SIAM J. MATRIX ANAL. APPL.
Vol. 41, No. 2, pp. 928–956

A SUBSPACE FRAMEWORK FOR $\mathcal{H}_\infty$-NORM MINIMIZATION*

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

Abstract. We deal with the minimization of the $\mathcal{H}_\infty$-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 interpolatary approaches inspired by our recent work [Aliyev et al., *SIAM J. Matrix Anal. Appl.*, 38 (2017), pp. 1496–1516] for the computation of the $\mathcal{H}_\infty$-norm. In this work, we minimize the $\mathcal{H}_\infty$-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 and nondegeneracy assumptions. The fast convergence of the proposed frameworks in practice is illustrated by several large-scale systems.

Key words. $\mathcal{H}_\infty$-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

DOI. 10.1137/19M125892X

1. Introduction.

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

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

Here, for an open and bounded set $\Omega \subseteq \mathbb{R}^{\mathsf{d}}$, $E, A : \Omega \to \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $B : \Omega \to \mathbb{R}^{\mathsf{n}\times\mathsf{m}}$, $C : \Omega \to \mathbb{R}^{\mathsf{p}\times\mathsf{n}}$ are matrix-valued functions in the parameter-affine representation (cf. [2]) defined by

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

for given matrices $E_1, \ldots, E_{\kappa_E}, A_1, \ldots, A_{\kappa_A} \in \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $B_1, \ldots, B_{\kappa_B} \in \mathbb{R}^{\mathsf{n}\times\mathsf{m}}$, $C_1, \ldots, C_{\kappa_C} \in \mathbb{R}^{\mathsf{p}\times\mathsf{n}}$, and real-analytic functions $f_1, \ldots, f_{\kappa_E}, g_1, \ldots, g_{\kappa_A}, h_1, \ldots,$

*Received by the editors April 30, 2019; accepted for publication (in revised form) March 23, 2020; published electronically June 18, 2020.
https://doi.org/10.1137/19M125892X

†Azerbaijan National Academy of Sciences, Institute of Mathematics and Mechanics. B. Vahabzade 9, 1141, Baku, Azerbaijan and French-Azerbaijani University (UFAZ) Nizami str. 183, Baku, Azerbaijan (naliyev@ku.edu.tr).

‡Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany (benner@mpi-magdeburg.mpg.de).

§Koç University, Department of Mathematics, Rumeli Feneri Yolu 34450, Sarıyer, Istanbul, Turkey (emengi@ku.edu.tr).

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

h_{κ_B}, $k_1, \ldots, k_{\kappa_C} : \Omega \to \mathbb{R}$. The functions $x(\cdot;\mu) : \mathbb{R} \to \mathbb{R}^{\mathsf{n}}$, $u(\cdot;\mu) : \mathbb{R} \to \mathbb{R}^{\mathsf{m}}$, and $y(\cdot;\mu) : \mathbb{R} \to \mathbb{R}^{\mathsf{p}}$ 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 fixed μ, the function $H[\mu](s)$ is real-rational in the indeterminate s, consequently, we use the notation $H[\mu](s) \in \mathbb{R}(s)^{\mathsf{p}\times\mathsf{m}}$. Observe that, since $H[\mu]$ is rational, it is analytic almost everywhere in $\mathbb{C}$.

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

$$\mathcal{L}_\infty^{\mathsf{p}\times\mathsf{m}} := \left\{ H(s) \in \mathbb{R}(s)^{\mathsf{p}\times\mathsf{m}} \;\middle|\; \sup_{\omega\in\mathbb{R}} \|H(\mathrm{i}\omega)\|_2 < \infty \right\},$$
$$\mathcal{H}_\infty^{\mathsf{p}\times\mathsf{m}} := \left\{ H(s) \in \mathbb{R}(s)^{\mathsf{p}\times\mathsf{m}} \;\middle|\; \sup_{\lambda\in\mathbb{C}^+} \|H(\lambda)\|_2 < \infty \right\},$$

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

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

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

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

In this work, we consider the problem of minimizing the $\mathcal{H}_\infty$-norm of $H[\mu]$ over μ that belongs to a compact subset $\underline{\Omega}$ of Ω, but keeping the assumption that $H[\mu] \in \mathcal{H}_\infty^{\mathsf{p}\times\mathsf{m}}$ 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 $\mathcal{H}_\infty$-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 $\mathcal{H}_\infty$-controller that minimizes the influence of a noisy input signal to the regulated output, which corresponds to minimizing the $\mathcal{H}_\infty$-norm of a closed-loop (parameter-dependent) transfer function; see, e.g., [21] 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 $\mathcal{H}_\infty$-norm minimization arise in the optimization of dynamic flow networks [11], parameter identification [20], and model reduction [19].

