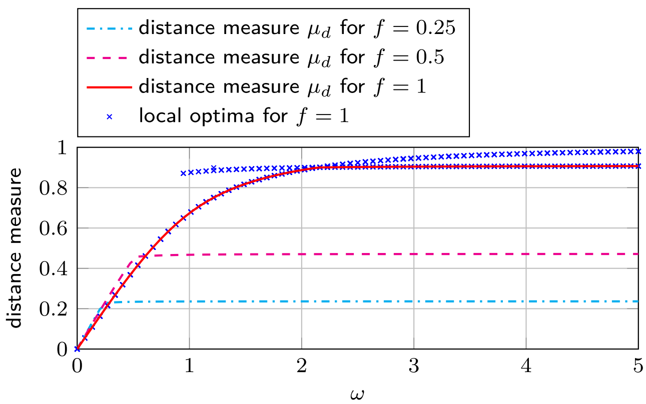

Fig. 2. Local optima and distance measure $\mu_d(\mathbf{A}, \mathbf{C})$ obtained using the descent algorithm with $\alpha = 1, \beta = \gamma = 0.5$ for (48) with $\bar{p}_1 = \bar{p}_2 = 2$ and different values of the focal length f as a function of the angular speed ω.

with respect to parameter perturbations that do not exceed the value of the distance measure. The depicted results furthermore suggest that the observability of the system is improved with increasing values of angular velocity ω or focal length f. Both facts match the intuition that reconstruction of the point's position becomes harder as the point moves slower or the size of its camera image decreases.

IX. CONCLUSION

The problem of assessing perspective observability and the observability of Riccati systems in terms of the distance to unobservability was considered. It was shown that in the latter case, the distance may be zero even for some observable systems. The reason is that the set of observable Riccati systems, unlike the set of perspectively observable systems, is not open. Nevertheless, the distances for the two observability notions were shown to coincide, and a distance measure for obtaining them was proposed. A nonlinear optimization problem for this distance measure was derived and a descent algorithm for its numerical solution was discussed. The effectiveness of the proposed approach for quantifying Riccati observability was shown using a practically motivated example.

Future work may focus on restricting the form of permitted perturbations, for example, to compute the distance measure only among systems exhibiting certain structural properties.

APPENDIX

Proof of Lemma 1: Suppose that matrices $\mathbf{V}, \mathbf{S}$, and a nonzero vector $\mathbf{u}$ satisfying (18) exist. Let $\lambda_1, \ldots, \lambda_{d+1}$ be the eigenvalues of $\mathbf{S}$. Then, one has due to Cayley–Hamilton

$$\prod_{i=1}^{d+1}(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{V} = \mathbf{V}\prod_{i=1}^{d+1}(\lambda_i \mathbf{I} - \mathbf{S}) = \mathbf{0} \tag{49}$$

and thus

$$\begin{bmatrix} \prod_{i=1}^{d+1}(\lambda_i \mathbf{I} - \mathbf{A}) \\ \mathbf{C} \end{bmatrix} \mathbf{V}\mathbf{u} = \mathbf{0}. \tag{50}$$

Since $\|\mathbf{V}\mathbf{u}\| = \|\mathbf{u}\| \neq 0$, one has $\mathbf{V}\mathbf{u} \neq \mathbf{0}$; hence, $(\mathbf{A}, \mathbf{C})$ is not perspectively observable according to Theorem 1.

Conversely, assume that $(\mathbf{A}, \mathbf{C})$ is not perspectively observable and let $\mathbf{u}$ be any unit-length vector. By Theorem 1, there exist $\lambda_1, \ldots, \lambda_{d+1} \in \mathbb{C}$ and a unit-length vector $\mathbf{w}$ that satisfy

$$\prod_{i=1}^{d+1}(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{w} = \mathbf{0}, \quad \mathbf{C}\mathbf{w} = \mathbf{0}. \tag{51}$$

