A Trisection Algorithm for Estimating Distance Measures for Strong Observability and Strong Detectability

Roland Falkensteiner [iD], Richard Seeber [iD], and Martin Horn [iD], *Member, IEEE*

Abstract—**For reliably analyzing the properties of strong observability and strong detectability of a system, continuous distance measures can be used. When calculating these measures, it is necessary to find the global minimum of a nonconvex target function. The main contribution of this article is an optimization algorithm that guarantees to find this global minimum in a fast and efficient way by exploiting the special structure of the optimization problem. Using this optimization algorithm, the distance measures can be reliably calculated. The numerical properties and the usefulness of the algorithm in practical applications are illustrated by means of a numerical example.**

Index Terms—**Linear systems, mechatronics, observability measures, optimization algorithms.**

I. INTRODUCTION

In real-world control applications, unknown external disturbances are often present. The correct estimation of internal system states in presence of disturbances and the reconstruction of these disturbances is of particular importance in applications, like fault detection or robust control. Many external disturbances can be modeled as unknown inputs. For state estimation in presence of unknown inputs, strong observability or strong detectability, as defined in [1], are essential system properties. Standard criteria for these properties have binary character and their reliable numerical evaluation is difficult, but important, for example, in sensor placement [2]–[4]. Therefore, measures for reliable evaluation of strong observability and strong detectability are of major interest.

For the evaluation of standard observability, or its dual controllability, of a system, a similar problem has been extensively studied. Several continuous measures for observability and controllability have been proposed, see [5] for an overview. These continuous measures can be categorized into modal, energy, and distance measures. Modal measures have been introduced in [6], and give information about the controllability or observability of every mode separately. Energy measures analyze how much energy is necessary to bring the system into a desired state, see [7]. Distance measures have been introduced in [8], and yield the smallest perturbation in system matrices rendering the system uncontrollable or unobservable.

In [9], the idea of distance measures was extended to systems with unknown inputs and distance measures, for strong observability and strong detectability, was introduced. These measures give the distance to the closest non-strongly observable or non-strongly detectable system, which is calculated by minimizing the smallest singular value of the Rosenbrock matrix. The target function of this minimization problem is nonconvex and can have multiple local minima. However, for the correct determination of the measures, knowledge of the global minimum is absolutely necessary.

For the calculation of the distance measures for standard controllability and observability, special algorithms have been proposed. In [10], an iterative algorithm is presented, where estimations for the smallest singular value and the corresponding singular vector were refined in every iteration step. This algorithm only finds the local minimum and the results depend on the initial conditions. In [11], an algorithm for finding the minimum along a straight line using a bisection algorithm was proposed. In combination with an appropriate grid, this algorithm is able to find the global minimum. The algorithm in [12] uses a similar approach, but directly determines the global minimum by a trisection algorithm with the so-called Gu's test as condition. In [13], a quasi-Newton optimization is presented that finds a minimum and uses Gu's test to check if it is the global minimum.

For the calculation of the distance measures for strong observability and strong detectability, as presented in [9], special algorithms have not been studied yet. Several standard optimization algorithms can be used, for example, grid search [14], quasi-Newton optimization [13], [15], or genetic algorithms [16]. However, these optimization algorithms do not guarantee to find the global minimum, and high computational costs are necessary for increasing the degree of confidence in a solution. A concept closely related to the distance measures for strong observability and strong detectability is the minimum phase radius, as defined in [17]. In [18], an algorithm for computing this radius is presented, which also does not guarantee to find the global minimum of the considered target function.

In this article, an optimization algorithm for the reliable and efficient calculation of measures for strong observability and strong detectability is presented, which guarantees to find the *global* minimum of the target function. This algorithm is based on a trisection approach in combination with Gu's test, as presented in [12] and [19], for standard controllability, and extends these ideas to the strong observability case. The properties of the algorithm are discussed and a numerical example is given, in order to show its practical usefulness.

This article is structured as follows. Section II gives preliminaries. Section III introduces the theoretical foundation for the development of the optimization algorithm. Section IV gives the optimization algorithms based on the previous theorems. In Section V, a numerical example is presented.

II. PRELIMINARIES

Consider a linear system

$$\Sigma : \quad \begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{E}\mathbf{w} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{F}\mathbf{w} \end{aligned} \tag{1}$$

with states $\mathbf{x} \in \mathbb{R}^n$, unknown inputs $\mathbf{w} \in \mathbb{R}^q$, outputs $\mathbf{y} \in \mathbb{R}^p$, and $\mathbf{A}, \mathbf{E}, \mathbf{C}$, and $\mathbf{F}$ being constant matrices of appropriate dimensions. It is assumed that no prior knowledge about the unknown inputs $\mathbf{w}$ is available, that $\operatorname{rank} \mathbf{E} = q$, and that more measurements than unknown inputs are available ($p \geq q$). Without loss of generality, any known input is omitted. The matrices describing system Σ are combined in the matrix

$$\mathbf{S} = \begin{bmatrix} \mathbf{A} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix}. \tag{2}$$

The operators $\mathbf{M}^{\mathrm{H}}$, $\|\mathbf{M}\|$, and $\sigma_{\min}(\mathbf{M})$ denote the complex conjugate transpose, the spectral norm, and the smallest singular value of a matrix $\mathbf{M}$, respectively. The complex conjugate of $c \in \mathbb{C}$ is denoted by $\bar{c}$ and its real part by $\operatorname{Re}(c)$. Furthermore, the complex right open-half-plane is denoted by $\mathbb{C}^+ = \{c \in \mathbb{C} | \operatorname{Re}(c) > 0\}$.

A. Strong Observability and Strong Detectability

Strong observability or *strong detectability* were introduced in [1], and are necessary system properties for state estimation in the presence of unknown inputs. These properties are connected to the transmission zeros of a system. Transmission zeros are the values $\lambda_0 \in \mathbb{C}$, for which the matrix

$$\mathbf{R}(\mathbf{S}, \lambda) = \begin{bmatrix} \mathbf{A} - \lambda \mathbf{I} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix} \tag{3}$$

satisfies $\operatorname{rank}(\mathbf{R}(\mathbf{S}, \lambda_0)) < n + q$, as defined in [20]. According to [1], a system Σ is *strongly detectable* if and only if (iff) it has no transmission zeros λ_0 in the positive complex half plane $\mathbb{C}^+$, i.e., iff $\operatorname{Re}(\lambda_0) < 0$ holds for all λ_0, where $\mathbf{R}(\mathbf{S}, \lambda_0)$ is rank deficient. A system Σ is *strongly observable* iff it has no transmission zeros at all.

B. Distance Measures for Strong Observability and Strong Detectability

The rank condition for strong observability and strong detectability has a binary character, and therefore, its reliable numerical evaluation is difficult. In particular, numerical errors while checking the conditions or small parametric perturbations of the system matrices could change the result. In order to circumvent these issues, the continuous measures

$$\mu_{\mathrm{SO}}(\mathbf{S}) = \inf_{\mathbf{\Delta}} \{ \|\mathbf{\Delta}\| : (\mathbf{S} + \mathbf{\Delta}) \text{ non-SO} \} \tag{4}$$

$$= \inf_{\lambda \in \mathbb{C}} \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda)) \tag{5}$$

$$\mu_{\mathrm{SD}}(\mathbf{S}) = \inf_{\mathbf{\Delta}} \{ \|\mathbf{\Delta}\| : (\mathbf{S} + \mathbf{\Delta}) \text{ non-SD} \} \tag{6}$$

$$= \inf_{\lambda \in \mathbb{C}^+} \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda)) \tag{7}$$

and their calculation are introduced in [9, Theorem 2]. The measures are defined as the minimum norm parametric perturbation $\mathbf{\Delta}$ that renders the system non-strongly observable or non-strongly detectable. This perturbation is calculated as the smallest singular value of $\mathbf{R}(\mathbf{S}, \lambda)$ in the corresponding part of the complex plane and can be considered the distance from system $\mathbf{S}$ to a perturbed system with a transmission zero at λ.