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., $\mathsf{n} \gg \mathsf{m}, \mathsf{p}$. Here, we present subspace frameworks that are inspired by our previous work [1]. The proposed frameworks converge rapidly 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 $\mathcal{H}_\infty$-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 $\mathcal{H}_\infty$-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 hold not only 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 under some smoothness and nondegeneracy assumptions at the minimizer. The performance of the 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 $\mathcal{H}_\infty$-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 $\mathcal{H}_\infty$-norm computation in [1]. We choose two subspaces $\mathcal{V}$, $\mathcal{W} \subseteq \mathbb{C}^{\mathsf{n}}$ of the same dimension, as well as matrices $V, W \in \mathbb{C}^{\mathsf{n}\times\mathsf{k}}$ whose columns form orthonormal bases for these subspaces, and define the reduced system in terms of the matrix-valued functions

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

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 $\left\|H^{\mathcal{V},\mathcal{W}}[\mu]\right\|_{\mathcal{H}_\infty}$ for appropriate choices of the subspaces $\mathcal{V}$, $\mathcal{W}$.

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

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

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 $H[\mu](\mathrm{i}\omega)$ and $H^{\mathcal{V},\mathcal{W}}[\mu](\mathrm{i}\omega)$, respectively. At every iteration, the basic framework minimizes the $\mathcal{H}_\infty$-norm of a reduced problem for a given pair of subspaces in line 3. Then it first computes an ω such that

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

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

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

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

$\|H[\mu]\|_{\mathcal{H}_\infty} = \sigma(\omega, \mu)$ 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], [4, Theorem 1].

Lemma 2.1 (interpolation properties for the basic algorithm). *The following assertions hold regarding Algorithm* 2.1 *for each* $j = 1, \ldots, k$*:*

(i) *It holds that* $\left\|H\big[\mu^{(j)}\big]\right\|_{\mathcal{H}_\infty} = \sigma\big(\mu^{(j)}, \omega^{(j)}\big) = \sigma^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu^{(j)}, \omega^{(j)}\big)$.

(ii) *It holds that* $\sigma_2\big(\mu^{(j)}, \omega^{(j)}\big) = \sigma_2^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu^{(j)}, \omega^{(j)}\big)$.

(iii) *If the largest singular value* $\sigma\big(\mu^{(j)}, \omega^{(j)}\big)$ *of* $H[\mu^{(j)}](\mathrm{i}\omega^{(j)})$ *is simple, then*

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

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

Note that in part (iv) of the lemma above $\sigma_\omega\big(\mu^{(j)}, \omega^{(j)}\big) = 0$ holds even if $\sigma\big(\mu^{(j)}, \omega^{(j)}\big)$ is not simple, since $\omega^{(j)}$ is a maximizer of $\sigma\big(\mu^{(j)}, \cdot\big)$, and as a result $\sigma\big(\mu^{(j)}, \cdot\big)$ is differentiable at $\omega^{(j)}$ regardless of its multiplicity (see, for instance, the arguments right before Theorem 2.3 in [6]). The equality $\sigma_\omega^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu^{(j)}, \omega^{(j)}\big) = 0$ follows from the interpolation properties between $H[\mu](\mathrm{i}\omega)$, $H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)$ 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 rth column of the $\mathsf{d} \times \mathsf{d}$ 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

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

Input: Matrices $E_1, \ldots, E_{\kappa_E} \in \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $A_1, \ldots, A_{\kappa_A} \in \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $B_1, \ldots, B_{\kappa_B} \in \mathbb{R}^{\mathsf{n}\times\mathsf{m}}$, $C_1, \ldots, C_{\kappa_C} \in \mathbb{R}^{\mathsf{p}\times\mathsf{n}}$ and functions $f_1, \ldots, f_{\kappa_E}$, $g_1, \ldots, g_{\kappa_A}$, $h_1, \ldots, h_{\kappa_B}$, $k_1, \ldots, k_{\kappa_C}$ as in (1.2).
Output: Sequences $\{\mu^{(k)}\}$, $\{\omega^{(k)}\}$.