Let $q(s)$ be the minimal polynomial of $\mathbf{w}$, i.e., the polynomial with minimal degree such that $q(\mathbf{A})\mathbf{w} = \mathbf{0}$ holds, and let $\mathcal{W}$ be the space spanned by the vectors $\mathbf{w}, \mathbf{A}\mathbf{w}, \ldots, \mathbf{A}^d\mathbf{w}$. This space is $\mathbf{A}$-invariant, i.e., $\mathbf{A}\mathcal{W} = \mathcal{W}$, and has dimension $\dim \mathcal{W} = \deg q(s) \leq d + 1$ due to (51). Using the following lemma from [35], an $\mathbf{A}$-invariant subspace $\mathcal{V}$ containing $\mathcal{W}$ with $\dim \mathcal{V} = d + 1$ is constructed.

Lemma 3 ([35, p. 202]): Let $\mathcal{W}$ be an $\mathbf{A}$-invariant subspace of $\mathbb{C}^n$ satisfying $\dim \mathcal{W} < n$. Then, there exists a vector $\mathbf{h} \notin \mathcal{W}$ such that $(\mathbf{A} - c\mathbf{I})\mathbf{h} \in \mathcal{W}$ for some eigenvalue c of $\mathbf{A}$.

With $\mathbf{h}$ as in this lemma, the space spanned by $\mathbf{h}$ and vectors in $\mathcal{W}$ is $\mathbf{A}$-invariant and has dimension $\dim \mathcal{W} + 1$; thus, repeated application eventually yields $\mathcal{V}$. Let $\mathbf{V}$ be a matrix, whose columns form an orthonormal basis for $\mathcal{V}$ and which satisfies $\mathbf{V}\mathbf{u} = \mathbf{w} \in \mathcal{W} \subseteq \mathcal{V}$. This matrix then fulfills $\mathbf{V}^{\mathrm{H}}\mathbf{V} = \mathbf{I}$ and $\mathbf{C}\mathbf{V}\mathbf{u} = \mathbf{C}\mathbf{w} = \mathbf{0}$. Due to the $\mathbf{A}$-invariance of $\mathcal{V}$, a matrix $\mathbf{S}$ furthermore exists such that $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{S}$ holds.

Proof of Lemma 2: For $\mathbf{W} \in \mathbb{C}^{(d+1)\times d}$, consider the solution

$$\mathbf{X}(t) = e^{\mathbf{A}t}\mathbf{V}\mathbf{W} = \mathbf{V}e^{\mathbf{S}t}\mathbf{W}, \qquad \mathbf{Y}(t) = \mathbf{C}\mathbf{V}e^{\mathbf{S}t}\mathbf{W} \tag{52}$$

of the system $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$, $\mathbf{Y} = \mathbf{C}\mathbf{X}$. With an appropriate partitioning $\mathbf{X} = [\mathbf{X}_1^{\mathrm{T}} \;\; \mathbf{X}_2^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{Y} = [\mathbf{Y}_1^{\mathrm{T}} \;\; \mathbf{Y}_2^{\mathrm{T}}]^{\mathrm{T}}$, it is straightforward to verify that $\mathbf{P}(t) = \mathbf{X}_1(t)\mathbf{X}_2(t)^{-1}$, $\mathbf{Q}(t) = \mathbf{Y}_1(t)\mathbf{Y}_2(t)^{-1}$ is a solution of the Riccati system (1), cf. [10]. For almost all $\mathbf{W}$, one has $\operatorname{rank} \mathbf{X}_2(0) = \operatorname{rank} \mathbf{Y}_2(0) = d$ due to (32), guaranteeing existence of these solutions on a nonempty time interval. Moreover, $\mathbf{Q}(t)$ and $\mathbf{P}(0)$ only depend on the images of $\mathbf{Y}(t)$ and $\mathbf{X}(0)$, respectively. Since $\operatorname{rank} \mathbf{C}\mathbf{V} \leq d$ and $\operatorname{rank} \mathbf{V} = d + 1$, the same $\mathbf{Q}(t)$ but different $\mathbf{P}(0)$ are obtained for almost all $\mathbf{W}$, proving loss of Riccati observability. ■

