

A SUBSPACE FRAMEWORK FOR \mathcal{H}_∞ -NORM MINIMIZATION

NICAT ALIYEV*, PETER BENNER†, EMRE MENGI‡, AND MATTHIAS VOIGT§

Abstract. We deal with the minimization of the \mathcal{H}_∞ -norm of the transfer function of a parameter-dependent descriptor system over the set of admissible parameter values. Subspace frameworks are proposed for such minimization problems where the involved systems are of large order. The proposed algorithms are greedy interpolatory approaches inspired by our recent work [Aliyev et al., SIAM J. Matrix Anal. Appl., 38(4):1496–1516, 2017] for the computation of the \mathcal{H}_∞ -norm. In this work, we minimize the \mathcal{H}_∞ -norm of a reduced-order parameter-dependent system obtained by two-sided restrictions onto certain subspaces. Then we expand the subspaces so that Hermite interpolation properties hold between the full and reduced-order system at the optimal parameter value for the reduced order system. We formally establish the superlinear convergence of the subspace frameworks under some smoothness assumptions. The fast convergence of the proposed frameworks in practice is illustrated by several large-scale systems.

Key words. \mathcal{H}_∞ -norm, large-scale, singular values, Hermite interpolation, descriptor systems, model order reduction, greedy search, reduced basis.

AMS subject classifications. 34K17, 65D05, 65F15, 65L80, 90C06, 90C26, 90C31, 93C05, 93D09

1. Introduction. In this work we are concerned with the minimization of the \mathcal{H}_∞ -norm of a parameter-dependent descriptor system of the form

$$(1.1) \quad \begin{aligned} \frac{d}{dt}E(\mu)x(\mu; t) &= A(\mu)x(\mu; t) + B(\mu)u(\mu; t), \\ y(\mu; t) &= C(\mu)x(\mu; t). \end{aligned}$$

Here, for an open and bounded set $\Omega \subseteq \mathbb{R}^d$, $E, A : \Omega \rightarrow \mathbb{R}^{n \times n}$, $B : \Omega \rightarrow \mathbb{R}^{n \times m}$, $C : \Omega \rightarrow \mathbb{R}^{p \times n}$ are matrix-valued functions defined by

$$(1.2) \quad \begin{aligned} E(\mu) &:= f_1(\mu)E_1 + \dots + f_{\kappa_E}(\mu)E_{\kappa_E}, \\ A(\mu) &:= g_1(\mu)A_1 + \dots + g_{\kappa_A}(\mu)A_{\kappa_A}, \\ B(\mu) &:= h_1(\mu)B_1 + \dots + h_{\kappa_B}(\mu)B_{\kappa_B}, \\ C(\mu) &:= k_1(\mu)C_1 + \dots + k_{\kappa_C}(\mu)C_{\kappa_C}, \end{aligned}$$

for given matrices $E_1, \dots, E_{\kappa_E}, A_1, \dots, A_{\kappa_A} \in \mathbb{R}^{n \times n}$, $B_1, \dots, B_{\kappa_B} \in \mathbb{R}^{n \times m}$, $C_1, \dots, C_{\kappa_C} \in \mathbb{R}^{p \times n}$, and real-analytic functions $f_1, \dots, f_{\kappa_E}, g_1, \dots, g_{\kappa_A}, h_1, \dots, h_{\kappa_B}, k_1, \dots, k_{\kappa_C} : \Omega \rightarrow \mathbb{R}$. The functions $x(\mu; \cdot) : \mathbb{R} \rightarrow \mathbb{R}^n$, $u(\mu; \cdot) : \mathbb{R} \rightarrow \mathbb{R}^m$, and $y(\mu; \cdot) : \mathbb{R} \rightarrow \mathbb{R}^p$ are called (generalized) state, input, and output, respectively. If for a *fixed* $\mu \in \Omega$, the matrix pencil $sE(\mu) - A(\mu)$ is *regular* (that is, there exists a $\lambda \in \mathbb{C}$ with $\det(\lambda E(\mu) - A(\mu)) \neq 0$), we define the transfer function of (1.1) by

$$H[\mu](s) := C(\mu)D(\mu, s)^{-1}B(\mu) \quad \text{with} \quad D(\mu, s) := sE(\mu) - A(\mu).$$

For a *fixed* μ , the function $H[\mu](s)$ is real-rational in the indeterminate s , consequently, we use the notation $H[\mu](s) \in \mathbb{R}(s)^{p \times m}$. Observe that, since $H[\mu]$ is rational, it is analytic almost everywhere in \mathbb{C} .

*E-Mail: naliyev@ku.edu.tr.

†Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany, E-Mail: benner@mpi-magdeburg.mpg.de.

‡Koç University, Department of Mathematics, Rumeli Feneri Yolu 34450, Sarıyer, Istanbul, Turkey, E-Mail: emengi@ku.edu.tr.

§Corresponding author, Universität Hamburg, Fachbereich Mathematik, Bereich Optimierung und Approximation, Bundesstraße 55, 20146 Hamburg, Germany, E-Mail: matthias.voigt@uni-hamburg.de and Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, 10623 Berlin, Germany, E-Mail: mvoigt@math.tu-berlin.de.

We define the following normed spaces of real-rational functions:

$$\begin{aligned}\mathcal{L}_\infty^{\mathbf{p} \times \mathbf{m}} &:= \left\{ H(s) \in \mathbb{R}(s)^{\mathbf{p} \times \mathbf{m}} \mid \sup_{\omega \in \mathbb{R}} \|H(i\omega)\|_2 < \infty \right\}, \\ \mathcal{H}_\infty^{\mathbf{p} \times \mathbf{m}} &:= \left\{ H(s) \in \mathbb{R}(s)^{\mathbf{p} \times \mathbf{m}} \mid \sup_{\lambda \in \mathbb{C}^+} \|H(\lambda)\|_2 < \infty \right\},\end{aligned}$$

where $\mathbb{C}^+ := \{\lambda \in \mathbb{C} \mid \text{Re}(\lambda) > 0\}$. For $H \in \mathcal{L}_\infty^{\mathbf{p} \times \mathbf{m}}$, the \mathcal{L}_∞ -norm is defined by

$$\|H\|_{\mathcal{L}_\infty} := \sup_{\omega \in \mathbb{R}} \|H(i\omega)\|_2 = \sup_{\omega \in \mathbb{R}} \sigma(H(i\omega)),$$

where $\sigma(\cdot)$ denotes the largest singular value of its matrix argument. We assume throughout this text that the functions under consideration are in $\mathcal{H}_\infty^{\mathbf{p} \times \mathbf{m}}$. For such a function $H \in \mathcal{H}_\infty^{\mathbf{p} \times \mathbf{m}}$, by employing the maximum principle for analytic functions, one can show that the \mathcal{H}_∞ -norm is equivalent to the \mathcal{L}_∞ -norm, that is

$$\|H\|_{\mathcal{H}_\infty} := \sup_{s \in \mathbb{C}^+} \|H(s)\|_2 = \sup_{s \in \partial\mathbb{C}^+} \|H(s)\|_2 = \sup_{\omega \in \mathbb{R}} \sigma(H(i\omega)).$$

In this work, we consider the problem of minimizing the \mathcal{H}_∞ -norm of $H[\mu]$ over μ that belongs to a compact subset $\underline{\Omega}$ of Ω , but keeping the assumption that $H[\mu] \in \mathcal{H}_\infty^{\mathbf{p} \times \mathbf{m}}$ for every $\mu \in \underline{\Omega}$. The latter assumption holds for all of the examples that we consider later in this paper; most of these examples arise from real applications. Formally, we aim to determine $\mu_* \in \underline{\Omega}$ such that

$$\|H[\mu_*]\|_{\mathcal{H}_\infty} = \min_{\mu \in \underline{\Omega}} \|H[\mu]\|_{\mathcal{H}_\infty}.$$

Minimizing the \mathcal{H}_∞ -norm of a parameter-dependent system is an important task in control engineering. For example, the parameter vector μ may consist of the design variables of a feedback controller. Then it is desirable to design an optimal \mathcal{H}_∞ -controller that minimizes the influence of a noisy input signal to the regulated output, which corresponds to minimizing the \mathcal{H}_∞ -norm of a closed-loop (parameter-dependent) transfer function, see, e. g., [19] and the references therein. Note that in the latter application, it is normally further imposed that the controller stabilizes the closed-loop system. This condition does not play a prominent role here, but efficient stability checks would be needed for controller design. Other applications for \mathcal{H}_∞ -norm minimization arise in the optimization of dynamic flow networks [9], parameter identification [18], and model reduction [17].

We focus on the large-scale setting, that is when n is large. We additionally impose the condition that the numbers of inputs and outputs are relatively small, i. e., $n \gg m, p$. Here we present subspace frameworks that are inspired by our previous work [1]. The proposed frameworks converge fast with respect to the subspace dimension. We provide a theoretical analysis which explains this convergence behavior and confirm our theoretical findings in practice by means of several numerical experiments.

Outline. The subspace frameworks are formally introduced in the next section. We first provide a basic greedy framework for \mathcal{H}_∞ -norm minimization in Algorithm 2.1. This framework reduces the order of the full-order system by employing two-sided restrictions to certain subspaces. It performs the \mathcal{H}_∞ -norm minimization on the reduced system, then expands the restriction subspaces so that Hermite interpolation properties hold between the full and reduced-order system at the optimal parameter value for the reduced system. An extension of the basic framework is proposed in Algorithm 2.2. There Hermite interpolation properties do not only hold at the optimal parameter value for the reduced system, but also at nearby points. In Section 3, we formally show that the basic subspace framework when there is only one parameter, and the extended framework converge with a superlinear rate

Algorithm 2.1 The basic greedy algorithm for \mathcal{H}_∞ -norm minimization

Input: Matrices $E_1, \dots, E_{\kappa_E} \in \mathbb{R}^{n \times n}$, $A_1, \dots, A_{\kappa_A} \in \mathbb{R}^{n \times n}$, $B_1, \dots, B_{\kappa_B} \in \mathbb{R}^{n \times m}$, $C_1, \dots, C_{\kappa_C} \in \mathbb{R}^{p \times n}$ and functions f_1, \dots, f_{κ_E} , g_1, \dots, g_{κ_A} , h_1, \dots, h_{κ_B} , k_1, \dots, k_{κ_C} as in (1.2).

Output: Sequences $\{\mu^{(k)}\}$, $\{\omega^{(k)}\}$.

- 1: Choose initial subspace $\mathcal{V}_0, \mathcal{W}_0 \subseteq \mathbb{C}^n$.
 - 2: **for** $k = 1, 2, \dots$ **do**
 - 3: $\mu^{(k)} \leftarrow \arg \min_{\mu \in \Omega} \|H^{\mathcal{V}_{k-1}, \mathcal{W}_{k-1}}[\mu]\|_{\mathcal{H}_\infty}$.
 - 4: $\omega^{(k)} \leftarrow \arg \max_{\omega \in \mathbb{R} \cup \{\infty\}} \sigma(\mu^{(k)}, \omega)$.
 - 5: **if** $m = p$ **then**
 - 6: $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)})$.
 - 7: $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^*$.
 - 8: **else if** $m < p$ **then**
 - 9: $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)})$.
 - 10: $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^* H[\mu^{(k)}](i\omega^{(k)})$.
 - 11: **else**
 - 12: $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)}) H[\mu^{(k)}](i\omega^{(k)})^*$
 - 13: $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^*$.
 - 14: **end if**
 - 15: $\mathcal{V}_k \leftarrow \mathcal{V}_{k-1} \oplus \text{Col}(\tilde{V}_k)$ and $\mathcal{W}_k \leftarrow \mathcal{W}_{k-1} \oplus \text{Col}(\tilde{W}_k)$.
 - 16: **end for**
-

under some smoothness assumptions at the minimizer. The performance of proposed basic subspace framework and its rate of convergence are illustrated for several examples in Section 4. As we report in the end, with the proposed subspace frameworks, only a few seconds are required for the minimization of the \mathcal{H}_∞ -norm of a parameter-dependent system of order 10^4 , in contrast to an approach that does not make use of reductions.

2. Subspace Frameworks. To deal with the large-scale problems described in the introduction, we employ two-sided restrictions in the flavor of the practice we followed for large-scale \mathcal{H}_∞ -norm computation in [1]. We choose two subspaces $\mathcal{V}, \mathcal{W} \subseteq \mathbb{C}^n$ of the same dimension, as well as matrices $V, W \in \mathbb{C}^{n \times k}$ whose columns form orthonormal bases for these subspaces, and define the reduced problem by

$$\begin{aligned} E^{V,W}(\mu) &:= f_1(\mu)W^*E_1V + \dots + f_{\kappa_E}(\mu)W^*E_{\kappa_E}V, \\ A^{V,W}(\mu) &:= g_1(\mu)W^*A_1V + \dots + g_{\kappa_A}(\mu)W^*A_{\kappa_A}V, \\ B^W(\mu) &:= h_1(\mu)W^*B_1 + \dots + h_{\kappa_B}(\mu)W^*B_{\kappa_B}, \\ C^V(\mu) &:= k_1(\mu)C_1V + \dots + k_{\kappa_C}(\mu)C_{\kappa_C}V. \end{aligned}$$

Associated with this system, there is the reduced transfer function

$$H^{\mathcal{V}, \mathcal{W}}[\mu](s) := C^V(\mu)D^{V,W}(\mu, s)^{-1}B^W(\mu) \quad \text{with} \quad D^{V,W}(\mu, s) := sE^{V,W}(\mu) - A^{V,W}(\mu)$$

which turns out to be independent of the particular choice of the bases for \mathcal{V} and \mathcal{W} . Our subspace frameworks are based on the repeated minimization of $\|H^{\mathcal{V}, \mathcal{W}}(\mu)\|_{\mathcal{H}_\infty}$ for appropriate choices of the subspaces \mathcal{V}, \mathcal{W} .