In the distance measures calculation, the minimization problem (5) or (7) needs to be solved. The corresponding target function

$$f(\lambda) = \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda)) \tag{8}$$

can have multiple local minima, because multiple systems with different transmission zeros λ but equal distance to Σ may exist. Therefore, $f(\lambda)$ is a nonconvex function, which can include *local* minima that give only upper bounds for the *global* minimum. For the correct calculation of the distance measures, it is therefore necessary to have an optimization algorithm, which reliably finds the *global* minimum of $f(\lambda)$.

III. THEORETICAL FOUNDATION FOR A GLOBAL OPTIMIZATION ALGORITHM

In order to reliably find the *global* minimum of $f(\lambda)$ despite its nonconvexity, an algorithm is proposed to verify the existence of successively tighter, lower, or upper bounds in every iteration step. In this section, first a theorem for the existence of such bounds is given, then a proposition for their verification is derived, and finally a corollary of this proposition is presented for a lower dimension case.

A. Fundamental Theorem

The following theorem allows to check if a pair of values $(q, q - \delta)$ contains a lower or upper bound of the global minimum.

Theorem 1: Consider an area $\mathcal{D}_\gamma = \{\lambda \in \mathbb{C} | \operatorname{Re}(\lambda) \geq \gamma\}$ with $\gamma \in (-\infty, 0] \cup \{-\infty\}$ and $q \in \mathbb{R}$. Let

$$\mu^*(\mathbf{S}) = \inf_{\lambda \in \mathcal{D}_\gamma} f(\lambda) \quad \mu_\gamma(\mathbf{S}) = \inf_{\lambda \in \partial \mathcal{D}_\gamma} f(\lambda) \tag{9}$$

denote the minimum of $f(\lambda)$ in $\mathcal{D}_\gamma$ and the minimum of $f(\lambda)$ along the border $\partial \mathcal{D}_\gamma$, respectively. Assume that $\mu_\gamma(\mathbf{S}) \geq q \geq \mu^*(\mathbf{S})$ holds. If $\delta \in \mathbb{R}$ satisfies $0 \leq \delta \leq q - \mu^*(\mathbf{S})$, then there exists a $\lambda_{\mathrm{int}} = \alpha_{\mathrm{int}} + j\beta_{\mathrm{int}}$ with $\alpha_{\mathrm{int}} \geq (\gamma + \delta)$ such that

$$f(\lambda_{\mathrm{int}} - \delta) = q \quad f(\lambda_{\mathrm{int}} + \delta) = q \tag{10}$$

holds.

The following implications are obtained from Theorem 1 and bounds for μ^* can be derived, provided that $q \leq \mu_\gamma$ holds:

$$\text{Theorem 1}: q - \delta \geq \mu^*$$

$$\Downarrow$$

$$(10): \begin{array}{l} f(\alpha_{\mathrm{int}} + j\beta_{\mathrm{int}} - \delta) = q \\ f(\alpha_{\mathrm{int}} + j\beta_{\mathrm{int}} + \delta) = q \end{array} \Rightarrow q \geq \mu^*.$$

If λ_{int} exist, for which (10) holds, then

$$U = q \geq \mu^* \tag{11}$$

yields an upper bound. If no such point exists, then the assumptions of Theorem 1 are contradicted, and

$$L = q - \delta < \mu^* \tag{12}$$

yields a lower bound.

Remark: Note that Theorem 1 is formulated such that bounds for either μ_{SO} or μ_{SD} can be deduced by means of a parameter γ. For the determination of μ_{SO}, set $\gamma = -\infty$ to ensure that $\partial \mathcal{D}_\gamma$ includes $\mathbb{C}$. For the determination of μ_{SD}, set $\gamma = 0$ to restrict $\mathcal{D}_\gamma$ to $\mathbb{C}^+$.
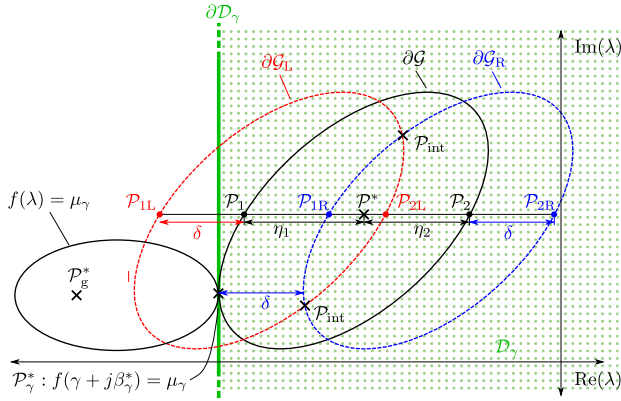
Fig. 1. Sketch for Theorem 1.

Proof: A full proof of Theorem 1 is given in Appendix A. In the following, a short summary of the proof is presented. Fig. 1 shows the geometrical relations of the proof as a visual aid.

Due to the continuity of $f(\lambda)$, it is possible to find a closed curve $\partial \mathcal{D}_\gamma$, where $f(\lambda) = q$ holds with the point of minimum $\mathcal{P}^*$, at which $f(\lambda^*) = \mu^*(\mathbf{S})$ holds, in its interior. Consider two points $\mathcal{P}_1$ and $\mathcal{P}_2$ located on $\partial \mathcal{D}_\gamma$ with the same imaginary part as $\mathcal{P}^*$. There exists a lower bound for the distances η_1 and η_2 from the two points $\mathcal{P}_1$ and $\mathcal{P}_2$ to $\mathcal{P}^*$. The lower bound can be derived from the fact that the change of function value of $f(\lambda)$ is bounded by $f(\lambda^* + \eta) - f(\lambda^*) \leq |\eta|$ as can be shown by using Weyl's inequality [21]. By inserting the function values for $f(\lambda^* + \eta)$ and $f(\lambda^*)$, the lower bounds are given by $\eta_1 \geq q - \mu_{\mathrm{SO}}(\mathbf{S})$ and $\eta_2 \geq q - \mu_{\mathrm{SO}}(\mathbf{S})$, respectively. When $\partial \mathcal{G}$ is shifted left and right by δ, new versions $\partial \mathcal{G}_{\mathrm{L}}$ and $\partial \mathcal{G}_{\mathrm{R}}$ are obtained, respectively. There exist points of intersection between $\partial \mathcal{G}_{\mathrm{L}}$ and $\partial \mathcal{G}_{\mathrm{R}}$ at $\mathcal{P}_{\mathrm{int}} = (\alpha_{\mathrm{int}}, \beta_{\mathrm{int}})$, where (10) holds since $\delta \leq q - \mu_{\mathrm{SO}}(\mathbf{S})$ holds. The right-shifted version $\partial \mathcal{G}_{\mathrm{R}}$ is strictly in $\mathcal{D}_\gamma$, with $\mathrm{Re}(\partial \mathcal{G}_{\mathrm{R}}) \geq (\gamma + \delta)$, as shown in Fig. 1, and therefore any point of intersection $\mathcal{P}_{\mathrm{int}}$ is in $\mathcal{D}_\gamma$ with $\alpha_{\mathrm{int}} > (\gamma + \delta)$. ■

B. Verification of Intersection

The challenge when using Theorem 1 for finding bounds on μ_{SO} and μ_{SD} is now to verify if a point of intersection $\mathcal{P}_{\mathrm{int}}$ does exist in order to show that (10) holds. For this verification the following proposition is introduced.