Proof of Proposition 1: Denote by $\mathbf{e}_1$ and $\mathbf{e}_2$ the first and second standard-basis vector, respectively, and by $\mathbf{1}$ a vector of ones, each with length as appropriate from context. Consider the Riccati system

$$\dot{\mathbf{P}} = \mathbf{A}_{11}\mathbf{P} - \mathbf{P}\mathbf{A}_{21}\mathbf{P}, \qquad \mathbf{Q} = \mathbf{C}_{11}\mathbf{P} \tag{53}$$

with $\mathbf{P} \in \mathbb{C}^{(n-d)\times d}$, $\mathbf{Q} \in \mathbb{C}^{(m-d)\times(n-d)}$ and parameters

$$\mathbf{A}_{21} = \varepsilon\mathbf{1} \cdot (-\mathbf{e}_1^{\mathrm{T}} + \mathbf{e}_2^{\mathrm{T}}), \qquad \mathbf{C}_{11} = \mathbf{1} \cdot \mathbf{1}^{\mathrm{T}} \tag{54}$$

and $\mathbf{A}_{11} = \operatorname{diag}(-1, 1, 2, \ldots, n - d - 1)$. For $\varepsilon = 0$, this is an observable linear time-invariant system with matrix-valued state $\mathbf{P}$, while for every $\varepsilon \neq 0$, the two constant solutions $\mathbf{P}^{(1)}(t) = \frac{1}{\varepsilon d}\mathbf{e}_1 \cdot \mathbf{1}^{\mathrm{T}}$ and $\mathbf{P}^{(2)}(t) = \frac{1}{\varepsilon d}\mathbf{e}_2 \cdot \mathbf{1}^{\mathrm{T}}$ yield the same constant output $\mathbf{Q}(t)$. ■

ACKNOWLEDGMENT

The authors would like to thank their colleague Markus Tranninger for several remarks that helped improve the article.

References

[1] M. Green and D. J. N. Limebeer. *Linear Robust Control*. Mineola, NY, USA: Dover Publications, 1995.

[2] P.-J. Bristeau, N. Petit, and L. Praly, "Design of a navigation filter by analysis of local observability," in *Proc. 49th IEEE Conf. Decis. Control*, 2010, pp. 1298–1305.

[3] B. K. Ghosh, H. Inaba, and S. Takahasbi, "Identification of Riccati dynamics under perspective and orthographic observations," *IEEE Trans. Autom. Control*, vol. 45, no. 7, pp. 1267–1278, Jul. 2000.

[4] M. Unel and B. K. Ghosh, "Dynamic models of planar algebraic curves," in *Proc. 40th IEEE Conf. Decis. Control*, 2001, vol. 2, pp. 1304–1309.

[5] J. P. Hespanha, "State estimation and control for systems with perspective outputs," in *Proc. IEEE Conf. Decis. Control*, 2002, vol. 2, pp. 2208–2213.

[6] X. Chen and H. Kano, "A new state observer for perspective systems," *IEEE Trans. Autom. Control*, vol. 47, no. 4, pp. 658–663, Apr. 2002.

[7] W. E. Dixon, Y. Fang, D. M. Dawson, and T. J. Flynn, "Range identification for perspective vision systems," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2232–2238, Dec. 2003.

[8] R. Abdursul, H. Inaba, and B. K. Ghosh, "Nonlinear observers for perspective time-invariant linear systems," *Automatica*, vol. 40, no. 3, pp. 481–490, 2004.

[9] J. Keshavan and J. S. Humbert, "Robust motion recovery from noisy stereoscopic measurements," in *Proc. Amer. Control Conf.*, 2018, pp. 2035–2040.

[10] J. Rosenthal, "An observability criterion for dynamic systems governed by Riccati differential equations," *IEEE Trans. Autom. Control*, vol. 41, no. 3, pp 434–436, Mar. 1996.

[11] B. K. Ghosh, M. Jankovic, and Y. T. Wu, "Some problems in perspective system theory and its application to machine vision," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 1992, pp. 139–146.

[12] B. K. Ghosh, E. P. Loucks, and M. Jankovic, "An introduction to perspective observability and recursive identification problems in machine vision," in *Proc. IEEE Conf. Decis. Control*, 1994, vol. 4, pp. 3229–3234.