The basic greedy framework is given in Algorithm 2.1 and throughout the rest of this work, we use the short-hand notations

$$\sigma(\mu, \omega) := \sigma(H[\mu](i\omega)) \quad \text{and} \quad \sigma^{\mathcal{V}, \mathcal{W}}(\mu, \omega) := \sigma(H^{\mathcal{V}, \mathcal{W}}[\mu](i\omega)).$$

We will also make frequent use of certain partial derivatives of these functions, where we denote the variables that we differentiate by subscripts, e. g., $\sigma_\omega(\cdot, \cdot)$ denotes the first partial

derivative with respect to the argument ω , whereas $\sigma_\mu(\cdot, \cdot)$ denotes the gradient with respect to μ . Additionally, we reserve the notations $\sigma_2(\mu, \omega)$ and $\sigma_2^{\mathcal{V}, \mathcal{W}}(\mu, \omega)$ for the second largest singular values of $H[\mu](i\omega)$ and $H^{\mathcal{V}, \mathcal{W}}[\mu](i\omega)$, respectively. At every iteration, the basic framework minimizes the \mathcal{H}_∞ -norm of a reduced problem for a given pair of subspaces in line 3. Then it first computes an ω such that $\|H[\mu]\|_{\mathcal{H}_\infty} = \sigma(\omega, \mu)$ in line 4 at the optimal μ value for the reduced problem, and expands the subspaces so that the following Hermite interpolation properties hold at the optimal μ, ω , which are immediate from [1, Theorem 2.1], [3, Theorem 1].

LEMMA 2.1. *The following assertions hold regarding Algorithm 2.1 for each $j = 1, \dots, k$:*

- (i) *It holds that $\|H[\mu^{(j)}]\|_{\mathcal{H}_\infty} = \sigma(\mu^{(j)}, \omega^{(j)}) = \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j)}, \omega^{(j)})$.*
- (ii) *It holds that $\sigma_2(\mu^{(j)}, \omega^{(j)}) = \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j)}, \omega^{(j)})$.*
- (iii) *If the largest singular value $\sigma(\mu^{(j)}, \omega^{(j)})$ of $H[\mu^{(j)}](i\omega^{(j)})$ is simple, then*

$$\nabla \|H[\mu^{(j)}]\|_{\mathcal{H}_\infty} = \sigma_\mu(\mu^{(j)}, \omega^{(j)}) = \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j)}, \omega^{(j)}).$$

- (iv) *We have $\sigma_\omega(\mu^{(j)}, \omega^{(j)}) = \sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j)}, \omega^{(j)}) = 0$.*

Note that in part (iv) of the lemma above $\sigma_\omega(\mu^{(j)}, \omega^{(j)}) = 0$ holds even if $\sigma(\mu^{(j)}, \omega^{(j)})$ is not simple, since $\omega^{(j)}$ is a maximizer of $\sigma(\mu^{(j)}, \cdot)$. The equality $\sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j)}, \omega^{(j)}) = 0$ follows from the interpolation properties between $H[\mu](i\omega)$, $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$ and their first derivatives at $\mu = \mu^{(j)}$, $\omega = \omega^{(j)}$.

We also propose an extended version of the basic greedy framework in Algorithm 2.2. For its description we define $e_{rq} := 1/\sqrt{2}(e_r + e_q)$ if $r \neq q$ and $e_{rr} := e_r$, where e_r is the r -th column of the $\mathbf{d} \times \mathbf{d}$ identity matrix. The description may look complicated at first, but the only main difference is that it includes additional vectors in the subspaces in lines 16–35 to interpolate not only at the minimizers of the reduced problems, but also at nearby points. The motivation for the inclusion of these additional vectors is to draw a theoretical conclusion about the accuracy of the second derivatives of the reduced singular value functions $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ in approximating $\sigma(\cdot, \cdot)$ in the multivariate case. In practice, we observe that both Algorithm 2.1 and Algorithm 2.2 converge rapidly. But in the multivariate case, the inclusion of the additional vectors in the subspaces in Algorithm 2.2 makes its rate of convergence analysis neater. The interpolation properties of the extended framework are listed in the next result. Once again, these properties are immediate from [1, Theorem 2.1].

LEMMA 2.2. *The iterates $\{\mu^{(k)}\}$, $\{\omega^{(k)}\}$ by Algorithm 2.2 satisfy the assertions (i)–(iii) of Lemma 2.1 for each $j = 1, \dots, k$. Additionally, for each $j = 1, \dots, k$, $r = 1, \dots, \mathbf{d}$, and $q = r, \dots, \mathbf{d}$, we have the following:*

- (i) *It holds that $\|H[\mu^{(j, rq)}]\|_{\mathcal{H}_\infty} = \sigma(\mu^{(j, rq)}, \omega^{(j, rq)}) = \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j, rq)}, \omega^{(j, rq)})$.*
- (ii) *If the largest singular value $\sigma(\mu^{(j, rq)}, \omega^{(j, rq)})$ of $H[\mu^{(j, rq)}](i\omega^{(j, rq)})$ is simple, then*

$$\nabla \|H[\mu^{(j, rq)}]\|_{\mathcal{H}_\infty} = \sigma_\mu(\mu^{(j, rq)}, \omega^{(j, rq)}) = \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j, rq)}, \omega^{(j, rq)}).$$

- (iii) *It holds that $\sigma_\omega(\mu^{(j, rq)}, \omega^{(j, rq)}) = \sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(j, rq)}, \omega^{(j, rq)}) = 0$.*

Before we start with the rate of convergence analysis, a few comments regarding the two algorithms are in order.

REMARK 2.3. (i) *The distinctions of cases in lines 5–14 in Algorithm 2.1 and lines 5–14 and 22–31 are done such that the subspaces \mathcal{V}_k and \mathcal{W}_k have the same dimension. This is needed in order to obtain a regular reduced matrix pencil $D^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, s)$ and a well-defined reduced transfer function $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k)}](s)$. In practice, a regularization procedure can be performed [11] to obtain a regular reduced matrix pencil, even if the*

Algorithm 2.2 The extended greedy algorithm for \mathcal{H}_∞ -norm minimization

Input: Matrices $E_1, \dots, E_{\kappa_E} \in \mathbb{R}^{n \times n}$, $A_1, \dots, A_{\kappa_A} \in \mathbb{R}^{n \times n}$, $B_1, \dots, B_{\kappa_B} \in \mathbb{R}^{n \times m}$, $C_1, \dots, C_{\kappa_C} \in \mathbb{R}^{p \times n}$ and functions f_1, \dots, f_{κ_E} , g_1, \dots, g_{κ_A} , h_1, \dots, h_{κ_B} , k_1, \dots, k_{κ_C} as in (1.2).

Output: Sequences $\{\mu^{(k)}\}$, $\{\omega^{(k)}\}$.

```

1: Choose initial subspace  $\mathcal{V}_0, \mathcal{W}_0 \subseteq \mathbb{C}^n$ .
2: for  $k = 1, 2, \dots$  do
3:    $\mu^{(k)} \leftarrow \arg \min_{\mu \in \Omega} \|H^{\mathcal{V}_{k-1}, \mathcal{W}_{k-1}}[\mu]\|_{\mathcal{H}_\infty}$ .
4:    $\omega^{(k)} \leftarrow \arg \max_{\omega \in \mathbb{R} \cup \{\infty\}} \sigma(\mu^{(k)}, \omega)$ .
5:   if  $m = p$  then
6:      $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)})$ .
7:      $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^*$ .
8:   else if  $m < p$  then
9:      $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)})$ .
10:     $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^* H[\mu^{(k)}](i\omega^{(k)})$ .
11:   else
12:      $\tilde{V}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-1} B(\mu^{(k)}) H[\mu^{(k)}](i\omega^{(k)})^*$ .
13:      $\tilde{W}_k \leftarrow D(\mu^{(k)}, i\omega^{(k)})^{-*} C(\mu^{(k)})^*$ .
14:   end if
15:    $\mathcal{V}_k \leftarrow \mathcal{V}_{k-1} \oplus \text{Col}(\tilde{V}_k)$  and  $\mathcal{W}_k \leftarrow \mathcal{W}_{k-1} \oplus \text{Col}(\tilde{W}_k)$ .
16:   if  $k \geq 2$  then
17:      $h^{(k)} \leftarrow \|\mu^{(k)} - \mu^{(k-1)}\|_2$ .
18:     for  $r = 1, 2, \dots, d$  do
19:       for  $q = r, \dots, d$  do
20:          $\mu^{(k,rq)} \leftarrow \mu^{(k)} + h^{(k)} e_{rq}$ .
21:          $\omega^{(k,rq)} \leftarrow \arg \max_{\omega \in \mathbb{R} \cup \{\infty\}} \sigma(\mu^{(k,rq)}, \omega)$ .
22:         if  $m = p$  then
23:            $\tilde{V}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-1} B(\mu^{(k,rq)})$ .
24:            $\tilde{W}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-*} C(\mu^{(k,rq)})^*$ .
25:         else if  $m < p$  then
26:            $\tilde{V}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-1} B(\mu^{(k,rq)})$ .
27:            $\tilde{W}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-*} C(\mu^{(k,rq)})^* H[\mu^{(k,rq)}](i\omega^{(k,rq)})$ .
28:         else
29:            $\tilde{V}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-1} B(\mu^{(k,rq)}) H[\mu^{(k,rq)}](i\omega^{(k,rq)})^*$ .
30:            $\tilde{W}_k^{(rq)} \leftarrow D(\mu^{(k,rq)}, i\omega^{(k,rq)})^{-*} C(\mu^{(k,rq)})^*$ .
31:         end if
32:          $\mathcal{V}_k \leftarrow \mathcal{V}_k \oplus \text{Col}(\tilde{V}_k^{(rq)})$  and  $\mathcal{W}_k \leftarrow \mathcal{W}_k \oplus \text{Col}(\tilde{W}_k^{(rq)})$ .
33:       end for
34:     end for
35:   end if
36: end for

```

above distinctions of cases are not carried out. In the above algorithms we make the silent assumption that the transfer functions $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k)}](s)$ are well-defined and in $\mathcal{L}_\infty^{p \times m}$ for all k . Note that the reduced dynamical systems associated with the transfer functions $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k)}](s)$ are not necessarily asymptotically stable, so the transfer functions are not necessarily in $\mathcal{H}_\infty^{p \times m}$. However, for the algorithm, the latter does not lead to any problem.

(ii) In this paper, we only consider parameter-dependent linear time-invariant systems.

Efficient algorithms for the computation of the \mathcal{L}_∞ -norm however, have also been recently considered for transfer functions of a more general class of systems [1, 15]. The results presented here can be transferred to this more general situation without any changes in the algorithm description.

3. Rate of Convergence Analysis. In this section, we perform a rate of convergence analysis for Algorithms 2.1 and 2.2. We view $\mu^{(k+1)}, \mu^{(k)}, \mu^{(k-1)}$ as functions of $\mu^{(1)}$. Letting μ_* be a local or a global minimizer of $\|H[\cdot]\|_{\mathcal{H}_\infty}$, we assume $\mu^{(k+1)}, \mu^{(k)}, \mu^{(k-1)} \rightarrow \mu_*$ as $\mu^{(1)} \rightarrow \mu_*$. Our main result is a superlinear convergence result, i. e., for all $k \geq 2$ there exists a constant C , independent of $\mu^{(1)}$, such that

$$\|\mu^{(k+1)} - \mu_*\|_2 \leq C \left(\|\mu^{(k)} - \mu_*\|_2 \cdot \max \{ \|\mu^{(k)} - \mu_*\|_2, \|\mu^{(k-1)} - \mu_*\|_2 \} \right)$$

for all $\mu^{(1)}$ sufficiently close to μ_* .

The analysis here addresses the smooth setting, that is, throughout this section we assume the following:

ASSUMPTION 3.1 (Smoothness). *(i) The supremum of $\sigma(\mu_*, \cdot)$ is attained uniquely, say at ω_* , and (ii) $\sigma(\mu_*, \omega_*) > 0$ is a simple singular value of $H[\mu_*](i\omega_*)$.*

Many of the results in this section are established uniformly over every $\mu^{(1)}$ that is sufficiently close to μ_* . The dependence of the reduced transfer function $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)$, as well as the reduced singular value functions $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot), \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ on $\mu^{(1)}$ is implicitly given through the subspaces $\mathcal{V}_k, \mathcal{W}_k$ or equivalently, the matrices V_k, W_k whose columns form orthonormal bases for these subspaces. We start with uniform Lipschitz continuity results for these functions with respect to $\mu^{(1)}$. Note that in this result and in the subsequent discussions, $\sigma_{\min}(\cdot)$ denotes the smallest singular value of its matrix argument, whereas

$$\overline{\mathcal{B}}(\tilde{\mu}, \eta) := \{ \mu \in \mathbb{R}^d \mid \|\mu - \tilde{\mu}\|_2 \leq \eta \} \quad \text{and} \quad \overline{\mathcal{B}}(\tilde{\omega}, \eta) := \{ \omega \in \mathbb{R} \mid |\omega - \tilde{\omega}| \leq \eta \}$$

for given $\tilde{\mu} \in \mathbb{R}^d, \tilde{\omega} \in \mathbb{R}$, and $\eta > 0$.