Proposition 2: Let $\alpha, \delta \in \mathbb{R}$ and $\gamma \in (-\infty, 0] \cup \{-\infty\}$. Consider the matrices

$$\mathbf{M}(\alpha) = \begin{bmatrix} -\alpha \mathbf{I} + \mathbf{A} - \mathbf{EGF}^{\mathrm{T}}\mathbf{C} & -\mathbf{EGE}^{\mathrm{T}} - \mathbf{I} \\ q^2\mathbf{I} - \mathbf{C}^{\mathrm{T}}\mathbf{C} + \mathbf{C}^{\mathrm{T}}\mathbf{FGF}^{\mathrm{T}}\mathbf{C} & \alpha \mathbf{I} - \mathbf{A}^{\mathrm{T}} + \mathbf{C}^{\mathrm{T}}\mathbf{FGE}^{\mathrm{T}} \end{bmatrix} \quad (13)$$

$$\mathbf{N} = \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (14)$$

$$\mathbf{\Pi}(\delta) = \mathbf{I} \otimes \mathbf{M}(\delta) + \mathbf{M}(-\delta) \otimes \mathbf{I} \quad (15)$$

$$\mathbf{\Gamma} = -\mathbf{I} \otimes \mathbf{N} - \mathbf{N} \otimes \mathbf{I}. \quad (16)$$

There exists $\lambda_{\mathrm{int}} = \alpha_{\mathrm{int}} + j\beta_{\mathrm{int}}$ satisfying (10), iff the matrix pair $\mathbf{\Pi}(\delta)$, $\mathbf{\Gamma}$ has a purely real-valued eigenvalue $\alpha_{\mathrm{int}} \geq (\gamma + \delta)$, i.e., $\det(\mathbf{\Pi}(\delta) - \alpha_{\mathrm{int}}\mathbf{\Gamma}) = 0$, and the corresponding matrices

$$\mathbf{M}_{\mathrm{a}} = \mathbf{M}(\alpha_{\mathrm{int}} + \delta) \qquad \mathbf{M}_{\mathrm{b}} = \mathbf{M}(\alpha_{\mathrm{int}} - \delta) \quad (17)$$

share a common purely imaginary eigenvalue $j\beta_{\mathrm{int}}$.

Proof: A proof is given in Appendix B.

Remark: For the verification of the existence of $\mathcal{P}_{\mathrm{int}}$ and consequently λ_{int} satisfying (10) in Theorem 1, the following steps are necessary. First, check if the matrix pair $\mathbf{\Pi}(\delta)$, $\mathbf{\Gamma}$ has a purely real-valued eigenvalue α_i, for which $\alpha_i \geq (\gamma + \delta)$ holds. Second, for every α_i, calculate the eigenvalues of $\mathbf{M}(\alpha_i + \delta)$ and $\mathbf{M}(\alpha_i - \delta)$. If they share a common purely imaginary eigenvalue $j\beta_{\mathrm{int}}$ for a $\alpha_{\mathrm{int}} = \alpha_i$, then there exists a point $\mathcal{P}_{\mathrm{int}} = (\alpha_{\mathrm{int}}, \beta_{\mathrm{int}})$ and λ_{int} satisfying (10).

C. Minimization Along the Border

Applying Theorem 1 requires the knowledge of μ_γ, the minimum value of $f(\lambda)$ on the border $\partial \mathcal{D}_\gamma$ located at $\mathcal{P}^*_\gamma = (\alpha^*_\gamma + j\beta^*_\gamma)$. It is necessary to find an algorithm for minimizing $f(\lambda)$ along $\partial \mathcal{D}_\gamma$, i.e., a line with $\lambda = \gamma + j\beta$ with $\beta \in \mathbb{R}$. For the construction of such an algorithm, a corollary for an upper bound of μ_γ can be deduced from Proposition 2 by inserting $\alpha_{\mathrm{int}} = \gamma$ and $\delta = 0$. In case $\gamma = -\infty$ is chosen, $\mu_\gamma = \sigma_{\min}(\mathbf{F})$ holds, as given in Appendix A1. Otherwise, the following corollary can be used.

Corollary 3: Let $\gamma \in (-\infty, 0]$. There exists $\lambda_{\mathrm{int}} = \gamma + j\beta_{\mathrm{int}}$, for which (10) holds, iff the matrix $\mathbf{M}(\gamma)$, as defined in (13), has a purely imaginary eigenvalue $j\beta_{\mathrm{int}}$.

Proof: Corollary 3 is a special case of Proposition 2, where $\alpha_{\mathrm{int}} = \gamma$ and $\delta = 0$. Then, $\mathbf{M}_{\mathrm{a}} = \mathbf{M}_{\mathrm{b}} = \mathbf{M}(\gamma)$ holds and $\mathbf{M}_{\mathrm{a}}$ and $\mathbf{M}_{\mathrm{b}}$ share all eigenvalues. Since $\mathbf{M}(\gamma)$ has a purely imaginary eigenvalue at $j\beta$, $\det(\mathbf{\Pi}(0) - \gamma\mathbf{\Gamma}) = 0$ holds, because

$$\begin{aligned} &\det(\mathbf{\Pi}(0) - \gamma\mathbf{\Gamma}) \\ &= \det(\mathbf{I} \otimes \mathbf{M}(0) + \mathbf{M}(0) \otimes \mathbf{I} + \mathbf{I} \otimes \gamma\mathbf{N} + \gamma\mathbf{N} \otimes \mathbf{I}) \\ &= \det(\mathbf{I} \otimes (\mathbf{M}(0) + \gamma\mathbf{N}) + (\mathbf{M}(0) + \gamma\mathbf{N}) \otimes \mathbf{I}) \\ &= \det(\mathbf{I} \otimes \mathbf{M}(\gamma) + \mathbf{M}(\gamma) \otimes \mathbf{I}) \end{aligned} \quad (18)$$

and

$$\begin{aligned} &(\mathbf{I} \otimes \mathbf{M}(\gamma) + \mathbf{M}(\gamma) \otimes \mathbf{I})(\mathbf{v} \otimes \mathbf{w}) \\ &= (\mathbf{v} \otimes \mathbf{M}(\gamma)\mathbf{w}) + (\mathbf{M}(\gamma)\mathbf{v} \otimes \mathbf{w}) \\ &= j\beta(\mathbf{v} \otimes \mathbf{w}) - j\beta(\mathbf{v} \otimes \mathbf{w}) = 0 \end{aligned} \quad (19)$$

with eigenvectors $\mathbf{w}$ and $\mathbf{v}$ of $\mathbf{M}(\gamma)$ for the eigenvalues $j\beta$ and $-j\beta$, respectively.

Since $\delta = 0$ holds, (10) reduces to

$$f(\gamma + j\beta_{\mathrm{int}}) = q \quad (20)$$

and can be used as a bound for a bisection algorithm.

IV. ALGORITHMS

In this section, a trisection algorithm for calculation of μ_{SO} and μ_{SD} is presented. The steps of the algorithm are described, the conditions for the bounds and their values are given, and the initial condition of the algorithm is discussed. In addition, a bisection algorithm for calculation of μ_γ is given.

A. Trisection Algorithm for the Calculation of Strong Observability and Strong Detectability

Using the results presented in Section III, an algorithm for finding the global minimum μ^* of $f(\lambda)$, given by Algorithm 1, can be constructed by the following steps. First, a possible range $\mu^* \in [L_0, U_0]$ is initialized, such that it includes the global minimum. The definition of L_0 and U_0 is given in the next paragraph. Afterwards, in every iteration step i, the parameters

$$\delta = \frac{1}{3}(U_{i-1} - L_{i-1}) \qquad q = \frac{2}{3}(U_{i-1} - L_{i-1}) \quad (21)$$

Algorithm 1: Estimation of μ_{SO} or μ_{SD}.

initialize for μ_{SO}: $\gamma \leftarrow -\infty$, $L_0 \leftarrow 0$, $U_0 \leftarrow \sigma_{\min}(\mathbf{F})$, $i = 1$
       for μ_{SD}: $\gamma \leftarrow 0$, $L_0 \leftarrow 0$, $U_0 \leftarrow \mu_\gamma$, $i = 1$
repeat
  $\delta \leftarrow \frac{1}{3}(U_{i-1} - L_{i-1})$
  $q \leftarrow \frac{2}{3}(U_{i-1} - L_{i-1})$
  if conditions for Proposition verified **then**
    $L_i \leftarrow L_{i-1}$
    $U_i \leftarrow q$
  else
    $L_i \leftarrow q - \delta$
    $U_i \leftarrow U_{i-1}$
  end if
until $U_i - L_i < \epsilon$

Algorithm 2: Estimation of μ_γ.

initialize $L_0 \leftarrow 0$, $U_0 \leftarrow \sigma_{\min}(F)$, $i = 1$
repeat
  $q \leftarrow \frac{1}{2}(U_{i-1} - L_{i-1})$
  if conditions for Corollary 3 hold **then**
    $L_i \leftarrow L_{i-1}$
    $U_i \leftarrow q$
  else
    $L_i \leftarrow q$
    $U_i \leftarrow U_{i-1}$
  end if
until $U_i - L_i < \epsilon$

---

are chosen, such that the resulting bounds split the range of the previous step in thirds, hence, the name trisection algorithm. By verifying if a point of intersection exists using Proposition 2, either (11) or (12) can be used for adapting L_i or U_i accordingly. The stopping criterion for the algorithm is given by $U_i - L_i < \epsilon$. The range $[L_i, U_i]$ always includes μ^*, and therefore it can be guaranteed to find μ^* with a precision of ϵ.