Choose initial subspace $\mathcal{V}_0, \mathcal{W}_0 \subseteq \mathbb{C}^{\mathsf{n}}$.
for $k = 1, 2, \ldots$ **do**
  $\mu^{(k)} \leftarrow \arg\min_{\mu\in\underline{\Omega}} \left\| H^{\mathcal{V}_{k-1},\mathcal{W}_{k-1}}[\mu] \right\|_{\mathcal{H}_\infty}$.
  $\omega^{(k)} \leftarrow \arg\max_{\omega\in\mathbb{R}\cup\{\infty\}} \sigma\left(\mu^{(k)}, \omega\right)$.
  if $\mathsf{m} = \mathsf{p}$ **then**
    $\widetilde{V}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-1} B\big(\mu^{(k)}\big)$.
    $\widetilde{W}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-*} C\big(\mu^{(k)}\big)^*$.
  else if $\mathsf{m} < \mathsf{p}$ **then**
    $\widetilde{V}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-1} B\big(\mu^{(k)}\big)$.
    $\widetilde{W}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-*} C\big(\mu^{(k)}\big)^* H\big[\mu^{(k)}\big]\big(\mathrm{i}\omega^{(k)}\big)$.
  else
    $\widetilde{V}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-1} B\big(\mu^{(k)}\big) H\big[\mu^{(k)}\big]\big(\mathrm{i}\omega^{(k)}\big)^*$.
    $\widetilde{W}_k \leftarrow D\big(\mu^{(k)}, \mathrm{i}\omega^{(k)}\big)^{-*} C\big(\mu^{(k)}\big)^*$.
  end if
  $\mathcal{V}_k \leftarrow \mathcal{V}_{k-1} \oplus \mathrm{Col}\big(\widetilde{V}_k\big)$ and $\mathcal{W}_k \leftarrow \mathcal{W}_{k-1} \oplus \mathrm{Col}\big(\widetilde{W}_k\big)$.
  if $k \geq 2$ **then**
    $h^{(k)} \leftarrow \left\|\mu^{(k)} - \mu^{(k-1)}\right\|_2$.
    for $r = 1, 2, \ldots, \mathsf{d}$ **do**
      for $q = r, \ldots, \mathsf{d}$ **do**
        $\mu^{(k,rq)} \leftarrow \mu^{(k)} + h^{(k)} e_{rq}$.
        $\omega^{(k,rq)} \leftarrow \arg\max_{\omega\in\mathbb{R}\cup\{\infty\}} \sigma\big(\mu^{(k,rq)}, \omega\big)$.
        if $\mathsf{m} = \mathsf{p}$ **then**
          $\widetilde{V}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-1} B\big(\mu^{(k,rq)}\big)$.
          $\widetilde{W}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-*} C\big(\mu^{(k,rq)}\big)^*$.
        else if $\mathsf{m} < \mathsf{p}$ **then**
          $\widetilde{V}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-1} B\big(\mu^{(k,rq)}\big)$.
          $\widetilde{W}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-*} C\big(\mu^{(k,rq)}\big)^* H\big[\mu^{(k,rq)}\big]\big(\mathrm{i}\omega^{(k,rq)}\big)$.
        else
          $\widetilde{V}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-1} B\big(\mu^{(k,rq)}\big) H\big[\mu^{(k,rq)}\big]\big(\mathrm{i}\omega^{(k,rq)}\big)^*$.
          $\widetilde{W}_k^{(rq)} \leftarrow D\big(\mu^{(k,rq)}, \mathrm{i}\omega^{(k,rq)}\big)^{-*} C\big(\mu^{(k,rq)}\big)^*$.
        end if
        $\mathcal{V}_k \leftarrow \mathcal{V}_k \oplus \mathrm{Col}\big(\widetilde{V}_k^{(rq)}\big)$ and $\mathcal{W}_k \leftarrow \mathcal{W}_k \oplus \mathrm{Col}\big(\widetilde{W}_k^{(rq)}\big)$.
      end for
    end for
  end if
end for