[13] B. K. Ghosh and J. Rosenthal, "A generalized Popov–Belevitch–Hautus test of observability," *IEEE Trans. Autom. Control*, vol. 40, no. 1, pp. 176–180, Jan. 1995.

[14] W. P. Dayawansa and C. F. Martin, "Observability of the Riccati equation," in *Proc. 34th IEEE Conf. Decis. Control*, 1995, vol. 1, pp. 317–320.

[15] G. Dirr, U. Helmke, and J. Jordan, "Control and observation of the matrix Riccati differential equation," in *Emergent Problems in Nonlinear Systems and Control*, Berlin, Germany: Springer, 2009, pp. 169–184.

[16] S. Moschik, M. Stadler, and N. Dourdoumas, "On testing the perspective observability of time-continuous linear time-invariant systems," *at-Automatisierungstechnik*, (in German), vol. 60, no. 12, pp. 735–742, 2012.

[17] D. Xu, "Controllability and observability of matrix Riccati type differential systems," in *Proc. 10th World Congr. Intell. Control Automat.*, 2012, pp. 1135–1138.

[18] L. Litz, "Modal measures of controllability, observability, dominance—Connections, drawbacks and a new approach," *Regelungstechnik*, (in German), vol. 31, no. 5, pp. 148–158, 1983.

[19] R. Eising, "The distance between a system and the set of uncontrollable systems," in *Mathematical Theory of Networks and Systems*. Berlin, Germany: Springer, 1984, pp. 303–314.

[20] A. M. A. Hamdan and A. H. Nayfeh, "Measures of modal controllability and observability for first- and second-order linear systems," *J. Guidance, Control, Dyn.*, vol. 12, no. 3, pp. 421–428, 1989.

[21] M. Tarokh, "Measures for controllability, observability and fixed modes," *IEEE Trans. Autom. Control*, vol. 37, no. 8, pp. 1268–1273, Aug. 1992.

[22] K. B. Lim and W. Gawronski, "Actuator and sensor placement for control of flexible structures," in *Control and Dynamic Systems*. San Diego, CA, USA: Academic Press, Inc., 1993, vol. 57, pp. 109–152.

[23] J.-P. Gauthier and I. A. Kupka, "Observability and observers for nonlinear systems," *SIAM J. Control Optim.*, vol. 32, no. 4, pp. 975–994, 1994.

[24] C. C. Paige, "Properties of numerical algorithms related to computing controllability," *IEEE Trans. Autom. Control*, vol. AC-26, no. 1, pp. 130–138, Feb. 1981.

[25] R. Seeber, N. Dourdoumas, and M. Horn, "A modal measure for observability and perspective observability of linear time-invariant systems," *at-Automatisierungstechnik*, (in German), vol. 65, no. 10, pp. 683–694, 2017.

[26] R. Falkensteiner, R. Seeber, M. Reichhartinger, and M. Horn, "Distance measures for strong observability and strong detectability of systems with direct feedthrough," *IEEE Control Syst. Lett.*, vol. 5, pp. 565–570, Apr. 2021.

[27] R. Seeber, "Structural and quantitative assessment of the perspective observability of linear time-invariant systems," Ph.D. thesis, Inst. Automat. Control, Graz Univ. Technol., (in German), Oct. 2015.

[28] M. Gu, E. Mengi, M. L. Overton, J. Xia, and J. Zhu, "Fast methods for estimating the distance to uncontrollability," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 2, pp. 477–502, 2006.

[29] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[30] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[31] Z. Wen and W. Yin, "A feasible method for optimization with orthogonality constraints," *Math. Program.*, vol. 142, no. 1–2, pp. 397–434, 2013.

[32] P. H. Schönemann, "On the formal differentiation of traces and determinants," *Multivariate Behav. Res.*, vol. 20, no. 2, pp. 113–139, 1985.

[33] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

[34] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, 1966.

[35] K. Hoffman and R. Kunze. *Linear Algebra*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice Hall, 1971.