The values L_0 and U_0 are chosen as $L_0 = 0$ and $U_0 = \mu_\gamma$, respectively, because $0 \le \mu^* \le \mu_\gamma$ always holds. Furthermore, selecting $U_0 = \mu_\gamma$ ensures that the condition $q \le \mu_\gamma$ in Theorem 1 holds in every iteration step of the algorithm.

The value of γ can be chosen, such that Algorithm 1 calculates either μ_{SO} or μ_{SD}. For μ_{SO}, select $\gamma = -\infty$. In this case, $\mu_\gamma = \lim_{|\lambda| \to \infty} f(\lambda) = \sigma_{\min}(\mathbf{F})$ holds, as given in Appendix A1. For μ_{SD}, select $\gamma = 0$. The calculation of μ_γ is discussed in the next section.

The possible range for μ^* is reduced by a factor of $2/3$ in every iteration step, and can be calculated as $U_i - L_i = (U_0 - L_0)(2/3)^i$. Starting from this equation and inserting $L_0 = 0$, the necessary number of iteration steps for finding the minimum μ^* with precision ϵ is given by

$$i_{\mathrm{tri}} = \left\lceil \frac{\ln(\epsilon/U_0)}{\ln(2/3)} \right\rceil. \tag{22}$$

Remark: A reduction of computation cost is possible by reformulating the generalized eigenvalue problem of size $(2n)^2 \times (2n)^2$ in (51). Similar to the considerations in [19], (51) can be reduced to a $2n \times 2n$ standard eigenvalue problem by splitting $\mathbf{X}$ up into submatrices, and further exploiting symmetry properties. This modification is given in Appendix B1.

B. Bisection Algorithm for Computation of Initial Bound

From Corollary 3, it is possible to construct an algorithm for finding the minimum μ_γ given by Algorithm 2. First, a possible range $\mu_\gamma \in [L_0, U_0]$ is defined, with $L_0 = 0$ and $U_0 = \sigma_{\min}(\mathbf{F})$ to ensure that μ_γ is included, as given in Appendix A1. Afterwards, in every iteration step i a parameter

$$q = \frac{1}{2}(U_{i-1} - L_{i-1}) \tag{23}$$

is chosen, splitting the range of the previous step in half. For deriving the bounds for the bisection algorithm, the following set of equivalences can be drawn from Corollary 3:

$$\mu_\gamma \le q < \sigma_{\min}(\mathbf{F})$$

$$\Updownarrow$$

$$(20): f(\gamma + j\beta_{\mathrm{int}}) = q$$

$$\Updownarrow$$

Corollary 3 : $j\beta_{\mathrm{int}}$ purely imaginary EV of $\mathbf{M}(\gamma)$.

If $\mathbf{M}(\gamma)$ has a purely imaginary eigenvalue, then $U_i = q \ge \mu_\gamma$ is a new upper bound for μ_γ. Conversely, $L_i = q < \mu_\gamma$ is a new lower bound for μ_γ, if $\mathbf{M}(\gamma)$ does not have a purely imaginary eigenvalue. The stopping criterion for the algorithm is given by $U_i - L_i < \epsilon$. Therefore, it can be guaranteed to find μ_γ with a precision of ϵ.

The possible range for μ_γ is reduced by a factor of $1/2$ in every iteration step. Following the same considerations as in the previous section, the necessary number of iteration steps for finding the minimum along the imaginary axis μ_γ with precision ϵ is given by

$$i_{\mathrm{bi}} = \left\lceil \frac{\ln(\epsilon/U_0)}{\ln(1/2)} \right\rceil. \tag{24}$$

V. NUMERICAL EXAMPLE

In order to show the practical usefulness of the presented algorithms, they are implemented in MATLAB and tested for a numerical example. For comparison with a state-of-the-art global optimization algorithm, the genetic algorithm ga, included in the MATLAB Global Optimization Toolbox, is used. In order to find the global minimum with acceptable confidence, an initial population size of 10^3 and a minimum improvement of the fitness function of 10^{-10} are chosen for the ga algorithm.

As a practical application, the strong observability and strong detectability of the steel beam, shown in Fig. 2, are analyzed for different sensor configurations. One end of the beam is fixed, and on the free end an unknown force F is acting. At two locations z_1 and z_2, the displacement, the acceleration, or both can be measured.

A linear modal model of the cantilevered beam, as presented for example in [22], is given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\boldsymbol{\Omega}^2 & -\boldsymbol{\Gamma} \end{bmatrix}, \qquad \mathbf{E} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Phi}_{\mathrm{w}} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\Phi}_{\mathrm{y}}^{\mathsf{T}} & \mathbf{0} \\ -\boldsymbol{\Phi}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Omega}^2 & -\boldsymbol{\Phi}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Gamma} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Phi}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}_{\mathrm{w}} \end{bmatrix} \tag{25}$$

where $\boldsymbol{\Omega} = \mathrm{diag}(\omega_{0,1}, \omega_{0,2}, \ldots, \omega_{0,k})$ is a diagonal matrix containing the resonance frequencies $\omega_{0,j}$, $(j = 1, \ldots, k)$, $\boldsymbol{\Gamma} =$
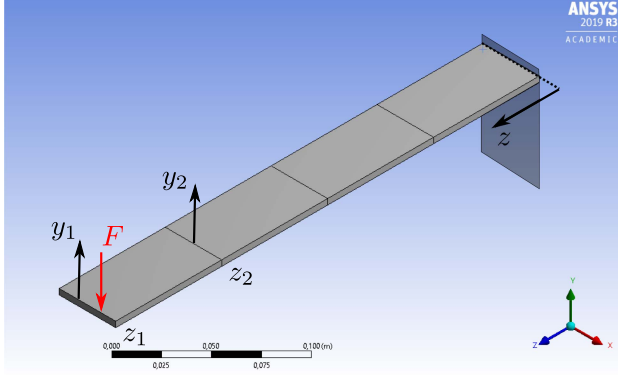
Fig. 2. Beam with positions of applied force F and measurements y (image used courtesy of ANSYS, Inc.).

Fig. 3. Mesh plot of $f(\lambda)$, closed curves $f(\lambda) = q_i$, and position of minimum for beam with acceleration and position measurement at z_1.

TABLE I
MODAL PARAMETERS FOR SYSTEM (25)

mode	$\omega_{0,j}$ in $\frac{\text{rad}}{\text{s}} 10^3$	γ_j in $\frac{1}{\text{s}}$	Φ_w	φ_{y1}	φ_{y2}
1	0.244	0.063	3.27	3.27	2.12
2	1.72	0.063	-3.45	-3.45	0.77
3	4.86	0.067	5.12	5.12	-3.76
4	6.61	0.070	3.34	3.34	-3.06
5	21.0	0.133	-5.89	-5.89	-0.59
6	27.7	0.185	-1.68	-1.68	2.26
7	57.0	0.581	5.69	5.69	4.53
8	112	2.064	-3.96	-3.96	-4.84

TABLE II
μ_{SO}, μ_{SD}, AND CORRESPONDING SPEEDUP OF CALCULATION TIMES GIVEN BY THE RATIO $\frac{t_{\text{gen}}}{t_{\text{tri}}}$