LEMMA 3.2 (Uniform Lipschitz continuity). *Suppose that, for some $\beta > 0$, the point $\mu^{(1)}$ is such that $\sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, i\omega_*)) \geq \beta$. Then there exist constants $\eta_\mu, \eta_\omega, \gamma$ independent of $\mu^{(1)}$ such that*

- (i) $\|H^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](i\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma \|\tilde{\mu} - \mu\|_2$
 $\forall \tilde{\mu}, \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega);$
- (ii) $\|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\tilde{\omega}) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma |\tilde{\omega} - \omega|$
 $\forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \tilde{\omega}, \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega);$
- (iii) $|\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \omega) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq \gamma \|\tilde{\mu} - \mu\|_2,$
 $|\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \omega) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq \gamma \|\tilde{\mu} - \mu\|_2$
 $\forall \tilde{\mu}, \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega),$
- (iv) $|\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \tilde{\omega}) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq \gamma |\tilde{\omega} - \omega|,$
 $|\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \tilde{\omega}) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq \gamma |\tilde{\omega} - \omega|$
 $\forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \tilde{\omega}, \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega).$

Proof. By Weyl's theorem [8, Theorem 4.3.1], for every $\mu \in \underline{\Omega}$ and $\omega \in \mathbb{R}$ we have

$$\begin{aligned} |\sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)) - \sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, \omega_*))| &\leq \|D^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) - D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, \omega_*)\|_2 \\ &= \|W_k^*(D(\mu, \omega) - D(\mu_*, \omega_*))V_k\|_2 \\ &\leq \|D(\mu, \omega) - D(\mu_*, \omega_*)\|_2 \\ &\leq \nu(\|\mu - \mu_*\|_2 + |\omega - \omega_*|) \end{aligned}$$

for some $\nu > 0$, where the last inequality is due to the fact that $D(\cdot, \cdot)$ is continuously differentiable in a neighborhood of $\underline{\Omega} \times \mathbb{R}$. This uniform Lipschitz continuity property of $\sigma_{\min}(D^{V_k, W_k}(\cdot, \cdot))$, combined with $\sigma_{\min}(D^{V_k, W_k}(\mu_*, i\omega_*)) \geq \beta$, implies the existence of η_μ, η_ω independent of $\mu^{(1)}$ such that

$$\sigma_{\min}(D^{V_k, W_k}(\mu, \omega)) \geq \beta/2 \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega).$$

It follows that $(\mu, \omega) \mapsto H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$ is differentiable $\forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \forall \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$.

(i) For every $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$ and $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$, by the product and chain rule we obtain

$$(3.1) \quad \begin{aligned} \frac{\partial H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)}{\partial \mu_j} &= \frac{\partial C^{V_k}(\mu)}{\partial \mu_j} D^{V_k, W_k}(\mu, i\omega)^{-1} B^{W_k}(\mu) + \\ &C^{V_k}(\mu) D^{V_k, W_k}(\mu, i\omega)^{-1} \frac{\partial D^{V_k, W_k}(\mu, i\omega)}{\partial \mu_j} D^{V_k, W_k}(\mu, i\omega)^{-1} B^{W_k}(\mu) + \\ &C^{V_k}(\mu) D^{V_k, W_k}(\mu, i\omega)^{-1} \frac{\partial B^{W_k}(\mu)}{\partial \mu_j} \end{aligned}$$

for $j = 1, \dots, d$. Setting

$$\begin{aligned} M'_{D,j} &:= \max \left\{ \left\| \frac{\partial D(\mu, i\omega)}{\partial \mu_j} \right\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega) \right\} \\ M'_{C,j} &:= \max \left\{ \left\| \frac{\partial C(\mu)}{\partial \mu_j} \right\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\}, \quad M_C := \max \{ \|C(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \}, \\ M'_{B,j} &:= \max \left\{ \left\| \frac{\partial B(\mu)}{\partial \mu_j} \right\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\}, \quad M_B := \max \{ \|B(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \}, \end{aligned}$$

and exploiting

$$\begin{aligned} \left\| \frac{\partial D(\mu, i\omega)}{\partial \mu_j} \right\|_2 &\geq \left\| \frac{\partial D^{V_k, W_k}(\mu, i\omega)}{\partial \mu_j} \right\|_2, \quad \|C(\mu)\|_2 \geq \|C^{V_k}(\mu)\|_2, \\ \|B(\mu)\|_2 &\geq \|B^{W_k}(\mu)\|_2, \quad \left\| \frac{\partial C(\mu)}{\partial \mu_j} \right\|_2 \geq \left\| \frac{\partial C^{V_k}(\mu)}{\partial \mu_j} \right\|_2, \quad \left\| \frac{\partial B(\mu)}{\partial \mu_j} \right\|_2 \geq \left\| \frac{\partial B^{W_k}(\mu)}{\partial \mu_j} \right\|_2, \end{aligned}$$

as well as $\sigma_{\min}(D^{V_k, W_k}(\mu, \omega)) \geq \beta/2$, we deduce from (3.1) that

$$\left\| \frac{\partial H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)}{\partial \mu_j} \right\|_2 \leq 2 \frac{M'_{C,j} M_B}{\beta} + 4 \frac{M_C M'_{D,j} M_B}{\beta^2} + 2 \frac{M_C M'_{B,j}}{\beta} =: M_j$$

for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$, $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$ and $j = 1, \dots, d$, where M_j does not depend on $\mu^{(1)}$. This in particular implies $|\partial H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)/\partial \mu_j|_{k\ell} \leq M_j$ for $k = 1, \dots, p$, $\ell = 1, \dots, m$. With $M := \max\{M_j \mid j = 1, \dots, d\}$, for every $\tilde{\mu}, \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$, by the mean value theorem we obtain

$$\begin{aligned} | [H^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](i\omega)]_{k\ell} - [H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)]_{k\ell} | &\leq |\nabla_\mu [H^{\mathcal{V}_k, \mathcal{W}_k}[\hat{\mu}](i\omega)]_{k\ell}^\top (\tilde{\mu} - \mu)| \\ &\leq \sum_{j=1}^d M_j |\tilde{\mu}_j - \mu_j| \leq dM \|\tilde{\mu} - \mu\|_2 \end{aligned}$$

for some $\hat{\mu} \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$, where $\nabla_\mu [H^{\mathcal{V}_k, \mathcal{W}_k}[\hat{\mu}](i\omega)]_{k\ell}$ denotes the gradient of $\mu \mapsto [H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)]_{k\ell}$ at $\hat{\mu}$. It follows that

$$\|H^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](i\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \sqrt{pmd} M \|\tilde{\mu} - \mu\|_2,$$

where the Lipschitz constant $\sqrt{pmd}M$ is independent of $\mu^{(1)}$ as desired.

- (ii) A similar proof as in part (i) applies but now by differentiating the function $(\mu, \omega) \mapsto H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$ with respect to ω instead of μ_j .
- (iii) By Weyl's theorem [8, Theorem 4.3.1] and part (i) we have

$$\begin{aligned} |\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \omega) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| &\leq \|H^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](i\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma \|\tilde{\mu} - \mu\|_2, \\ |\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \omega) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| &\leq \|H^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](i\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma \|\tilde{\mu} - \mu\|_2 \end{aligned}$$

for all $\tilde{\mu}, \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$ and $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$, hence we get the result.

- (iv) Weyl's theorem and part (ii) combined imply

$$\begin{aligned} |\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \tilde{\omega}) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| &\leq \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\tilde{\omega}) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma |\tilde{\omega} - \omega|, \\ |\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \tilde{\omega}) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| &\leq \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\tilde{\omega}) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\|_2 \leq \gamma |\tilde{\omega} - \omega| \end{aligned}$$

for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$ and $\tilde{\omega}, \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$ as claimed. \square

The lemma below asserts uniform upper bounds on the derivatives of the largest singular value function for the reduced problem provided $D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, \omega_*)$ is away from singularity. Its proof is inspired by [10, Proposition 2.9], and given in the appendix.

LEMMA 3.3. *Suppose that Assumption 3.1 holds and that $\|\mu^{(1)} - \mu_*\|_2$ is small enough. Additionally, assume that for some $\beta > 0$, the point $\mu^{(1)}$ is such that*

$$\sigma_{\min}(D(\mu_*, i\omega_*)) \geq \beta \quad \text{and} \quad \sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, i\omega_*)) \geq \beta.$$

Then there exist a U and constants $\eta_\mu, \eta_\omega > 0$ independent of $\mu^{(1)}$ such that

$$\begin{aligned} |\sigma_{\chi_1}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq U, \quad |\sigma_{\chi_1 \chi_2}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq U, \quad |\sigma_{\chi_1 \chi_2 \chi_3}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq U \\ \forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \quad \forall \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega) \end{aligned}$$

for all $\chi_1, \chi_2, \chi_3 \in \{\omega\} \cup \{\mu_j \mid j = 1, \dots, d\}$.

The next result draws two important conclusions. First, the maximizers of $\sigma(\mu, \cdot)$ and $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ can be expressed as smooth functions of μ in a neighborhood of μ_* . Second, $\|H[\cdot]\|_{\mathcal{H}_\infty}$, as well as its reduced counter-parts generated by the algorithms are smooth locally around μ_* .

PROPOSITION 3.4. *Suppose that Assumption 3.1 holds and that $\|\mu^{(1)} - \mu_*\|_2$ is small enough. Furthermore, assume for some $\delta < 0$ and $\beta > 0$ that the point $\mu^{(1)}$ is such that*

$$(3.2) \quad \sigma_{\omega\omega}(\mu_*, \omega_*) \leq \delta \quad \text{and} \quad \sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, \omega_*) \leq \delta,$$

as well as

$$\sigma_{\min}(D(\mu_*, i\omega_*)) \geq \beta \quad \text{and} \quad \sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, i\omega_*)) \geq \beta.$$

Then for some $\eta_{\mu,0}, \eta_{\omega,0}, \varepsilon > 0$ independent of $\mu^{(1)}$, the following assertions hold:

- (i) *There exists a unique continuous function $\omega : \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \rightarrow \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ that is three times continuously differentiable in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ such that*

$$\omega(\mu_*) = \omega_* \quad \text{and} \quad \sigma_\omega(\mu, \omega(\mu)) = 0 \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}).$$

Furthermore, $\sigma_{\omega\omega}(\mu, \omega(\mu)) \leq \delta/2$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$.

- (ii) *There exists a unique continuous function $\omega^{\mathcal{V}_k, \mathcal{W}_k} : \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \rightarrow \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ that is three times continuously differentiable in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ such that*

$$\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) = \omega^{(k)} \quad \text{and} \quad \sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) = 0 \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}).$$

Additionally, $\sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \leq \delta/2$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$.

(iii) We have

$$\sigma(\mu, \omega(\mu)) - \sigma_2(\mu, \omega(\mu)) \geq \varepsilon,$$

and the unique global maximizer of $\sigma(\mu, \cdot)$ is given by $\omega(\mu)$. In particular, for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ it holds that

$$\sigma(\mu, \omega(\mu)) = \|H[\mu]\|_{\mathcal{H}_\infty}.$$

(iv) We have

$$\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \geq \varepsilon,$$

and the unique global maximizer and stationary point of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ in $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ is $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$.

Proof. As argued in the opening of the proof of Lemma 3.3, we have

$$(3.3) \quad \sigma(\mu, \omega) - \sigma_2(\mu, \omega) \geq \widehat{\varepsilon} \quad \forall (\mu, \omega) \in \overline{\mathcal{B}}(\mu_*, \widehat{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widehat{\eta}_\omega)$$

for some $\widehat{\varepsilon} > 0$, $\widehat{\eta}_\mu > 0$, $\widehat{\eta}_\omega > 0$, and

$$(3.4) \quad \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \geq \varepsilon \quad \forall (\mu, \omega) \in \overline{\mathcal{B}}(\mu_*, \widetilde{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widetilde{\eta}_\omega)$$

for some $\varepsilon \in (0, \widehat{\varepsilon})$, $\widetilde{\eta}_\mu > 0$, $\widetilde{\eta}_\omega > 0$. An important point here is that ε , $\widetilde{\eta}_\mu$, $\widetilde{\eta}_\omega$ do not depend on $\mu^{(1)}$. The function $\sigma(\cdot, \cdot)$ is real analytic in the interior of $\overline{\mathcal{B}}(\mu_*, \widehat{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widehat{\eta}_\omega)$, whereas $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ is real analytic in the interior of $\overline{\mathcal{B}}(\mu_*, \widetilde{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widetilde{\eta}_\omega)$. Moreover, by Lemma 3.3, there exists a $\widetilde{\delta} > 0$ such that

$$(3.5) \quad |\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)| \leq \widetilde{\delta}$$

holds uniformly for all (μ, ω) in a neighborhood of (μ_*, ω_*) , where $\widetilde{\delta}$ and the neighborhood are independent of $\mu^{(1)}$. Now we prove the four statements of the proposition:

(i) Since $\sigma(\cdot, \cdot)$ is real analytic with continuous second derivatives in a neighborhood of (μ_*, ω_*) , its second derivative $\sigma_{\omega\omega}(\cdot, \cdot)$ must be bounded from above by $\delta/2$ in another neighborhood of (μ_*, ω_*) . Then the assertion follows immediately from the implicit function theorem.

(ii) Due to (3.5) the condition $\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \leq \delta/2$ must hold in an open neighborhood \mathcal{N} of (μ_*, ω_*) independent of $\mu^{(1)}$. Additionally, observe that $\omega^{(k)} \rightarrow \omega_*$, $\mu^{(k)} \rightarrow \mu_*$ as $\mu^{(1)} \rightarrow \mu_*$ due to

$$\sigma(\mu_*, \omega_*) = \|H[\mu_*]\|_{\mathcal{H}_\infty} = \lim_{\mu^{(1)} \rightarrow \mu_*} \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} = \lim_{\mu^{(1)} \rightarrow \mu_*} \sigma(\mu^{(k)}, \omega^{(k)}),$$

as well as the uniqueness of ω_* as the maximizer of $\sigma(\mu_*, \cdot)$ and the continuity of $\sigma(\cdot, \cdot)$. We assume $\|\mu^{(1)} - \mu_*\|_2$ is small enough so that $(\mu^{(k)}, \omega^{(k)}) \in \mathcal{N}$, particularly $\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) \leq \delta/2 < 0$. Now the assertion again follows from the implicit function theorem. The uniformity of the radii $\eta_{\mu,0}$, $\eta_{\omega,0}$ over all $\mu^{(1)}$ (as long as $\|\mu^{(1)} - \mu_*\|_2$ is small enough) follows from the uniform upper bound $\delta/2$ on the second derivatives.

(iii) Assume $\eta_{\mu,0} \leq \widehat{\eta}_\mu$, $\eta_{\omega,0} \leq \widehat{\eta}_\omega$ without loss of generality, where $\widehat{\eta}_\mu$, $\widehat{\eta}_\omega$ are as in (3.3). But then for $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \subseteq \overline{\mathcal{B}}(\mu_*, \widehat{\eta}_\mu)$, we have $\omega(\mu) \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0}) \subseteq \overline{\mathcal{B}}(\omega_*, \widehat{\eta}_\omega)$. Hence, (3.3) implies $\sigma(\mu, \omega(\mu)) - \sigma_2(\mu, \omega(\mu)) \geq \widehat{\varepsilon} > \varepsilon$.

To show that $\omega(\mu)$ is the unique global maximizer of $\sigma(\mu, \cdot)$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, we introduce

$$\delta_1(\mu) := \sup\{\sigma(\mu, \omega) \mid \omega \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})\}, \quad \delta_2(\mu) := \sup\{\sigma(\mu, \omega) \mid \omega \in \mathbb{R} \setminus \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})\},$$

and let $\delta_* := \delta_1(\mu_*) - \delta_2(\mu_*) > 0$. As argued at the beginning of the proof of Lemma 3.3, there exists a neighborhood $\widetilde{\mathcal{N}}$ of (μ_*, ω_*) where the transfer function $(\mu, \omega) \mapsto H[\mu](i\omega)$ is

continuously differentiable. As a result, the largest singular value function $\sigma(\cdot, \cdot)$ is Lipschitz continuous, say with the Lipschitz constant ζ over $\tilde{\mathcal{N}}$ which we assume contains $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \times \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ without loss of generality. The functions $\delta_1(\cdot)$ and $\delta_2(\cdot)$ are also Lipschitz continuous with the Lipschitz constant ζ over $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, see [12, Lemma 8 (ii)] (that concerns the minimization of a smallest singular value rather than the maximization of a largest singular value as in here, but the proof over there can be modified in a straightforward manner). We furthermore assume $\eta_{\mu,0} < \delta_*/(4\zeta)$ without loss of generality (since we can choose $\eta_{\mu,0}$ as small as we wish), so

$$\delta_1(\mu) \geq \delta_1(\mu_*) - \delta_*/4 \quad \text{and} \quad \delta_2(\mu_*) \geq \delta_2(\mu) - \delta_*/4$$

for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ by the Lipschitz continuity of $\delta_1(\cdot)$ and $\delta_2(\cdot)$. These inequalities combined with $\delta_1(\mu_*) - \delta_2(\mu_*) = \delta_*$ yield

$$\delta_1(\mu) - \delta_2(\mu) \geq \delta_1(\mu_*) - \delta_2(\mu_*) - \delta_*/2 = \delta_*/2$$

for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$. This means that any global maximizer $\tilde{\omega}(\mu)$ of $\sigma(\mu, \cdot)$ lies in the interior of $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$. Since $\sigma(\cdot, \cdot)$ is differentiable in a neighborhood of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \times \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$, we must have $\sigma_\omega(\mu, \tilde{\omega}(\mu)) = 0$. We conclude from part (i) that $\tilde{\omega}(\mu) = \omega(\mu)$ and that it is the unique global maximizer of $\sigma(\mu, \cdot)$.

(iv) We assume without loss of generality that $\eta_{\mu,0} \leq \tilde{\eta}_\mu$ and $\eta_{\omega,0} \leq \tilde{\eta}_\omega$. Consequently, $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu) \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0}) \subseteq \overline{\mathcal{B}}(\omega_*, \tilde{\eta}_\omega)$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \subseteq \overline{\mathcal{B}}(\mu_*, \tilde{\eta}_\mu)$, so (3.4) yields $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \geq \varepsilon$ for such μ . The uniqueness of $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ as the stationary point of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ in $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ is immediate from the implicit function theorem. Additionally, without loss of generality, we can assume $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \times \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0}) \subseteq \mathcal{N}$ where \mathcal{N} is the neighborhood of (μ_*, ω_*) as in part (ii) over which $\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \leq \delta/2 < 0$. This means $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ is strictly concave in $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$. Thus, the unique stationary point $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ must also be the unique global maximizer of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ in $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$. \square

REMARK 3.5. *The second condition in (3.2) can be dropped in theory by including additional vectors in the subspaces and doubling the subspace dimensions as follows: The interpolation property*

$$\sigma_\omega(\mu^{(k)}, \omega^{(k-1)}) = \sigma_{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k-1)})$$

can be achieved, for instance, by the inclusions

$$\begin{aligned} \text{Col} \left(\left(i\omega^{(k-1)} E(\mu^{(k)}) - A(\mu^{(k)}) \right)^{-1} B(\mu^{(k)}) \right) &\subseteq \mathcal{V}_k, \\ \text{Col} \left(\left(i\omega^{(k-1)} E(\mu^{(k)}) - A(\mu^{(k)}) \right)^{-*} C(\mu^{(k)})^* \right) &\subseteq \mathcal{W}_k, \end{aligned}$$

when $\mathbf{m} = \mathbf{p}$. By the mean value theorem, this would lead to

$$\begin{aligned} \sigma_{\omega\omega}(\mu^{(k)}, \xi^{(k)}) &= \frac{\sigma_\omega(\mu^{(k)}, \omega^{(k)}) - \sigma_\omega(\mu^{(k)}, \omega^{(k-1)})}{\omega^{(k)} - \omega^{(k-1)}} \\ &= \frac{\sigma_{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) - \sigma_{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k-1)})}{\omega^{(k)} - \omega^{(k-1)}} = \sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \tilde{\xi}^{(k)}) \end{aligned}$$

for some $\xi^{(k)}, \tilde{\xi}^{(k)}$ in the open interval with the end-points $\omega^{(k-1)}, \omega^{(k)}$ so that

$$\left| \sigma_{\omega\omega}(\mu^{(k)}, \omega^{(k)}) - \sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) \right| = \mathcal{O} \left(\left| \omega^{(k)} - \omega^{(k-1)} \right| \right).$$

Hence, by the continuity of the second derivatives with respect to ω and $(\mu^{(k)}, \omega^{(k)}) \rightarrow (\mu_*, \omega_*)$, the condition $\sigma_{\omega\omega}(\mu_*, \omega_*) =: \hat{\delta} < 0$ would imply $\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu_*, \omega_*) < \hat{\delta}/2 < 0$ provided $\mu^{(1)}$ is chosen sufficiently close to μ_* .

The main conclusion of the next result is that the higher-order derivatives of $\mu \mapsto \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))$ are uniformly bounded in absolute value by a constant.

LEMMA 3.6 (Uniform boundedness of higher-order derivatives). *Suppose that the assumptions of Proposition 3.4 hold. Furthermore, let $\eta_{\mu,0}$ and $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ be as in Proposition 3.4, and let*

$$(3.6) \quad \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) := \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)).$$

Then for every $\hat{\eta}_{\mu,0} \in (0, \eta_{\mu,0})$ there exists $\gamma > 0$ independent of $\mu^{(1)}$ such that for all $\mu \in \overline{\mathcal{B}}(\mu_*, \hat{\eta}_{\mu,0})$, we have

$$(i) \quad \left| \frac{\partial^2 \|H[\mu]\|_{\mathcal{H}_\infty}}{\partial \mu_q \partial \mu_r} \right| \leq \gamma \quad \text{and} \quad \left| \tilde{\sigma}_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) \right| \leq \gamma, \quad q, r = 1, \dots, d,$$

$$(ii) \quad \left| \frac{\partial^3 \|H[\mu]\|_{\mathcal{H}_\infty}}{\partial \mu_q \partial \mu_r \partial \mu_\ell} \right| \leq \gamma \quad \text{and} \quad \left| \tilde{\sigma}_{\mu_q \mu_r \mu_\ell}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) \right| \leq \gamma, \quad q, r, \ell = 1, \dots, d.$$

Proof. (i) By Proposition 3.4, the functions $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ are three times continuously differentiable in a neighborhood of $\overline{\mathcal{B}}(\mu_*, \hat{\eta}_{\mu,0})$. The first assertion, that is the boundedness of the second derivatives of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ in $\overline{\mathcal{B}}(\mu_*, \hat{\eta}_{\mu,0})$, is immediate. Let us prove the existence of a uniform $\gamma > 0$ such that

$$(3.7) \quad \left| \tilde{\sigma}_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) \right| \leq \gamma \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \hat{\eta}_{\mu,0})$$

for $q, r = 1, \dots, d$ independent of $\mu^{(1)}$. To this end, we first observe

$$(3.8) \quad \tilde{\sigma}_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) = \sigma_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) + \sigma_{\mu_q \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \omega_{\mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu).$$

The function $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ is implicitly defined by the equation $\sigma_{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) = 0$ for μ near $\mu^{(k)}$. Differentiating this equation with respect to μ_r yields

$$\omega_{\mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) = - \frac{\sigma_{\mu_r \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))}{\sigma_{\omega \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))},$$

which we plug into (3.8) to obtain

$$(3.9) \quad \tilde{\sigma}_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) = \sigma_{\mu_q \mu_r}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) - \frac{\sigma_{\mu_q \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \sigma_{\mu_r \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))}{\sigma_{\omega \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))}.$$

By part (ii) of Proposition 3.4, we have $\sigma_{\omega \omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)) \leq \delta/2 < 0$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ independent of $\mu^{(1)}$. Additionally, by Lemma 3.3, all mixed second derivatives of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ are bounded from above in absolute value uniformly in $\overline{\mathcal{B}}(\mu_*, \eta_\mu) \times \overline{\mathcal{B}}(\omega_*, \eta_\omega) \supseteq \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \times \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ (to be precise we assume the inclusion without loss of generality as we can choose $\eta_{\mu,0}, \eta_{\omega,0}$ as small as we wish), where the upper bound is independent of $\mu^{(1)}$. Hence, we conclude with (3.7) as desired.

(ii) The boundedness of the third derivatives of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ in $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ is immediate from three times continuous differentiability of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ in a neighborhood of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$.

The boundedness of the absolute values of the third derivatives of $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ uniformly by a constant independent of $\mu^{(1)}$ can be established in a similar way as in part (i). Specifically, by differentiating (3.9) with respect to μ_ℓ , it can be seen that $\tilde{\sigma}_{\mu_q \mu_r \mu_\ell}^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ is a ratio, where the expression in the numerator is a sum of products of the mixed second derivatives $\sigma_{\chi_1, \chi_2}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))$ and third derivatives $\sigma_{\chi_1, \chi_2 \chi_3}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))$ for $\chi_1, \chi_2, \chi_3 \in$

$\{\omega, \mu_q, \mu_r, \mu_\ell\}$, while the expression in the denominator is $\sigma_{\omega\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu))^3$. Hence, once again, the conclusion

$$\left| \tilde{\sigma}_{\mu_q \mu_r \mu_\ell}^{\mathcal{V}_k, \mathcal{W}_k}(\mu) \right| \leq \gamma \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \hat{\eta}_{\mu, 0})$$

for $q, r, \ell = 1, \dots, d$ for some γ independent of $\mu^{(1)}$ can be drawn from part **(ii)** of Proposition 3.4 and Lemma 3.3. \square

By Lemma 2.1 and Lemma 2.2, the reduced function $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ Hermite interpolates the original \mathcal{H}_∞ function $\|H[\cdot]\|_{\mathcal{H}_\infty}$ at $\mu = \mu^{(k)}, \mu^{(k-1)}$. Indeed, for the extended algorithm (Algorithm 2.2), these Hermite interpolation properties also hold at $\mu = \mu^{(k, rq)}$ for each $r = 1, \dots, d, q = r, \dots, d$ by Lemma 2.2. From these observations, by also employing Lemma 3.6, it is possible to conclude with an upper bound on the gap between the second derivatives of $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ and $\|H[\cdot]\|_{\mathcal{H}_\infty}$ near μ_* , which we formally state and prove next.

LEMMA 3.7 (Proximity of the second derivatives). *Suppose that the assumptions of Proposition 3.4 hold. Additionally, assume that $\nabla^2 \|H[\mu_*]\|_{\mathcal{H}_\infty}$ is invertible. Furthermore, let $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ be as in Proposition 3.4, and $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ be defined as in (3.6). Then there exists a $\zeta > 0$ such that the following statements hold for Algorithm 2.1 when $d = 1$ and for Algorithm 2.2 independent of $\mu^{(1)}$:*

- (i) We have $\left\| \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} - \nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right\|_2 \leq \zeta \|\mu^{(k)} - \mu^{(k-1)}\|_2$.
- (ii) Both $\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty}$ and $\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$ are invertible.
- (iii) We have $\left\| \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} - \left[\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \right\|_2 \leq \zeta \|\mu^{(k)} - \mu^{(k-1)}\|_2$.

Proof. **(i)** We focus on Algorithm 2.2 only (the proof for Algorithm 2.1 with $d = 1$ proceeds similarly by defining $h^{(k)} := \mu^{(k-1)} - \mu^{(k)}$). By parts **(iii)** and **(iv)** of Proposition 3.4, the functions $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ are three times differentiable in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu, 0})$ independent of $\mu^{(1)}$. Now choose $\mu^{(1)}$ close enough to μ_* so that $\overline{\mathcal{B}}(\mu^{(k)}, h^{(k)}) \subset \overline{\mathcal{B}}(\mu_*, \eta_{\mu, 0})$, as well as $\omega^{(k)}, \omega^{(k, rq)}$ belong to the interior of $\overline{\mathcal{B}}(\omega_*, \eta_{\omega, 0})$ for $r = 1, \dots, d$ and $q = r, \dots, d$ (observe that $\omega^{(k, rq)} \rightarrow \omega_*$ as $\mu^{(1)} \rightarrow \mu_*$ based on arguments similar to the ones for $\omega^{(k)} \rightarrow \omega_*$ as $\mu^{(1)} \rightarrow \mu_*$ given in the proof of part **(ii)** of Proposition 3.4).