Acc.	Disp.	μ_{SO}	$\frac{t_{\text{gen}}}{t_{\text{tri}}}$	μ_{SD}	$\frac{t_{\text{gen}}}{t_{\text{tri}}}$
z_1	-	$6.88 \cdot 10^{-11}$	2.62	$9.06 \cdot 10^{-11}$	4.07
z_1	z_1	$6.88 \cdot 10^{-11}$	3.38	$5.43 \cdot 10^{-6}$	5.41
z_1	z_2	$2.11 \cdot 10^{-4}$	3.17	$2.11 \cdot 10^{-4}$	3.80
z_1, z_2	z_1, z_2	$3.22 \cdot 10^{-4}$	3.21	$1.8 \cdot 10^{-3}$	3.70

$\text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_k)$ is a diagonal matrix containing the corresponding modal damping ratios γ_j, $(j = 1, \ldots, k)$, and k is the number of modes considered in the model. The matrices $\boldsymbol{\Phi}_{\text{w}} \in \mathbb{R}^{k \times 1}$ and $\boldsymbol{\Phi}_{\text{y}}^{\text{T}} = \begin{bmatrix} \boldsymbol{\varphi}_{\text{y1}} & \boldsymbol{\varphi}_{\text{y2}} \end{bmatrix}^{\text{T}} \in \mathbb{R}^{2 \times k}$ collect the mode shape vectors at the points z_1 and z_2, respectively. The mode shapes describe how sensitive the system is for excitation and measurement at a specific point. The beam is modeled by the first $k = 8$ modes and the resulting system is of order $n = 16$. The modal parameters for the system (25) are determined using the finite element simulation program Ansys Academic Student 2019 R3 and are given in Table I.

As an illustration of the iterative calculation process for the beam with acceleration and position measurement at z_1, Fig. 3 shows the mesh plot of $f(\lambda)$ and the closed curves with $f(\lambda) = q_i$. It can be observed how the closed curves first include multiple local minima but get closer and closer to the global minimum in every iteration step.

Using algorithms presented in Section IV with a precision of $\epsilon = 10^{-10}$, the distance measures for strong observability and strong detectability are calculated. The results are given in Table II. If only an
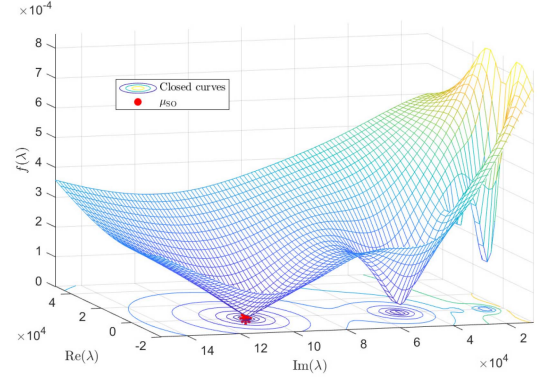
acceleration sensor is used, the system is neither strongly observable nor strongly detectable. If an acceleration sensor and a displacement sensor are used at the same position z_1, then the system is not strongly observable but strongly detectable. If acceleration sensors and displacement sensors are used at different locations, the system is strongly observable and strongly detectable. All these properties are correctly indicated by the corresponding values of μ_{SO} and μ_{SD}.

For the calculation of μ_{SO}, 70 iteration steps are necessary in every sensor configuration, as can be determined by (22). The number of iteration steps for the calculation of μ_{SD} ranges from 28 to 43 iteration steps because it depends on the start value given by the minimum along the imaginary axis. This minimum is first calculated by using the bisection algorithm in 42 iteration steps, as can be calculated using (24). In Table II, the ratio between the calculation times on off-the-shelf hardware (Notebook, Intel Core i5-8265 U, 1.80 GHz) of the trisection algorithm, with reduced eigenvalue problem t_{tri} and of the genetic algorithm t_{gen}, is given for different sensor configurations. It can be seen that the trisection algorithm finds the minimum roughly two to five times faster than the genetic algorithm.

VI. CONCLUSION

This article presented an algorithm for reliable and efficient calculation of measures for strong observability and strong detectability of linear and time-invariant systems. The challenge in calculating these measures was to solve a minimization problem with nonconvex target function. The main contribution of this article was an algorithm, which guarantees to find the *global* minimum of the target function for a given precision within a fixed number of iteration steps. The algorithm was constructed by exploiting the special structure of the target function. Using even more sophisticated numerical techniques, further reduction of the computational load may be possible. The computationally most expensive operation was finding the real eigenvalues of (51) or (59). In [19], a divide and conquer approach for efficient calculation of real eigenvalues was presented, which may be used in the presented algorithm as well.

APPENDIX

A. Proof of Theorem 1

The singular value decomposition is continuous in the change of matrix entries, and therefore $f(\lambda)$ is continuous in λ. Due to the fact

that

$$\mu_\gamma(\mathbf{S}) \geq q \geq \mu^*(\mathbf{S}) \tag{26}$$

holds, it follows that there exists a λ_1, such that $f(\lambda_1) = q$.

Since $f(\lambda)$ is continuous and (26) holds, there exist closed curves that satisfy $f(\lambda) = q$. The point of minimum $\mathcal{P}^* = (\alpha^*, \beta^*)$ in $\mathcal{D}_\gamma$, at which $f(\lambda^*) = f(\alpha^* + j\beta^*) = \mu^*$ holds, must lie within one of these closed curves. Let $\partial\mathcal{G}$ denotes the closed curve with the smallest area $\mathcal{G}$ with $\mathcal{P}^*$ in its interior. From (26), it follows that all points on $\partial\mathcal{G}$ are in $\mathcal{D}_\gamma$, as shown in Fig. 1. On a line parallel to the real axis with $\mathcal{P}^*$ on it, there exist two points $\mathcal{P}_1 = (\alpha^* - \eta_1, \beta^*)$ and $\mathcal{P}_2 = (\alpha^* + \eta_2, \beta^*)$ on $\partial\mathcal{G}$ with $\eta_1, \eta_2 \in \mathbb{R}, \eta_1 > 0$, and $\eta_2 > 0$. If the line crosses $\partial\mathcal{G}$ multiple times, then $\mathcal{P}_1$ and $\mathcal{P}_2$ are chosen as the points closest to $\mathcal{P}^*$. For the points $\mathcal{P}_1$ and $\mathcal{P}_2$

$$f(\lambda^* - \eta_1) = f(\lambda^* + \eta_2) = q \tag{27}$$

holds. The point $\mathcal{P}^*$ lies within $\mathcal{G}$ and also on the line segment between $\mathcal{P}_1$ and $\mathcal{P}_2$. Therefore, any point on this line segment lies within $\mathcal{G}$.

The change of singular values of a matrix $\mathbf{T} \in \mathbb{R}^{m \times n}$ is bounded by the norm of the perturbation $\mathbf{U} \in \mathbb{R}^{m \times n}$

$$|\sigma_k(\mathbf{T} + \mathbf{U}) - \sigma_k(\mathbf{T})| \leq \sigma_1(\mathbf{U}) = \|\mathbf{U}\| \tag{28}$$

as described by Weyl's inequality, see for example [21]. If

$$\mathbf{T} = \mathbf{R}(\mathbf{S}, \lambda) = \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} \eta\mathbf{I} & 0 \\ 0 & 0 \end{bmatrix} \tag{29}$$

with $\eta \in \mathbb{R}$ is inserted in (28) and the smallest singular value is considered, then

$$|\sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda + \eta)) - \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda))| \leq \left\| \begin{bmatrix} \eta\mathbf{I} & 0 \\ 0 & 0 \end{bmatrix} \right\| \tag{30}$$

$$\Leftrightarrow f(\lambda^* + \eta) - f(\lambda^*) \leq |\eta| \tag{31}$$

holds, and the change of function value is bounded by the norm of change of argument $|\eta|$. Combining (31) with the relation (27) for points $\mathcal{P}_1$ and $\mathcal{P}_2$, it is possible to give lower bounds for their distances η_1 and η_2 to the point of minimum $\mathcal{P}^*$ as

$$\eta_1 \geq q - \mu^*(\mathbf{S}) \qquad \eta_2 \geq q - \mu^*(\mathbf{S}). \tag{32}$$