It follows that the functions

$$\begin{aligned} \ell : [0, 1] &\rightarrow \mathbb{R}, & \ell(\alpha) &:= \|H(\mu^{(k)} + \alpha h^{(k)} e_{rq})\|_{\mathcal{H}_\infty}, \\ \tilde{\ell} : [0, 1] &\rightarrow \mathbb{R}, & \tilde{\ell}(\alpha) &:= \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)} + \alpha h^{(k)} e_{rq}) \end{aligned}$$

are continuous and three times differentiable in $(0, 1)$. Additionally, Lemma 2.2 implies that the following interpolation properties between these functions

$$(3.10) \quad \ell(0) = \tilde{\ell}(0), \quad \ell'(0) = \tilde{\ell}'(0) \quad \text{and} \quad \ell(1) = \tilde{\ell}(1)$$

are satisfied. To see the last equality at $\alpha = 1$, we observe

$$0 = \sigma_\omega(\mu^{(k, rq)}, \omega^{(k, rq)}) = \sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k, rq)}, \omega^{(k, rq)})$$

by Lemma 2.2, so part **(iv)** of Proposition 3.4, in particular the uniqueness of the stationary point $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu)$ of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \cdot)$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu, 0})$, implies $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k, rq)}) = \omega^{(k, rq)}$ (as $\mu^{(k, rq)} \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu, 0})$ and $\omega^{(k, rq)} \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega, 0})$). Hence, again by Lemma 2.2, we have

$$\begin{aligned} \ell(1) &= \|H[\mu^{(k, pq)}]\|_{\mathcal{H}_\infty} = \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k, rq)}, \omega^{(k, rq)}) \\ &= \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k, rq)}, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k, rq)})) = \tilde{\ell}(1). \end{aligned}$$

By employing the interpolation properties in (3.10) in the Taylor expansions

$$\begin{aligned}\ell(1) &= \ell(0) + \ell'(0) + \frac{1}{2}\ell''(0) + \frac{1}{6}\ell'''(\varepsilon), \\ \tilde{\ell}(1) &= \tilde{\ell}(0) + \tilde{\ell}'(0) + \frac{1}{2}\tilde{\ell}''(0) + \frac{1}{6}\tilde{\ell}'''(\tilde{\varepsilon})\end{aligned}$$

for some $\varepsilon, \tilde{\varepsilon} \in (0, 1)$, we obtain

$$(3.11) \quad \begin{aligned} & [h^{(k)}]^\top e_{rq} \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} - \nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right] e_{rq} \\ &= \ell''(0) - \tilde{\ell}''(0) = \frac{1}{3} \left(\tilde{\ell}'''(\tilde{\varepsilon}) - \ell'''(\varepsilon) \right) = \mathcal{O} \left([h^{(k)}]^3 \right), \end{aligned}$$

where the constant hidden in the Landau symbol \mathcal{O} is independent of $\mu^{(1)}$ due to Lemma 3.6. By considering particular values of $r = 1, \dots, d$ and $q = r, \dots, d$ in (3.11), we deduce

$$\left| \frac{\partial^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty}}{\partial \mu_r \partial \mu_q} - \frac{\partial^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})}{\partial \mu_r \partial \mu_q} \right| = \mathcal{O}(h^{(k)}).$$

Once again, the constant hidden in the Landau symbol \mathcal{O} does not depend on $\mu^{(1)}$ in the latter equation.

(ii) By the continuity of $\nabla^2 \|H[\cdot]\|_{\mathcal{H}_\infty}$ in the interior of $\bar{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, coupled with the assumption $\mu^{(k)} \rightarrow \mu_*$ as $\mu^{(1)} \rightarrow \mu_*$, we have $\lim_{\mu^{(1)} \rightarrow \mu_*} \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} = \nabla^2 \|H[\mu_*]\|_{\mathcal{H}_\infty}$. Consequently, $\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty}$ is invertible provided $\mu^{(1)}$ is sufficiently close to μ_* . In addition, from part (i) we get

$$\nabla^2 \|H[\mu_*]\|_{\mathcal{H}_\infty} = \lim_{\mu^{(1)} \rightarrow \mu_*} \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} = \lim_{\mu^{(1)} \rightarrow \mu_*} \nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}),$$

implying also the invertibility of $\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$ for $\mu^{(1)}$ close to μ_* .

(iii) This statement follows from part (i) by employing the adjugate formulas for the inverses of $\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty}$ as well as $\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$. For details, we refer to [10, Lemma 2.8, part (ii)]. \square

Now we are ready for the main rate-of-convergence result.

THEOREM 3.8 (Local superlinear convergence). *Suppose that the assumptions of Proposition 3.4 hold. In particular, let $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ be as in Proposition 3.4. Additionally, assume that the matrix $\nabla^2 \|H[\mu_*]\|_{\mathcal{H}_\infty}$ is invertible, the point μ_* is strictly in the interior of $\underline{\Omega}$, and that the function $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}, \cdot)$ has a unique global maximizer, say at $\tilde{\omega}^{(k+1)}$, with $\lim_{\mu^{(1)} \rightarrow \mu_*} \tilde{\omega}^{(k+1)} = \omega_*$. Regarding Algorithm 2.1 when $d = 1$ and Algorithm 2.2, the following statement holds: There exists a $C > 0$ independent of $\mu^{(1)}$ such that*

$$(3.12) \quad \frac{\|\mu^{(k+1)} - \mu_*\|_2}{\|\mu^{(k)} - \mu_*\|_2 \max \{ \|\mu^{(k)} - \mu_*\|_2, \|\mu^{(k-1)} - \mu_*\|_2 \}} \leq C.$$

Proof. By Proposition 3.4, both $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ defined by (3.6) are twice Lipschitz continuously differentiable in the interior of the ball $\bar{\mathcal{B}}(\mu_*, \eta_{\mu,0})$. Now suppose $\mu^{(1)}$ is close enough to μ_* so that $\mu^{(k+1)}, \mu^{(k)}, \mu^{(k-1)}$ lie in the interior of $\bar{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, whereas $\tilde{\omega}^{(k+1)}$ belongs to the interior of $\bar{\mathcal{B}}(\omega_*, \eta_{\omega,0})$, and

- (1) $\bar{\mathcal{B}}(\mu^{(k)}, h^{(k)}) \subset \bar{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ where $h^{(k)} := \|\mu^{(k)} - \mu^{(k-1)}\|_2$ (recall that $\mu^{(k)} \rightarrow \mu_*$ and $h^{(k)} \rightarrow 0$ as $\mu^{(1)} \rightarrow \mu_*$),
- (2) $\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty}$ and $\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$ are invertible (part (ii) of Lemma 3.7 ensures this as $\mu^{(1)}$ is chosen close to μ_*).

By an application of Taylor's theorem with integral remainder we obtain

$$0 = \nabla \|H[\mu_*]\|_{\mathcal{H}_\infty} = \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} + \int_0^1 \nabla^2 \|H[\mu^{(k)} + t(\mu_* - \mu^{(k)})]\|_{\mathcal{H}_\infty} (\mu_* - \mu^{(k)}) dt,$$

which implies

$$(3.13) \quad \begin{aligned} 0 &= \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} + (\mu_* - \mu^{(k)}) + \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} \\ &\quad \times \int_0^1 \left[\nabla^2 \|H[\mu^{(k)} + t(\mu_* - \mu^{(k)})]\|_{\mathcal{H}_\infty} - \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right] (\mu_* - \mu^{(k)}) dt. \end{aligned}$$

Now by exploiting the interpolation property, in particular part **(iv)** of Lemma 2.1, and recalling $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) = \omega^{(k)}$ due to part **(ii)** of Proposition 3.4, we get

$$\begin{aligned} \nabla \|H(\mu^{(k)})\|_{\mathcal{H}_\infty} &= \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) \\ &= \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})) \\ &= \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})) + \sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})) \nabla \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \\ &= \nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}), \end{aligned}$$

where we employ $\sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})) = 0$ for the third equality. Hence, equation (3.13) can be rearranged as

$$(3.14) \quad \begin{aligned} 0 &= \left[\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) + (\mu_* - \mu^{(k)}) + \\ &\quad \left\{ \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} - \left[\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \right\} \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} + \\ &\quad \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} \\ &\quad \times \int_0^1 \left[\nabla^2 \|H[\mu^{(k)} + t(\mu_* - \mu^{(k)})]\|_{\mathcal{H}_\infty} - \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right] (\mu_* - \mu^{(k)}) dt. \end{aligned}$$

Throughout the rest of the proof, by manipulating (3.14), we bound $\|\mu^{(k+1)} - \mu_*\|_2$ from above in terms of $\|\mu^{(k)} - \mu_*\|_2$ and $\|\mu^{(k-1)} - \mu_*\|_2$. Since $\tilde{\omega}^{(k+1)} \in \bar{\mathcal{B}}(\omega_*, \eta_{\omega, 0})$ is assumed to be the unique global maximizer of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}, \cdot)$, we must have $\omega^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}) = \tilde{\omega}^{(k+1)}$ by part **(iv)** of Proposition 3.4. It follows that

$$\nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}) = \sigma_\mu^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}, \tilde{\omega}^{(k+1)}) = \nabla \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k+1)}]\|_{\mathcal{H}_\infty} = 0,$$

where we use the fact that $\mu^{(k+1)}$ is a maximizer of $\|H^{\mathcal{V}_k, \mathcal{W}_k}[\cdot]\|_{\mathcal{H}_\infty}$ for the last equality. Moreover, a Taylor expansion yields

$$\begin{aligned} 0 &= \nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}) \\ &= \nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) + \nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) (\mu^{(k+1)} - \mu^{(k)}) + \mathcal{O} \left(\|\mu^{(k+1)} - \mu^{(k)}\|_2^2 \right), \end{aligned}$$

which in turn implies

$$(3.15) \quad \left[\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \nabla \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) = (\mu^{(k)} - \mu^{(k+1)}) + \mathcal{O} \left(\|\mu^{(k+1)} - \mu^{(k)}\|_2^2 \right).$$

Additionally, by another Taylor expansion,

$$\begin{aligned} 0 &= \nabla \|H[\mu_*]\|_{\mathcal{H}_\infty} \\ &= \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} + \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} (\mu_* - \mu^{(k)}) + \mathcal{O} \left(\|\mu^{(k)} - \mu_*\|_2^2 \right). \end{aligned}$$

Therefore, by using Lemma 3.7 and part (i) of Lemma 3.6, we see that

$$(3.16) \quad \left\| \left\{ \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} - \left[\nabla^2 \tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \right\} \cdot \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right\|_2 \\ \leq \zeta \|\mu^{(k)} - \mu^{(k-1)}\|_2 \cdot \left\| \nabla \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right\|_2 = \mathcal{O} \left(\|\mu^{(k)} - \mu^{(k-1)}\|_2 \cdot \|\mu^{(k)} - \mu_*\|_2 \right).$$

Finally, by exploiting the Lipschitz continuity of $\nabla^2 \|H[\cdot]\|_{\mathcal{H}_\infty}$ near μ_* , we obtain

$$(3.17) \quad \left\| \left[\nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right]^{-1} \right. \\ \left. \times \int_0^1 \left[\nabla^2 \|H[\mu^{(k)} + t(\mu_* - \mu^{(k)})]\|_{\mathcal{H}_\infty} - \nabla^2 \|H[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right] (\mu_* - \mu^{(k)}) dt \right\|_2 \\ = \mathcal{O} \left(\|\mu^{(k)} - \mu_*\|_2^2 \right).$$

Combining (3.14) with (3.15), (3.16), (3.17), and noting $\|\mu^{(k)} - \mu^{(k-1)}\|_2 \leq 2 \max \{ \|\mu^{(k)} - \mu_*\|_2, \|\mu^{(k-1)} - \mu_*\|_2 \}$, we finally obtain

$$\|\mu^{(k+1)} - \mu_*\|_2 \leq c_1 \max \{ \|\mu^{(k)} - \mu_*\|_2, \|\mu^{(k-1)} - \mu_*\|_2 \} \|\mu^{(k)} - \mu_*\|_2 + c_2 \|\mu^{(k)} - \mu_*\|_2^2$$

for some constants c_1, c_2 independent of $\mu^{(1)}$ from which (3.12) is immediate. \square

REMARK 3.9. *One important assumption for the rate of convergence result above is that the global minimizer μ_* is contained in the interior of $\underline{\Omega}$. Suppose $\underline{\Omega}$ is a box, and μ_* lies on the boundary of this box. Then one or more of the box constraints are active for the full-order problem at μ_* , and $\|H[\cdot]\|_{\mathcal{H}_\infty}$ is increasing in all directions pointing into the interior of $\underline{\Omega}$ in a ball $\bar{\mathcal{B}}(\mu_*, \eta)$ (as $\|H[\cdot]\|_{\mathcal{H}_\infty}$ is continuously differentiable in a neighborhood of μ_*). The same property holds to be true for the reduced function $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ in another ball $\bar{\mathcal{B}}(\mu_*, \tilde{\eta}) \subseteq \bar{\mathcal{B}}(\mu_*, \eta)$, due to the interpolation properties (specifically due to part (iii) of Lemma 2.1), and uniform upper bounds on the derivatives of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$, $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ (see in particular Lemma 3.3 and 3.6), provided $\mu^{(1)}$ is close enough to μ_* . Consequently, the same active box constraints for the original function $\|H[\cdot]\|_{\mathcal{H}_\infty}$ at μ_* have to be active for the reduced function $\tilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ at $\mu^{(k+1)}$. This means that the rate of convergence analysis above, in particular the proof of Theorem 3.8, is applicable by restricting μ to the variables that are not active at μ_* . If all of the constraints are active at μ_* , then $\mu^{(k+1)} = \mu_*$ in exact arithmetic.*

The minimizers for the examples arising from real applications on which we perform numerical experiments in the next section turn out to be on the boundary of the box, see, e. g., Example 4.1 where all of the three box constraints are active at the minimizer, or Example 4.3 where only one of the two box constraints is active, while the other is inactive. On the other hand, the minimizer for the synthetic example in the next section is usually in the interior, see Example 4.4.

4. Numerical Experiments. In this section, we present numerical results obtained by our MATLAB implementation of Algorithm 2.1 that we made available for download. We first discuss some important implementation details and the test setup in the next subsection. Then, we report the numerical results on several large-scale linear parameter-dependent systems which we describe in detail. All test examples are taken from the Model Order Reduction Wiki (MOR Wiki) website¹. Our numerical experiments have been performed on a machine with an 4 Intel[®] Core[™] i5-4590 CPUs with 3.30GHz each and 16GB RAM using Linux version 4.4.132-53-default and MATLAB version 9.4.0.813654 (R2018a).

¹available at https://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Main_Page.

4.1. Implementation Details and Test Setup. At each iteration of Algorithm 2.1, the \mathcal{L}_∞ -norm of the transfer function of a reduced parametrized system needs to be minimized. We have implemented and tested two optimization techniques to solve this global non-convex optimization problem:

- **eigopt**, a MATLAB implementation of the algorithm in [13], which is an adaptation of the algorithm in [5] for eigenvalue optimization. This MATLAB package creates a lower and an upper bound for the optimal value of a given eigenvalue function by employing piece-wise quadratic support functions, and terminates when the difference between these bounds is less than a prescribed tolerance. For reliability and efficiency, one should supply an appropriate global lower bound γ on the minimum eigenvalue of the Hessian of the eigenvalue function to be minimized to **eigopt**. This solver can be slow, if there are many parameters or if γ is very small. For our tests we always use $\gamma = -10000$.
- **GRANSO** [6], which is based on BFGS together with line searches ensuring the satisfaction of the weak Wolfe conditions. **GRANSO** converges to a locally optimal solution, that is not necessarily optimal globally, but works efficiently even when there are several parameters.

Algorithm 2.1 is terminated in practice when the relative distance between $\mu^{(k)}$ and $\mu^{(k-1)}$ is less than a prescribed tolerance for some $k > 1$, if the minimal \mathcal{L}_∞ -norm values for the reduced transfer functions at two consecutive iterations differ by less than a prescribed tolerance, or if the number of iterations exceeds a specified integer, more formally, we terminate, if

$$k > k_{\max} \quad \text{or} \quad \|\mu^{(k)} - \mu^{(k-1)}\|_2 < \varepsilon_1 \cdot \frac{1}{2} \|\mu^{(k)} + \mu^{(k-1)}\|_2 \quad \text{or} \\ \left| \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k+1)}]\|_{\mathcal{H}_\infty} - \|H^{\mathcal{V}_{k-1}, \mathcal{W}_{k-1}}[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right| < \\ \varepsilon_2 \cdot \frac{1}{2} \left\{ \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k+1)}]\|_{\mathcal{H}_\infty} + \|H^{\mathcal{V}_{k-1}, \mathcal{W}_{k-1}}[\mu^{(k)}]\|_{\mathcal{H}_\infty} \right\}.$$

In our numerical experiments, we set $\varepsilon_1 = \varepsilon_2 = 10^{-6}$ and $k_{\max} = 20$.

The absolute termination tolerance for the accuracy of the global optimizer computed by **eigopt** is 10^{-8} , whereas the tolerance for reaching (approximate) stationarity in **GRANSO** is set to 10^{-12} . Apart from these we use default options in **eigopt**, **GRANSO**, as well as our MATLAB routine **linorm_subsp** that implements the method from [1] for computing the \mathcal{L}_∞ -norm of the transfer function of a large-scale linear system. In **linorm_subsp** we call the FORTRAN routine **AB13HD.F** via a mex file that implements the method of [4] to compute the \mathcal{L}_∞ -norm of small-scale reduced systems. The latter is often faster and more reliable than the native MATLAB routine **norm** from the Control Systems Toolbox, that one could use for small-scale \mathcal{L}_∞ -norm computations as well. Our initial reduced order models are generated by 10 interpolation points (which consist of pairs of parameter values μ and frequencies ω) that are equidistantly aligned on a line in $\underline{\Omega} \times [0, \omega_{\max})$ where ω_{\max} is a problem-dependent maximum frequency. Further details on the implementation can be inferred from the code that we have made available for download.

4.2. Results for Real Examples. We first test our algorithm on the following four parameter-dependent descriptor systems, all of which originate from real applications.

EXAMPLE 4.1 (Thermal conduction (T2DAL_BCI), see [14]). *Our first example is a thermal conduction model in a chip production. For a compact and efficient model of thermal conduction, one should take into account different configurations of the boundary conditions. This gives the capability to the chip producers to assess how the change in the environment influences the temperature in the chip. A mathematical model of the thermal conduction is given by the heat equation where the heat exchange through the three device interfaces is modeled by convection boundary conditions. These boundary conditions introduce the parameters*

μ_1, μ_2, μ_3 , called the film coefficients, to describe the change in the temperature on the three device interfaces. After spatial discretization of the partial differential equation and by incorporating the boundary conditions one obtains a time-invariant linear system with transfer function

$$(4.1) \quad H[\mu_1, \mu_2, \mu_3](s) = C(sE - (A_0 + \mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3))^{-1} B$$

where $E \in \mathbb{R}^{4257 \times 4257}$ and $A_i \in \mathbb{R}^{4257 \times 4257}$, $i = 1, 2, 3$ are diagonal matrices arising from the discretization of convection boundary conditions on the i -th interface and $B \in \mathbb{R}^{4257 \times 1}$, $C \in \mathbb{R}^{7 \times 4257}$ are the input and output matrices, respectively. The specified box for the parameter $\mu := \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$ is $[1, 10^4] \times [1, 10^4] \times [1, 10^4]$.

We report on the results of Algorithm 2.1 on the T2DAL_BCI example for different setups in Table 4.1.

Table 4.1: Numerical results for the T2DAL_BCI example.

setup	n_{iter}	$(\mu_{1,*}, \mu_{2,*}, \mu_{3,*})$	$\ H[\mu_{1,*}, \mu_{2,*}, \mu_{3,*}]\ _{\mathcal{H}_\infty}$	time in s
eigopt	2	(1.0000e+4, 1.0000e+4, 1.0000e+4)	1.15429e+1	374.25
GRANSO	2	(1.0000e+4, 1.0000e+4, 1.0000e+4)	1.15429e+1	2.54

EXAMPLE 4.2 (Anemometer (`anemometer_1p` and `anemometer_3p`), see [2]). An anemometer is a device to measure heat flow which consists of a heater and temperature sensors placed near the heater. The temperature field is affected by the flow and hence a temperature difference occurs between the sensors. The measured temperature difference determines the velocity of the fluid flow. The mathematical model for the anemometer is given by the convection-diffusion equation

$$\rho c \frac{\partial T}{\partial t} = \nabla(\kappa \nabla T) - \rho c v \nabla T + q',$$

where ρ denotes the mass density, c is the specific heat, κ is the thermal conductivity, v is the fluid velocity, T is the temperature, and q' is the heat flow. A spatial discretization of the convection-diffusion equation above, for instance by the finite element method, yields a parametric linear system with the transfer function

$$H[v](s) = C(sE - (A_1 + v(A_2 - A_1)))^{-1} B$$

which depends on only the fluid velocity $v \in [0, 1]$; or a parametric system with the transfer function

$$H[c, \kappa, v](s) = C(s(E_1 + cE_2) - (A_1 + \kappa A_2 + cvA_3))^{-1} B$$

where three parameters $c \in [0, 1]$, $\kappa \in [1, 2]$, $v \in [0.1, 2]$ appear. The input and output matrices B and C above result from separating the spatial variables in q' . We refer to these one parameter and three parameter examples as `anemometer_1p` and `anemometer_3p`, respectively. In both cases, the order of the state space is 29008, there is a single input and a single output.

We report on the results of Algorithm 2.1 on the `anemometer_1p` and `anemometer_3p` examples for different setups in Tables 4.2 and 4.3, respectively.

EXAMPLE 4.3 (Scanning electrochemical microscopy (SECM), see [7]). Scanning electrochemical microscopy is a technique to analyze the electrochemical behavior of species (in different states of matter) at their interface. This example considers the chemical reaction between two species on an electrode. The species transport is described by the second Fick's law which leads to two partial diffusion equations with appropriate boundary conditions. A

Table 4.2: Numerical results for the `anemometer_1p` example.

setup	n_{iter}	v_*	$\ H[v_*]\ _{\mathcal{H}_\infty}$	time in s
<code>eigopt</code>	6	0.0000	1.32274e-2	32.68
<code>GRANSO</code>	6	-0.0000	1.32274e-2	30.91

Table 4.3: Numerical results for the `anemometer_3p` example.

setup	n_{iter}	(c_*, κ_*, v_*)	$\ H[c_*, \kappa_*, v_*]\ _{\mathcal{H}_\infty}$	time in s
<code>eigopt</code>	4	(0.0000, 2.0000, 1.0000e-1)	1.64723e-3	766.06
<code>GRANSO</code>	3	(0.0000, 2.0000, 8.3855e-1)	1.64723e-3	40.93

spatial discretization together with a boundary control then leads to a linear-time invariant system whose transfer function is

$$H[h_1, h_2](s) = C(sE - (h_1A_1 + h_2D_2 - A_3))^{-1}B,$$

where $E, A_1, A_2, A_3 \in \mathbb{R}^{16912 \times 16912}$, $B \in \mathbb{R}^{16912 \times 1}$, and $C \in \mathbb{R}^{5 \times 16912}$ and h_1, h_2 are the parameters of the problem. The experiment is performed in the box $[1, e^2] \times [1, e^2]$.

The results for the `SECM` example are summarized in Table 4.4

Table 4.4: Numerical results for the `SECM` example.

setup	n_{iter}	$(h_{1,*}, h_{2,*})$	$\ H[h_{1,*}, h_{2,*}]\ _{\mathcal{H}_\infty}$	time in s
<code>eigopt</code>	5	(1.0000, 4.1944)	1.85588	180.01
<code>GRANSO</code>	5	(1.0000, 4.2882)	1.85583	20.51

In all examples, we observe superlinear convergence in the final iterations. Specifically, for the `SECM` example, we report the errors with respect to the iteration number in Table 4.5. Four additional iterations after the construction of the initial reduced model suffice to find the minimal \mathcal{H}_∞ -norm with the specified relative tolerances. For most examples, in particular the ones with more than one parameter, using `GRANSO` is significantly faster than `eigopt`. On the other hand, in contrast to `GRANSO`, `eigopt` returns the global minimizer for the reduced problems and thus sometimes yields more reliable results. In particular, due to the local convergence issue with `GRANSO`, rarely the subspace framework equipped with `GRANSO` does not converge to the global minimizer of the full problem, while the one with `eigopt` does converge to the global minimizer of the full problem. This can for example be seen in the synthetic example discussed below.

To our knowledge, there is no reliable and efficient algorithm for large-scale \mathcal{H}_∞ -norm minimization in the literature which we can use for comparison purposes and verify the correctness of the results obtained. Instead, for each example above, we have computed the \mathcal{H}_∞ -norm of the system for various values of μ near the computed optimal parameter value μ_* . According to these computations, the optimal parameter values listed above seem to be at least locally optimal. For three of the examples, the plots of the \mathcal{H}_∞ -norm as a function of μ are illustrated in Figure 4.1

4.3. Results for Synthetic Examples. Next, we test our approach on synthetic examples of various orders taken from the MOR Wiki.

EXAMPLE 4.4 (Synthetic example). *We consider parametric single-input, single-output systems of order $n = 2q$ with transfer functions of the form*

$$(4.2) \quad H[\mu](s) = C(sI_n - \mu A_1 - A_0)^{-1}B,$$

Table 4.5: The minimizers for the reduced problems as well as the errors of the iterates of Algorithm 2.1 and the corresponding errors in the \mathcal{H}_∞ -norms are listed for the SECM example by using GRANSO for optimization. Here, the short-hands $f^{(k)} := \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k+1)}]\|_{\mathcal{L}_\infty}$ and $f_* := \|H[\mu_*]\|_{\mathcal{H}_\infty}$ are used.

k	$\mu^{(k+1)}$	$\ \mu^{(k+1)} - \mu_*\ _2$	$ f^{(k)} - f_* $
0	(1.0000, 1.2088)	3.08	2.61e-1
1	(1.6579, 7.3891)	3.17	2.72e-4
2	(1.4761, 6.3522)	2.12	1.51e-4
3	(1.0000, 4.2882)	1.55e-9	1.24e-12
4	(1.0000, 4.2882)	< 1e-12	< 1e-12

where the matrices $A_1, A_0 \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$ are given by

$$A_1 = \begin{bmatrix} A_{1,1} & & \\ & \ddots & \\ & & A_{1,m} \end{bmatrix}, \quad A_0 = \begin{bmatrix} A_{0,1} & & \\ & \ddots & \\ & & A_{0,m} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix}, \quad C = [C_1 \quad \dots \quad C_m]$$

with

$$A_{1,i} = \begin{bmatrix} a_i & 0 \\ 0 & a_i \end{bmatrix}, \quad A_{0,i} = \begin{bmatrix} 0 & b_i \\ -b_i & 0 \end{bmatrix}, \quad B_i = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad C_i = [1 \quad 0], \quad i = 1, \dots, m.$$

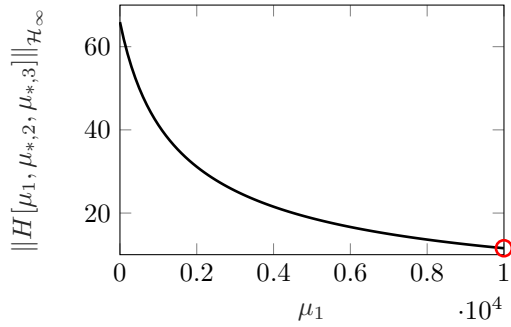
The numbers a_i and b_i are chosen equidistantly in the intervals $[-10^3, -10]$ and $[10, 10^3]$, respectively. The parameter μ is constrained to lie in the interval $[0.02, 1]$.

We perform our experiments on this synthetic example for several values of n varying in $10^2, \dots, 10^6$. For smaller values of n , a comparison of Algorithm 2.1 and the MATLAB package `eigopt` (for the unreduced problems) is provided in Table 4.7. This table indicates that with or without reduction we retrieve exactly the same optimal \mathcal{H}_∞ -norm values up to the prescribed tolerance $\varepsilon_2 = 10^{-6}$, yet the proposed subspace framework leads to speed-ups on the order of 10^3 , indeed the ratios of the runtimes increase quickly with respect to n .

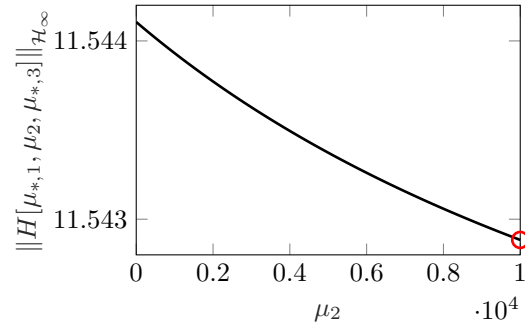
Larger examples are considered in Table 4.8, but only using the proposed subspace framework. It does not seem possible to solve these larger \mathcal{H}_∞ -norm minimization problems in a reasonable time without employing reductions. Even the examples of order 10^6 can be solved very fast. All examples up to order 10^5 can be solved with just two to four iterations, only for very large examples up to 9 iterations may be needed. Moreover, the largest fraction of the computation time is spent for solving large-scale linear systems.