If the curve $\partial\mathcal{G}$ is shifted left and right, i.e., in parallel to the imaginary axis, by a distance δ, two continuous closed curves

$$\partial\mathcal{G}_L = \{\lambda - \delta | \lambda \text{ on } \partial\mathcal{G}\} \tag{33a}$$

$$\partial\mathcal{G}_R = \{\lambda + \delta | \lambda \text{ on } \partial\mathcal{G}\} \tag{33b}$$

are obtained. On the curve shifted to the left $\partial\mathcal{G}_L$, there exist two points $\mathcal{P}_{1L} = (\alpha^* - \eta_1 - \delta, \beta^*)$ and $\mathcal{P}_{2L} = (\alpha^* + \eta_2 - \delta, \beta^*)$. On the curve shifted to the right $\partial\mathcal{G}_R$, there exist two points $\mathcal{P}_{1R} = (\alpha^* - \eta_1 + \delta, \beta^*)$ and $\mathcal{P}_{2R} = (\alpha^* + \eta_2 + \delta, \beta^*)$. The point $\mathcal{P}_{1L}$ is left of $\mathcal{P}_{1R}$ since their real parts satisfy

$$\mathrm{Re}(\lambda_{1L}) = \alpha^* - \eta_1 - \delta < \alpha^* - \eta_1 + \delta = \mathrm{Re}(\lambda_{1R}) \tag{34}$$

and therefore $\mathcal{P}_{1L}$ is in the exterior of $\partial\mathcal{G}_R$. Furthermore, $\mathcal{P}_{2L}$ is to the right of $\mathcal{P}_{1R}$, because

$$\mathrm{Re}(\lambda_{2L}) = \alpha^* + \eta_2 - \delta \geq \alpha^* - \eta_1 + \delta = \mathrm{Re}(\lambda_{1R}) \tag{35}$$

holds due to the fact that $\delta \leq q - \mu_{SO} \leq \eta_1$ and, analogously, $\delta \leq \eta_2$. From (34) and (35), it follows that two points $\mathcal{P}_{1L}$ and $\mathcal{P}_{2L}$ exist on

$\partial\mathcal{G}_L$, one in the interior and one in the exterior of $\partial\mathcal{G}_R$, respectively. Since $\partial\mathcal{G}_L$ and $\partial\mathcal{G}_R$ are continuous closed curves, there exists at least one point of intersection $\mathcal{P}_{int} = (\alpha_{int}, \beta_{int})$ of the two curves $\partial\mathcal{G}_L$ and $\partial\mathcal{G}_R$.

Since $\partial\mathcal{G}$ is strictly in $\mathcal{D}_\gamma$, also the right-shifted version $\partial\mathcal{G}_R$ is strictly in $\mathcal{D}_\gamma$, with $\mathrm{Re}(\partial\mathcal{G}_R) \geq (\gamma + \delta)$, as shown in Fig. 1. Therefore, any point of intersection $\mathcal{P}_{int}$ between $\partial\mathcal{G}_L$ and $\partial\mathcal{G}_R$ is in $\mathcal{D}_\gamma$ with $\alpha_{int} > (\gamma + \delta)$.

From the definition of the shifted curves (33), it follows that $\lambda_{int} + \delta$ and $\lambda_{int} - \delta$ are on $\partial\mathcal{G}$ and solutions for (10). Therefore, (10) has at least one solution λ_{int} in $\mathcal{D}_\gamma$ with $\alpha_{int} > (\gamma + \delta)$.

B. Limit of f

Lemma 4: For the function

$$f(\lambda) = \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda)) = \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix} \tag{36}$$

it holds that $\lim_{|\lambda| \to \infty} f(\lambda) = \sigma_{\min}(\mathbf{F})$.

Proof: Consider the transformation

$$\mathbf{Q}(\mathbf{S}, \lambda) = \begin{bmatrix} \sqrt{\frac{c}{\lambda}}\mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \sqrt{\frac{c}{\lambda}}\mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{c}{\lambda}\mathbf{A} - c\mathbf{I} & \sqrt{\frac{c}{\lambda}}\mathbf{E} \\ \sqrt{\frac{c}{\lambda}}\mathbf{C} & \mathbf{F} \end{bmatrix} \tag{37}$$

with $c = 2\sigma_{\min}(\mathbf{F})$ preserving rank $\mathbf{Q}(\mathbf{S}, \lambda) = \mathrm{rank}\,\mathbf{R}(\mathbf{S}, \lambda)$, if $\lambda \neq 0$. As $|\lambda|$ approaches infinity, (37) becomes

$$\lim_{|\lambda| \to \infty} \mathbf{Q}(\mathbf{S}, \lambda) = \mathbf{Q}_\infty(\mathbf{S}, \lambda) = \begin{bmatrix} -c\mathbf{I} & 0 \\ 0 & \mathbf{F} \end{bmatrix}. \tag{38}$$

The operations of finding the limit of the function and calculating the singular values can be interchanged, and it follows that $\lim_{|\lambda| \to \infty} f(\lambda) = \sigma_{\min}(\mathbf{Q}_\infty(\mathbf{S}, \lambda)) = \min\{c, \sigma_{\min}(\mathbf{F})\} = \sigma_{\min}(\mathbf{F})$ holds for the limit of $f(\lambda)$.

C. Proof of Proposition 2

From the definition of the singular value decomposition it follows that if $f(\lambda) = \sigma_{\min}(\mathbf{R}(\mathbf{S}, \lambda)) = q$ holds, then there exists a vector $\mathbf{z}$ such that

$$\mathbf{R}(\mathbf{S}, \lambda)^H \mathbf{R}(\mathbf{S}, \lambda)\mathbf{z} = q^2\mathbf{z} \tag{39}$$

holds. It is possible to split (39) up into the following two equations:

$$\begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & \mathbf{E} \\ \mathbf{C} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} \tag{40a}$$

$$\begin{bmatrix} \mathbf{A}^T - \bar{\lambda}\mathbf{I} & \mathbf{C}^T \\ \mathbf{E}^T & \mathbf{F}^T \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} q^2\mathbf{z}_1 \\ q^2\mathbf{z}_2 \end{bmatrix}. \tag{40b}$$

From the second line of (40a), $\mathbf{w} = \mathbf{C}\mathbf{z}_1 + \mathbf{F}\mathbf{z}_2$ can be calculated and inserted into the second line of (40b), giving

$$\mathbf{E}^T\mathbf{v} + \mathbf{F}^T\mathbf{C}\mathbf{z}_1 + \mathbf{F}^T\mathbf{F}\mathbf{z}_2 = q^2\mathbf{z}_2$$

$$\mathbf{z}_2 = -\mathbf{G}\mathbf{E}^T\mathbf{v} - \mathbf{G}\mathbf{F}^T\mathbf{C}\mathbf{z}_1 \tag{41}$$

where $\mathbf{G} = (\mathbf{F}^T\mathbf{F} - q^2\mathbf{I})^{-1}$. Since q satisfies $q < \sigma_{\min}(\mathbf{F})$ this inverse exists. Inserting $\mathbf{w}$ and $\mathbf{z}_2$ into the first lines of (40a) and (40b), one

can obtain

$$\begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} - \mathbf{EGF}^\mathrm{T}\mathbf{C} & -\mathbf{EGE}^\mathrm{T} - \mathbf{I} \\ q^2\mathbf{I} - \mathbf{C}^\mathrm{T}\mathbf{C} + \mathbf{C}^\mathrm{T}\mathbf{FGF}^\mathrm{T}\mathbf{C} & -\mathbf{A}^\mathrm{T} + \bar{\lambda}\mathbf{I} + \mathbf{C}^\mathrm{T}\mathbf{FGE}^\mathrm{T} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix} = \mathbf{0}. \tag{42}$$

The variable λ can be split up into its real and imaginary part $\lambda = \alpha + j\beta$. In (42), the imaginary part can be separated to get

$$\underbrace{\begin{bmatrix} -\alpha\mathbf{I} + \mathbf{A} - \mathbf{EGF}^\mathrm{T}\mathbf{C} & -\mathbf{EGE}^\mathrm{T} - \mathbf{I} \\ q^2\mathbf{I} - \mathbf{C}^\mathrm{T}\mathbf{C} + \mathbf{C}^\mathrm{T}\mathbf{FGF}^\mathrm{T}\mathbf{C} & \alpha\mathbf{I} - \mathbf{A}^\mathrm{T} + \mathbf{C}^\mathrm{T}\mathbf{FGE}^\mathrm{T} \end{bmatrix}}_{\mathbf{M}(\alpha)} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix}$$

$$= j\beta \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix}. \tag{43}$$

One can see that $j\beta$ is a purely imaginary eigenvalue of $\mathbf{M}(\alpha)$.

For the two shifted curves, $\lambda = \lambda_{\mathrm{int}} + \delta$ and $\lambda = \lambda_{\mathrm{int}} - \delta$ can be inserted into (43) to get

$$\mathbf{M}(\alpha_{\mathrm{int}} + \delta) \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix} = j\beta_{\mathrm{int}} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix} \tag{44a}$$

$$\mathbf{M}(\alpha_{\mathrm{int}} - \delta) \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix} = j\beta_{\mathrm{int}} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{v} \end{bmatrix}. \tag{44b}$$

Both $\mathbf{M}(\alpha_{\mathrm{int}} + \delta)$ and $\mathbf{M}(\alpha_{\mathrm{int}} - \delta)$ are Hamiltonian and $\mathbf{M}(\alpha_{\mathrm{int}} - \delta)$ and $-\mathbf{M}(\alpha_{\mathrm{int}} - \delta)^\mathrm{T}$ share the same eigenvalues. Exploiting the structure of $\mathbf{M}(\alpha)$, (44) can be separated into

$$\mathbf{M}(\alpha_{\mathrm{int}} + \delta) = \mathbf{M}(\delta) + \alpha_{\mathrm{int}}\mathbf{N} \tag{45a}$$

$$-\mathbf{M}(\alpha_{\mathrm{int}} - \delta)^\mathrm{T} = -\mathbf{M}(-\delta)^\mathrm{T} - \alpha_{\mathrm{int}}\mathbf{N} \tag{45b}$$

with $\mathbf{N} = \begin{bmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$.

Since $\mathbf{M}(\alpha_{\mathrm{int}} + \delta)$ and $-\mathbf{M}(\alpha_{\mathrm{int}} - \delta)^\mathrm{T}$ share a common purely imaginary eigenvalue at $j\beta_{\mathrm{int}}$, the Sylvester equation

$$(\mathbf{M}(\delta) + \alpha_{\mathrm{int}}\mathbf{N}) \mathbf{X} + \mathbf{X} (\mathbf{M}(-\delta)^\mathrm{T} + \alpha_{\mathrm{int}}\mathbf{N}) = \mathbf{0} \tag{46}$$

$$(\mathbf{M}(\delta)) \mathbf{X} + \mathbf{X} (\mathbf{M}(-\delta)^\mathrm{T}) = \alpha_{\mathrm{int}} (-\mathbf{NX} - \mathbf{XN}) \tag{47}$$

will be used. The two matrices share a common eigenvalue iff the Sylvester equation has a nonzero solution $\mathbf{X}$.

Using the Kroneckers product property $\mathrm{vec}(\mathbf{TXU}) = (\mathbf{U}^\mathrm{T} \otimes \mathbf{T})\,\mathrm{vec}(\mathbf{X})$, where $\mathrm{vec}(\mathbf{X})$ denotes the vector of stacked columns of $\mathbf{X}$, and $\mathbf{N}^\mathrm{T} = \mathbf{N}$, (47) can be rearranged into

$$(\mathbf{I} \otimes \mathbf{M}(\delta) + \mathbf{M}(-\delta) \otimes \mathbf{I}) \,\mathrm{vec}(\mathbf{X})$$
$$= \alpha_{\mathrm{int}} (-\mathbf{I} \otimes \mathbf{N} - \mathbf{N} \otimes \mathbf{I}) \,\mathrm{vec}(\mathbf{X}) \tag{48}$$

with

$$\mathbf{\Pi}(\delta) = \mathbf{I} \otimes \mathbf{M}(\delta) + \mathbf{M}(-\delta) \otimes \mathbf{I} \tag{49}$$

$$\mathbf{\Gamma} = -\mathbf{I} \otimes \mathbf{N} - \mathbf{N} \otimes \mathbf{I}. \tag{50}$$

This gives a generalized eigenvalue problem

$$(\mathbf{\Pi}(\delta) - \alpha_{\mathrm{int}}\mathbf{\Gamma}) \,\mathrm{vec}(\mathbf{X}) = \mathbf{0} \tag{51}$$

of size $(2n)^2 \times (2n)^2$. As a result, the problem of finding a nonzero solution $\mathbf{X}$ for (47) can be reduced to checking if the generalized eigenvalue problem (51) has a real solution α_{int}.

Conversely, suppose that α_{int} is a purely real solution of the generalized eigenvalue problem (51) and $j\beta_{\mathrm{int}}$ is a purely imaginary eigenvalue of $\mathbf{M}(\alpha_{\mathrm{int}} + \delta)$ and $\mathbf{M}(\alpha_{\mathrm{int}} - \delta)$. Then, the existence of α_{int} ensures that the matrices in (44) share a common eigenvalue, and there exists a vector $\mathbf{z} \neq \mathbf{0}$ such that

$$(\mathbf{M}(\alpha_{\mathrm{int}} \pm \delta) - j\beta_{\mathrm{int}}\mathbf{I})\mathbf{z} = 0 \tag{52}$$

holds. Equation (52) is equivalent to (43), and can be reformulated back into (39). It is clear that q is a singular value of $\mathbf{R}(\mathbf{S}, \lambda)$, and therefore condition (10) holds.

D. Reduction of Generalized Eigenvalue Problem

Starting from (47), the matrices $\mathbf{X}$, $\mathbf{M}(\delta)$, and $\mathbf{M}(-\delta)^\mathrm{T}$ can be partitioned into

$$\mathbf{M}(\delta) = \begin{bmatrix} -\delta\mathbf{I} + \mathbf{A} - \mathbf{EGF}^\mathrm{T}\mathbf{C} & -\mathbf{EGE}^\mathrm{T} - \mathbf{I} \\ q^2\mathbf{I} - \mathbf{C}^\mathrm{T}\mathbf{C} + \mathbf{C}^\mathrm{T}\mathbf{FGF}^\mathrm{T}\mathbf{C} & \delta\mathbf{I} - \mathbf{A}^\mathrm{T} + \mathbf{C}^\mathrm{T}\mathbf{FGE}^\mathrm{T} \end{bmatrix}$$

$$= \mathbf{M}_\mathrm{a} = \begin{bmatrix} \mathbf{M}_{\mathrm{a}11} & \mathbf{M}_{\mathrm{a}12} \\ \mathbf{M}_{\mathrm{a}21} & \mathbf{M}_{\mathrm{a}22} \end{bmatrix} \tag{53}$$

$$\mathbf{M}(-\delta)^\mathrm{T} = \begin{bmatrix} \delta\mathbf{I} + \mathbf{A} - \mathbf{EGF}^\mathrm{T}\mathbf{C} & q^2\mathbf{I} - \mathbf{C}^\mathrm{T}\mathbf{C} + \mathbf{C}^\mathrm{T}\mathbf{FGF}^\mathrm{T}\mathbf{C} \\ -\mathbf{EGE}^\mathrm{T} - \mathbf{I} & -\delta\mathbf{I} - \mathbf{A}^\mathrm{T} + \mathbf{C}^\mathrm{T}\mathbf{FGE}^\mathrm{T} \end{bmatrix}$$

$$= \mathbf{M}_\mathrm{b} = \begin{bmatrix} \mathbf{M}_{\mathrm{b}11} & \mathbf{M}_{\mathrm{b}12} \\ \mathbf{M}_{\mathrm{b}21} & \mathbf{M}_{\mathrm{b}22} \end{bmatrix} \tag{54}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \tag{55}$$

and (47) can be reformulated as

$$\begin{bmatrix} \mathbf{M}_{\mathrm{a}11} & \mathbf{M}_{\mathrm{a}12} \\ \mathbf{M}_{\mathrm{a}21} & \mathbf{M}_{\mathrm{a}22} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} + \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{M}_{\mathrm{b}11} & \mathbf{M}_{\mathrm{b}12} \\ \mathbf{M}_{\mathrm{b}21} & \mathbf{M}_{\mathrm{b}22} \end{bmatrix}$$

$$= \alpha_{\mathrm{int}} \left\{ \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} + \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \right\}. \tag{56}$$