Note that we have used `eigopt` for the optimization of the the small subproblems here which is guaranteed to return a global minimizer. We observe in practice that when the reduced problems are solved by a locally convergent algorithm, convergence to $\mu = 1$, a locally optimal solution, sometimes occurs. This is in particular the case for some values of n when the reduced problems are solved with GRANSO. Also note that for the computation of the \mathcal{L}_∞ -norm in this example we make use of the native MATLAB function `norm`, since the periodic QZ algorithm used for the eigenvalue computation in `AB13HD.f` does not converge always. Further, we have set $\gamma = -1000$ in `eigopt` – otherwise, the runtimes would be higher.

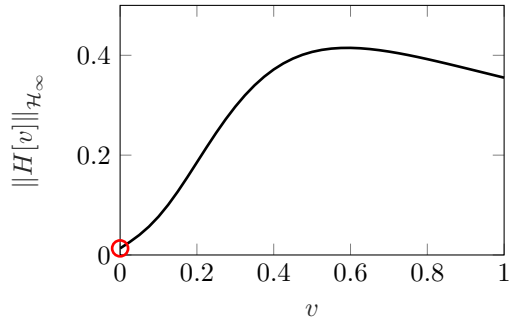
Finally, the progress of the subspace framework is displayed in Figure 4.2 on this synthetic example for $n = 500$. After one subspace iteration, the \mathcal{L}_∞ -norm of the reduced problem already closely resembles the one for the original problem around the minimizer. After two subspace iterations, it is even hard to distinguish the \mathcal{L}_∞ -norm functions for the reduced and original problems around the minimizer, except for a thin peak that occurs in the reduced problem. The progress of the iteration is further summarized in Table 4.6.



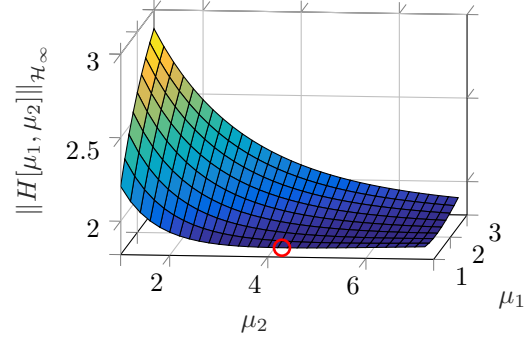
(a) Graph of $\mu_1 \mapsto \|H[\mu_1, \mu_{*,2}, \mu_{*,3}]\|_{\mathcal{H}_\infty}$ for the T2DAL_BCI example.



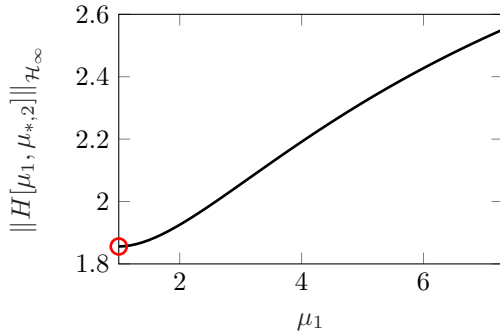
(b) Graph of $\mu_2 \mapsto \|H[\mu_{*,1}, \mu_2, \mu_{*,3}]\|_{\mathcal{H}_\infty}$ for the T2DAL_BCI example.



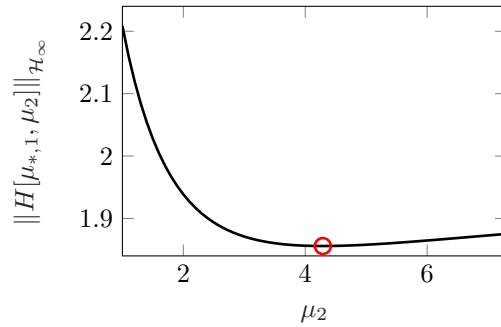
(c) Graph of $v \mapsto \|H[v]\|_{\mathcal{H}_\infty}$ for the anemometer_1p example.



(d) Graph of $(\mu_1, \mu_2) \mapsto \|H[\mu_1, \mu_2]\|_{\mathcal{H}_\infty}$ for the SECM example.



(e) Graph of $\mu_1 \mapsto \|H[\mu_1, \mu_{*,2}]\|_{\mathcal{H}_\infty}$ for the SECM example.



(f) Graph of $\mu_2 \mapsto \|H[\mu_{*,1}, \mu_2]\|_{\mathcal{H}_\infty}$ for the SECM example.

Fig. 4.1: The \mathcal{H}_∞ -norms for the different examples, where the computed minimal norm value is marked by a red circle. Note that in the captions and legends of (a)–(f), $\mu_{*,j}$ denotes the j th component of μ_* for $j = 1, 2$, or 3 .

5. Concluding Remarks. In this work we have developed new subspace restriction techniques to minimize the \mathcal{H}_∞ -norm of transfer functions of large-scale parameter-dependent linear systems. We have given a detailed analysis of the rate of convergence of these methods, demonstrated the validity of the deduced rate of convergence results in practice by various numerical examples, which could all be solved very efficiently. The methods presented here make the design of optimal \mathcal{H}_∞ -controllers for large-scale systems partly feasible. A fully feasible method to design optimal \mathcal{H}_∞ -controllers for large-scale systems should also take

Table 4.6: The minimizers for the reduced problems as well as the errors of the iterates of Algorithm 2.1 and the corresponding errors in the \mathcal{H}_∞ -norms are listed for the **synthetic** example for $n = 500$ by using `eigopt` for optimization. Here, again the short-hands $f^{(k)} := \|H^{\mathcal{V}_k, \mathcal{W}_k}[\mu^{(k+1)}]\|_{\mathcal{L}_\infty}$ and $f_* := \|H[\mu_*]\|_{\mathcal{H}_\infty}$ are used.

k	$\mu^{(k+1)}$	$\ \mu^{(k+1)} - \mu_*\ _2$	$ f^{(k)} - f_* $
0	0.5100	2.74e-1	0.12e-1
1	0.2354	2.94e-4	5.01e-7
2	0.2357	< 1e-12	< 1e-12

Table 4.7: Results of the numerical experiments on Example 4.4 for smaller values of n , where we list the number of subspace iterations n_{iter} , the optimal parameter values by Algorithm 2.1 and `eigopt`, and the corresponding minimal \mathcal{H}_∞ -norms, as well as the runtimes are listed. The optimal \mathcal{H}_∞ -norm values returned by Algorithm 2.1 are the same with those returned by `eigopt` at least up to six decimal digits.

n	n_{iter}	μ_*		$\ H[\mu_*]\ _{\mathcal{H}_\infty}$		runtime in s	
		Alg. 2.1	eigopt	Alg. 2.1	eigopt	Alg. 2.1	eigopt
100	2	1.000000	1.000000	0.317092	0.317092	1.33	6.98
200	2	1.000000	1.000000	0.549800	0.549800	0.82	52.33
400	3	0.270587	0.270549	0.969289	0.969289	3.85	455.07
600	4	0.212279	0.212255	1.337220	1.337219	3.06	1563.83
800	2	0.181492	0.181501	1.706940	1.706940	1.65	2635.76

stability considerations into account. We intend to address stability issues in future.

Code Availability. The MATLAB implementation of our algorithm and the computational results are publicly available under the URL <http://www.tu-berlin.de/?202212&L=1>.

Appendix A. Proof of Lemma 3.3. By the continuity of $(\mu, \omega) \mapsto \sigma_{\min}(D(\mu, \omega))$, there exists a neighborhood $\tilde{\mathcal{N}}$ of (μ_*, ω_*) such that $\sigma_{\min}(D(\mu, \omega)) \geq \beta/2$ for all $(\mu, \omega) \in \tilde{\mathcal{N}}$. Consequently, the mapping $(\mu, \omega) \mapsto H[\mu](i\omega)$ is continuously differentiable and $\sigma(\cdot, \cdot)$, $\sigma_2(\cdot, \cdot)$ are continuous in $\tilde{\mathcal{N}}$. The continuity of $\sigma(\cdot, \cdot)$, $\sigma_2(\cdot, \cdot)$ implies that $\sigma(\mu, \omega)$ remains a simple singular value of $H[\mu](i\omega)$, hence it is bounded away from zero in a neighborhood $\mathcal{N} \subseteq \tilde{\mathcal{N}}$ of (μ_*, ω_*) . Formally,

$$(A.1) \quad \sigma(\mu, \omega) - \sigma_2(\mu, \omega) \geq \hat{\varepsilon} \quad \forall (\mu, \omega) \in \mathcal{N}$$

for some $\hat{\varepsilon} > 0$.

Moreover, by employing the interpolation properties

$$\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) = \sigma(\mu^{(k)}, \omega^{(k)}) \quad \text{and} \quad \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}, \omega^{(k)}) = \sigma_2(\mu^{(k)}, \omega^{(k)}),$$

as well as the uniform Lipschitz continuity of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$, $\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ (i.e., parts **(iii)** and **(iv)** of Lemma 3.2), there exists a region $\bar{\mathcal{B}}(\mu_*, \tilde{\eta}_\mu) \times \bar{\mathcal{B}}(\omega_*, \tilde{\eta}_\omega)$ in which $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)$ is a simple, hence also a positive singular value of $H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$. More precisely, we have

$$(A.2) \quad \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \geq \varepsilon \quad \forall (\mu, \omega) \in \bar{\mathcal{B}}(\mu_*, \tilde{\eta}_\mu) \times \bar{\mathcal{B}}(\omega_*, \tilde{\eta}_\omega)$$

for some $\varepsilon \in (0, \hat{\varepsilon})$, where the constants ε , $\tilde{\eta}_\mu$, $\tilde{\eta}_\omega$ do not depend on $\mu^{(1)}$. However, here it is assumed that $\|\mu^{(1)} - \mu_*\|_2$ is small enough in order to ensure $\|\mu^{(k)} - \mu_*\|_2 \ll \hat{\varepsilon}$.

Table 4.8: The performance of Algorithm 2.1 on Example 4.4 for larger values of n , where we have used `eigopt` for the optimization of the reduced subproblems.

n	n_{iter}	μ_*	$\ H[\mu_*]\ _{\mathcal{H}_\infty}$	runtime in s
1000	4	0.157222	2.08316	3.61
2000	4	0.115748	4.08243	4.68
5000	2	0.113064	10.1718	2.15
10000	2	0.112964	20.3321	1.64
20000	2	0.113009	40.6554	1.37
50000	2	0.113066	101.628	1.70
100000	2	0.113090	203.248	2.69
200000	2	0.113102	406.490	5.04
500000	2	0.113111	1016.22	12.53
1000000	2	0.113113	2032.43	26.11

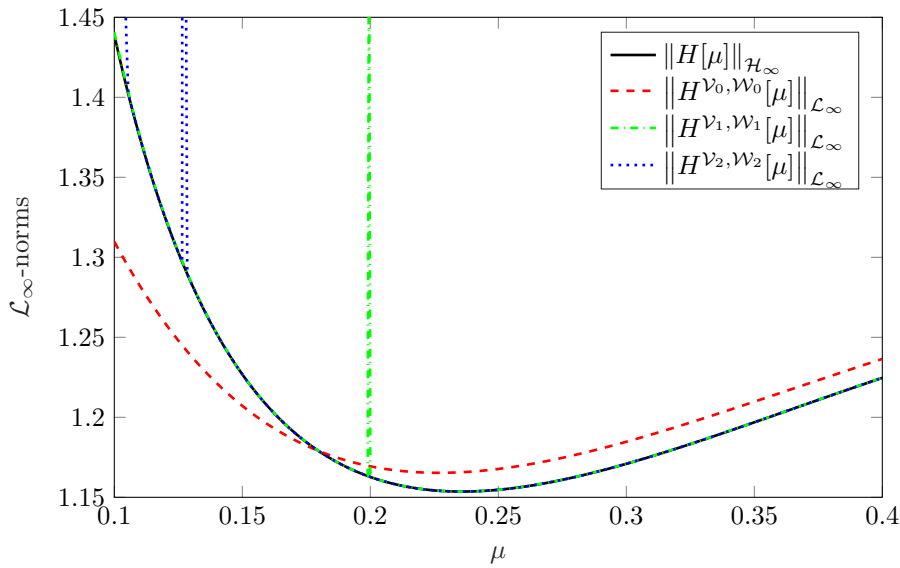


Fig. 4.2: The plots of the full function $\|H[\cdot]\|_{\mathcal{H}_\infty}$, as well as the reduced functions $\|H^{\mathcal{V}_0, \mathcal{W}_0}[\cdot]\|_{\mathcal{L}_\infty}$, $\|H^{\mathcal{V}_1, \mathcal{W}_1}[\cdot]\|_{\mathcal{L}_\infty}$, and $\|H^{\mathcal{V}_2, \mathcal{W}_2}[\cdot]\|_{\mathcal{L}_\infty}$ in the interval $[0.1, 0.4]$ for Example 4.4 with $n = 500$.

Let us now prove that $|\sigma_\omega(\cdot, \cdot)|$ and $|\sigma_{\omega\mu_1}(\cdot, \cdot)|$ are bounded from above uniformly in a neighborhood of (μ_*, ω_*) . Our approach is based on the analytic continuation of the mapping

$$(\mu, s) \mapsto \begin{bmatrix} 0 & H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s) \\ H_*^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s) & 0 \end{bmatrix} =: M^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)$$

into the complex plane for $(\mu, s) \in \mathbb{C}^d \times \mathbb{C}$ near $(\mu_*, i\omega_*)$, where

$$\begin{aligned} H_*^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s) &:= B_*^{W_k}(\mu) D_*^{\mathcal{V}_k, W_k}(\mu, s)^{-1} C_*^{\mathcal{V}_k}(\mu) \quad \text{with} \\ D_*^{\mathcal{V}_k, W_k}(\mu, s) &:= -s E_*^{\mathcal{V}_k, W_k}(\mu) - A_*^{\mathcal{V}_k, W_k}(\mu), \end{aligned}$$

and

$$\begin{aligned}
E_*^{V_k, W_k}(\mu) &:= f_1(\mu)(W_k^* E_1 V_k)^* + \dots + f_{\kappa_E}(\mu)(W_k^* E_{\kappa_E} V_k)^*, \\
A_*^{V_k, W_k}(\mu) &:= g_1(\mu)(W_k^* A_1 V_k)^* + \dots + g_{\kappa_A}(\mu)(W_k^* A_{\kappa_A} V_k)^*, \\
B_*^{W_k}(\mu) &:= h_1(\mu)(W_k^* B_1)^* + \dots + h_{\kappa_B}(\mu)(W_k^* B_{\kappa_B})^*, \\
C_*^{V_k}(\mu) &:= k_1(\mu)(C_1 V_k)^* + \dots + k_{\kappa_C}(\mu)(C_{\kappa_C} V_k)^*.
\end{aligned}$$

Note that $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)$ and $\sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)$ correspond to the largest and second largest eigenvalues of $M^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$ for real ω , as indeed $H_*^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega) = \{H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)\}^*$. These Hermiticity properties are lost, when we replace f_j, g_j, h_j, k_j with their analytic continuations $\hat{f}_j, \hat{g}_j, \hat{h}_j, \hat{k}_j$ or if we choose $s \notin i\mathbb{R} := \{i\omega \mid \omega \in \mathbb{R}\}$. Let us denote the resulting extensions of $H^{\mathcal{V}_k, \mathcal{W}_k}, H_*^{\mathcal{V}_k, \mathcal{W}_k}, M^{\mathcal{V}_k, \mathcal{W}_k}$ with $\hat{H}^{\mathcal{V}_k, \mathcal{W}_k}, \hat{H}_*^{\mathcal{V}_k, \mathcal{W}_k}, \hat{M}^{\mathcal{V}_k, \mathcal{W}_k}$. As the subsequent arguments are for these complex continuations, in the rest of the proof $\overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta) := \{\mu \in \mathbb{C}^d \mid \|\mu - \mu_*\|_2 \leq \eta\}$ and $\overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta) := \{s \in \mathbb{C} \mid |s - i\omega_*| \leq \eta\}$ now denote the balls in the complex Euclidean spaces for a given radius $\eta > 0$. It is straightforward to verify that the uniform Lipschitz continuity of $(\mu, \omega) \mapsto H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$ established in parts **(i)** and **(ii)** of Lemma 3.2 extend to its complex counter-part, in particular, there exist $\gamma, \hat{\eta}_\mu, \hat{\eta}_\omega$ which are independent of $\mu^{(1)}$ such that

$$\begin{aligned}
&\|\hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](\tilde{s}) - \hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2 \\
&\leq \|\hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](\tilde{s}) - \hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\tilde{s})\|_2 + \|\hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\tilde{s}) - \hat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2 \\
&\leq \gamma(\|\tilde{s} - s\|_2 + |\tilde{\omega} - \omega|)
\end{aligned}$$

for all $\tilde{\mu}, \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \hat{\eta}_\mu) \subset \mathbb{C}^d$, and for all $\tilde{s}, s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \hat{\eta}_\omega) \subset \mathbb{C}$. Analogous uniform Lipschitz continuity assertion also holds for $\hat{H}_*^{\mathcal{V}_k, \mathcal{W}_k}$. Consequently, there exist $\gamma, \hat{\eta}_\mu, \hat{\eta}_\omega$ which are independent of $\mu^{(1)}$ such that

$$\begin{aligned}
\text{(A.3)} \quad &\|\hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](\tilde{s}) - \hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2 \leq \gamma(\|\tilde{\mu} - \mu\|_2 + |\tilde{s} - s|) \\
&\forall \tilde{\mu}, \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \hat{\eta}_\mu) \subset \mathbb{C}^d, \quad \forall \tilde{s}, s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \hat{\eta}_\omega) \subset \mathbb{C}.
\end{aligned}$$

Now, for $(\mu, s) \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \hat{\eta}_\mu) \times \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \hat{\eta}_\omega)$, let us consider the eigenvalue $\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, s)$ of $\hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)$ corresponding to the eigenvalue $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega)$ of $M^{\mathcal{V}_k, \mathcal{W}_k}[\mu](i\omega)$, that is, $\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ is obtained by the analytic continuation of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ into the complex plane. This eigenvalue function is no more real-valued, since $\hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)$ is not necessarily a Hermitian matrix. However, by (A.2) and (A.3), as well as Theorem 5.1 in [16, Chapter 4], there exist $\eta_{\mu, m} \leq \min\{\hat{\eta}_\mu, \hat{\eta}_\mu\}$ and $\eta_{\omega, m} \leq \min\{\hat{\eta}_\omega, \hat{\eta}_\omega\}$ such that the eigenvalue $\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu, s)$ remains simple for all $\mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_{\mu, m})$ and all $s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_{\omega, m})$. We remark that $\eta_{\mu, m}$ and $\eta_{\omega, m}$ are independent of $\mathcal{V}_k, \mathcal{W}_k$ and hence are independent of $\mu^{(1)}$. Now let us consider any $\eta_\mu \in (0, \eta_{\mu, m})$ and any $\eta_\omega \in (0, \eta_{\omega, m})$. By the analyticity of $\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)$ in the interior of $\overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_{\mu, m}) \times \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_{\omega, m})$, for a given $\tilde{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)$ and $\tilde{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega/2)$, by Cauchy's integral formula we have

$$\text{(A.4)} \quad \hat{\sigma}_s^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \tilde{s}) = \frac{1}{2\pi i} \oint_{|s-\tilde{s}|=\eta_\omega/2} \frac{\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, s)}{(s-\tilde{s})^2} ds.$$

We claim that the term $\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, s)$ inside the integral in modulus is uniformly bounded from above. To this end, as $|\hat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, s)| \leq \|\hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](s)\|_2$, it suffices to show the uniform boundedness of $\|\hat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\tilde{\mu}](s)\|_2$. Letting $\beta := \sigma_{\min}(D(\mu_*, \omega_*))$ and following the arguments at the beginning of the proof of Lemma 3.2, there exists a neighborhood $\hat{\mathcal{N}} \subset \mathbb{C}^d \times \mathbb{C}$ of $(\mu_*, i\omega_*)$ such that $\sigma_{\min}(D^{\mathcal{V}_k, \mathcal{W}_k}(\mu, s)) \geq \beta/2$ for all $(\mu, s) \in \hat{\mathcal{N}}$. Without loss of generality,

we assume $\widehat{\mathcal{N}} = \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu) \times \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega)$ (as we can choose η_μ and η_ω as small as we wish). Hence,

$$\|\widehat{H}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2 \leq 2 \frac{\|C^{\mathcal{V}_k}(\mu)\|_2 \|B^{\mathcal{W}_k}(\mu)\|_2}{\beta} \leq 2 \frac{M_C M_B}{\beta} \quad \forall \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu), \quad \forall s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega),$$

where $M_C := \max \{\|C(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\}$, $M_B := \max \{\|B(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\}$. In an analogous fashion, the same upper bound also holds uniformly for $\|\widehat{H}_*^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2$ for all $\mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)$ and all $s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega)$, which gives rise to

$$\|\widehat{M}^{\mathcal{V}_k, \mathcal{W}_k}[\mu](s)\|_2 \leq 2 \frac{M_C M_B}{\beta} =: M \quad \forall \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu), \quad \forall s \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega).$$

We deduce from (A.4) that

$$|\widehat{\sigma}_s^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \tilde{s})| \leq \frac{1}{2\pi} \left\{ \max_{|s-\tilde{s}|=\eta_\omega/2} |\widehat{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, s)| \right\} \frac{1}{(\eta_\omega/2)^2} (2\pi\eta_\omega/2) \leq \frac{2M}{\eta_\omega} \quad \forall \tilde{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu), \quad \forall \tilde{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega/2),$$

hence also $|\sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\tilde{\mu}, \tilde{\omega})| \leq 2M/\eta_\omega$ for all $\tilde{\mu} \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$ and all $\tilde{\omega} \in \overline{\mathcal{B}}(\omega_*, \eta_\omega/2)$.

Now let us consider the mixed derivative $\sigma_{s\mu_1}^{\mathcal{V}_k, \mathcal{W}_k}$, specifically for a given $\widehat{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu/2)$ and $\widehat{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega/2)$, we have

$$(A.5) \quad \widehat{\sigma}_{s\mu_1}^{\mathcal{V}_k, \mathcal{W}_k}(\widehat{\mu}, \widehat{s}) = \frac{1}{2\pi i} \oint_{\mathcal{C}} \frac{\widehat{\sigma}_s^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \widehat{s})}{(\mu_1 - \widehat{\mu}_1)^2} d\mu_1,$$

where the contour integral is over $\mathcal{C} := \{\mu \in \mathbb{C}^d \mid |\mu_1 - \widehat{\mu}_1| = \eta_\mu/2, \mu_j = \widehat{\mu}_j, j = 2, \dots, d\}$. Taking the modulus of both sides in (A.5) yields

$$|\widehat{\sigma}_{s\mu_1}^{\mathcal{V}_k, \mathcal{W}_k}(\widehat{\mu}, \widehat{s})| \leq \frac{1}{2\pi} \left\{ \max_{\mu \in \mathcal{C}} |\widehat{\sigma}_s^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \widehat{s})| \right\} \frac{1}{(\eta_\mu/2)^2} (2\pi\eta_\mu/2) \leq \frac{4M}{\eta_\mu \eta_\omega} \quad \forall \widehat{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu/2), \quad \forall \widehat{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(i\omega_*, \eta_\omega/2).$$

The arguments above prove the uniform boundedness of $|\sigma_\omega^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)|$, $|\sigma_{\omega\mu_1}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot, \cdot)|$. The uniform boundedness of all other first three derivatives can be proven similarly. \square

REFERENCES

- [1] A. ALIYEV, P. BENNER, E. MENGI, P. SCHWERDTNER, AND M. VOIGT, *Large-scale computation of \mathcal{L}_∞ -norms by a greedy subspace method*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1496–1516.
- [2] U. BAUR, P. BENNER, A. GREINER, J. G. KORVINK, J. LIENEMANN, AND C. MOOSMANN, *Parameter preserving model order reduction for MEMS applications*, Math. Comput. Model. Dyn. Syst., 17 (2011), pp. 297–317.
- [3] C. BEATTIE AND S. GUGERCIN, *Interpolatory projection methods for structure-preserving model reduction*, Systems Control Lett., 58 (2009), pp. 225–232.
- [4] P. BENNER, V. SIMA, AND M. VOIGT, *\mathcal{L}_∞ -norm computation for continuous-time descriptor systems using structured matrix pencils*, IEEE Trans. Automat. Control, 57 (2012), pp. 233–238.
- [5] L. BREIMAN AND A. CUTLER, *A deterministic algorithm for global optimization*, Math. Program., 58 (1993), pp. 179–199.
- [6] F. E. CURTIS, T. MITCHELL, AND M. L. OVERTON, *A BFGS-SQP method for nonsmooth, nonconvex, constrained optimization and its evaluation using relative minimization profiles*, Optim. Method Softw., 32 (2017), pp. 148–181.

- [7] L. FENG, D. KOZIOL, E. B. RUDNYI, AND J. G. KORVINK, *Parametric model reduction for fast simulation of cyclic voltammograms*, *Sensor Lett.*, 4 (2006), pp. 165–173.
- [8] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, 1985.
- [9] A. JOHANSSON, J. WEI, H. SANDBERG, K. H. JOHANSSON, AND J. CHEN, *Optimization of the H_∞ -norm of dynamic flow networks*, Preprint arXiv:1710.03154, 2017.
- [10] F. KANGAL, K. MEERBERGEN, E. MENGI, AND W. MICHELS, *A subspace method for large scale eigenvalue optimization*, *SIAM J. Matrix Anal. Appl.*, 39 (2018), pp. 48–82.
- [11] A. J. MAYO AND A. C. ANTOULAS, *A framework for the solution of the generalized realization problem*, *Linear Algebra Appl.*, 425 (2007), pp. 634–662.
- [12] E. MENGI, *Large-scale and global maximization of the distance to instability*, *SIAM J. Matrix Anal. Appl.*, 39 (2018), pp. 1776–1809.
- [13] E. MENGI, E. A. YILDIRIM, AND M. KILIÇ, *Numerical optimization of eigenvalues of Hermitian matrix functions*, *SIAM J. Matrix Anal. Appl.*, 35 (2014), pp. 699–724.
- [14] E. B. RUDNYI AND J. G. KORVINK, *Boundary condition independent thermal model*, in *Dimension Reduction of Large-Scale Systems*, vol. 45 of *Lect. Notes Comput. Sci. Eng.*, Springer-Verlag, Berlin, Heidelberg, 2005, ch. 17, pp. 345–348.
- [15] P. SCHWERDTNER AND M. VOIGT, *Computation of the L_∞ -norm using rational interpolation*, *IFAC-PapersOnLine*, 51 (2018), pp. 84–89. Joint 9th IFAC Symposium on Robust Control Design and 2nd IFAC Workshop on Linear Parameter Varying Systems, Florianópolis, Brazil, 2018.
- [16] G. W. STEWART AND J. SUN, *Matrix Perturbation Theory*, Academic Press, 1990.
- [17] A. VARGA AND P. PARILLO, *Fast algorithms for solving H_∞ -norm minimization problems*, in *Proc. 40th IEEE Conference on Decision and Control*, Orlando, FL, USA, 2001, pp. 261–266.
- [18] D. VIZER, G. MERCÈRE, O. PROT, AND E. LAROCHE, *H_∞ -norm-based optimization for the identification of gray-box LTI state-space model parameters*, *Systems Control Lett.*, 92 (2016), pp. 34–41.
- [19] K. ZHOU, J. C. DOYLE, AND K. GLOVER, *Robust and Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1996.