Using the property of the Kronecker product, (56) can be vectorized into

$$\begin{bmatrix} \mathbf{K}_{\mathrm{a}11}U\mathbf{K}_{\mathrm{b}11} & \mathbf{K}_{\mathrm{b}21} & \mathbf{K}_{\mathrm{a}12} & \mathbf{0} \\ \mathbf{K}_{\mathrm{b}12} & \mathbf{K}_{\mathrm{a}11} + \mathbf{K}_{\mathrm{b}22} & \mathbf{0} & \mathbf{K}_{\mathrm{a}12} \\ \mathbf{K}_{\mathrm{a}21} & \mathbf{0} & \mathbf{K}_{\mathrm{a}22} + \mathbf{K}_{\mathrm{b}11} & \mathbf{K}_{\mathrm{b}21} \\ \mathbf{0} & -\mathbf{K}_{\mathrm{a}21} & -\mathbf{K}_{\mathrm{b}12} & -\mathbf{K}_{\mathrm{a}22} - \mathbf{K}_{\mathrm{b}22} \end{bmatrix}$$

$$\begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathrm{vec}(\mathbf{X}_{12}) \\ \mathrm{vec}(\mathbf{X}_{21}) \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix} = 2\alpha_{\mathrm{int}} \begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathbf{0} \\ \mathbf{0} \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix} \tag{57}$$

where $\mathbf{K}_{\mathrm{a}xy} = (\mathbf{I} \otimes \mathbf{M}_{\mathrm{a}xy})$ and $\mathbf{K}_{\mathrm{b}xy} = (\mathbf{M}_{\mathrm{b}xy}^\mathrm{T} \otimes \mathbf{I})$. From second and third lines of (57), $\mathrm{vec}(\mathbf{X}_{12})$ and $\mathrm{vec}(\mathbf{X}_{21})$ can be calculated and

inserted into first and fourth lines, to get

$$
\left\{
\begin{bmatrix} \mathbf{K}_{a11}+\mathbf{K}_{b11} & \mathbf{0} \\ \mathbf{0} & -\mathbf{K}_{a22}-\mathbf{K}_{b22} \end{bmatrix} - \begin{bmatrix} \mathbf{K}_{b21} & \mathbf{K}_{a12} \\ -\mathbf{K}_{a21} & -\mathbf{K}_{b12} \end{bmatrix}
\right.
$$

$$
\left.
\begin{bmatrix} \mathbf{K}_{a11}+\mathbf{K}_{b22} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{a22}+\mathbf{K}_{b11} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{K}_{b12} & \mathbf{K}_{a12} \\ \mathbf{K}_{a21} & \mathbf{K}_{b21} \end{bmatrix}
\right\}
$$

$$
\begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix} = 2\alpha_{\mathrm{int}} \begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix}. \tag{58}
$$

Therefore, α_{int} can be found by solving the eigenvalue problem

$$
\mathbf{\Pi}_{\mathrm{red}} \begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix} = \alpha_{\mathrm{int}} \begin{bmatrix} \mathrm{vec}(\mathbf{X}_{11}) \\ \mathrm{vec}(\mathbf{X}_{22}) \end{bmatrix} \tag{59}
$$

with

$$
\mathbf{\Pi}_{\mathrm{red}} = \left\{
\begin{bmatrix} \mathbf{K}_{a11}+\mathbf{K}_{b11} & \mathbf{0} \\ \mathbf{0} & -\mathbf{K}_{a22}-\mathbf{K}_{b22} \end{bmatrix} - \begin{bmatrix} \mathbf{K}_{b21} & \mathbf{K}_{a12} \\ -\mathbf{K}_{a21} & -\mathbf{K}_{b12} \end{bmatrix}
\right.
$$

$$
\left.
\begin{bmatrix} \mathbf{K}_{a11}+\mathbf{K}_{b22} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{a22}+\mathbf{K}_{b11} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{K}_{b12} & \mathbf{K}_{a12} \\ \mathbf{K}_{a21} & \mathbf{K}_{b21} \end{bmatrix}
\right\} \frac{1}{2}. \tag{60}
$$

As a result, instead of the $(2n)^2 \times (2n)^2$ generalized eigenvalue problem in (51), a $2n \times 2n$ eigenvalue problem given by (59) may be solved to find α_{int}.

ACKNOWLEDGMENT

REFERENCES

[1] M. Hautus, "Strong detectability and observers," *Linear Algebra Appl.*, vol. 50, pp. 353–368, Apr. 1983.

[2] W. Gawronski and K. Lim, "Balanced actuator and sensor placement for flexible structures," in *Proc. Guid., Navig., Control Conf.*, 1995, pp. 1871–1880.

[3] S. D. Bopardikar, "A randomized approach to sensor placement with observability assurance," *Automatica*, vol. 123, Jan. 2021, Art. no. 109340.

[4] J. Milosevic, A. Teixeira, K. H. Johansson, and H. Sandberg, "Actuator security indices based on perfect undetectability: Computation, robustness, and sensor placement," *IEEE Trans. Autom. Control*, vol. 65, no. 9, pp. 3816–3831, Sep. 2020.

[5] S. Moschik, "Steuerbarkeitsmaße für lineare zeitinvariante systeme," Ph.D. dissertation, Inst. Automat. Control, Graz Univ. Technol., Graz, Austria, 2011.

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

[7] P. Müller and H. Weber, "Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems," *Automatica*, vol. 8, no. 3, pp. 237–246, May 1972.

[8] R. Eising, "Between controllable and uncontrollable," *Syst. Control Lett.*, vol. 4, no. 5, pp. 263–264, Jul. 1984.

[9] 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, no. 2, pp. 565–570, Apr. 2021.

[10] M. Wicks and R. DeCarlo, "Computing the distance to an uncontrollable system," in *Proc. 28th IEEE Conf. Decis. Control*, 1989, pp. 2361–2366.

[11] M. Gao and M. Neumann, "A global minimum search algorithm for estimating the distance to uncontrollability," *Linear Algebra Appl.*, vol. 188-189, pp. 305–350, Jul. 1993.

[12] M. Gu, "New methods for estimating the distance to uncontrollability," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 3, pp. 989–1003, Jan. 2000.

[13] J. V. Burke, A. S. Lewis, and M. L. Overton, "Pseudospectral components and the distance to uncontrollability," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 350–361, Jan. 2004.

[14] R. Byers, "Detecting nearly uncontrollable pairs," in *Proc. Numer. Methods Proc. Int. Symp.*, 1990, pp. 447–457.

[15] M. Danilova *et al.*, "Recent theoretical advances in non-convex optimization," 2020, *arXiv:2012.06188.*

[16] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Boston, MA, USA: Addison Wesley, 1989.

[17] S. Lam and E. J. Davison, "The transmission zero at s radius and the minimum phase radius of LTI systems," *IFAC Proc. Volumes*, vol. 41, no. 2, pp. 6371–6376, 2008.

[18] S. Lam and E. J. Davison, "A fast algorithm to compute the controllability, decentralized fixed-mode, and minimum-phase radius of LTI systems," in *Proc. 47th IEEE Conf. Decis. Control*, 2008, pp. 508–513.

[19] 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, Jan. 2006.

[20] E. J. Davison and S. H. Wang, "Properties and calculation of transmission zeros of linear multivariable systems," *Automatica*, vol. 10, no. 6, pp. 643–658, 1974.

[21] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2013.

[22] K. Maes, E. Lourens, K. Van Nimmen, E. Reynders, G. De Roeck, and G. Lombaert, "Design of sensor networks for instantaneous inversion of modally reduced order models in structural dynamics," *Mech. Syst. Signal Process.*, vol. 52-53, no. 1, pp. 628–644, 2015.