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 (interpolation properties for the extended algorithm). *The iterates* $\{\mu^{(k)}\}$, $\{\omega^{(k)}\}$ *by Algorithm* 2.2 *satisfy the assertions* (i)–(iv) *of Lemma* 2.1 *for each* $j = 1, \ldots, k$. *Additionally, for each* $j = 1, \ldots, k$, $r = 1, \ldots, \mathsf{d}$, *and* $q = r, \ldots, \mathsf{d}$, *we have the following:*

(i) *It holds that* $\left\| H\left[\mu^{(j,rq)}\right]\right\|_{\mathcal{H}_\infty} = \sigma\left(\mu^{(j,rq)}, \omega^{(j,rq)}\right) = \sigma^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu^{(j,rq)}, \omega^{(j,rq)}\right)$.

(ii) *If the largest singular value* $\sigma\left(\mu^{(j,rq)}, \omega^{(j,rq)}\right)$ *of* $H[\mu^{(j,rq)}](\mathrm{i}\omega^{(j,rq)})$ *is simple, then*

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

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

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 in Algorithm 2.2 are done such that the subspaces $\mathcal{V}_k$ and $\mathcal{W}_k$ have the same dimension (otherwise, the subspace of lower dimension must be extended by additional basis vectors to achieve this condition). This is needed in order to obtain a regular reduced matrix pencil $D^{V_k,W_k}\left(\mu^{(k)}, s\right)$ and a well-defined reduced transfer function $H^{\mathcal{V}_k,\mathcal{W}_k}\left[\mu^{(k)}\right](s)$. In practice, a regularization procedure can be performed [13] to obtain a regular reduced matrix pencil. In the above algorithms, we make the silent assumption that the transfer functions $H^{\mathcal{V}_k,\mathcal{W}_k}\left[\mu^{(k)}\right](s)$ are well-defined and in $\mathcal{L}_\infty^{\mathsf{p}\times\mathsf{m}}$ for all k. Note that the reduced dynamical systems associated with the transfer functions $H^{\mathcal{V}_k,\mathcal{W}_k}\left[\mu\right](s)$ are not necessarily asymptotically stable, so the transfer functions are not necessarily in $\mathcal{H}_\infty^{\mathsf{p}\times\mathsf{m}}$. However, for the algorithms, 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 $\mathcal{L}_\infty$-norm, however, have also been recently considered for transfer functions of a more general class of systems [1, 17]. 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. Section 3.2 below introduces functions associated with the reduced systems that interpolate $\|H[\cdot]\|_{\mathcal{H}_\infty}$, and presents their essential differentiability properties. We include a proper derivation of these differentiability properties in Appendix A, as the derivation involves technicalities. Section 3.3 establishes the main superlinear convergence result by exploiting the interpolation properties, in particular by making an analogy with quasi-Newton methods for unconstrained optimization.

Throughout the rest of this text, $\sigma_{\min}(\cdot)$ denotes the smallest singular value of its matrix argument, whereas

$$\overline{\mathcal{B}}\left(\widetilde{\mu}, \eta\right) := \left\{\mu \in \mathbb{R}^{\mathsf{d}} \;\middle|\; \left\|\mu - \widetilde{\mu}\right\|_2 \le \eta\right\} \quad \text{and} \quad \overline{\mathcal{B}}\left(\widetilde{\omega}, \eta\right) := \left\{\omega \in \mathbb{R} \;\middle|\; \left|\omega - \widetilde{\omega}\right| \le \eta\right\}$$

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

3.1. Assumptions and a summary of the main result. It is assumed throughout the section that we have three consecutive iterates of the algorithms $\mu^{(k+1)}, \mu^{(k)}, \mu^{(k-1)}$ at hand, and they are sufficiently close to a local or a global minimizer μ_* of $\|H[\cdot]\|_{\mathcal{H}_\infty}$, where the following smoothness assumptions hold.

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 $H[\mu_*](\mathrm{i}\omega_*)$.

Results are proven uniformly over all subspaces and orthonormal bases for them as long as they satisfy the following nondegeneracy conditions.

Assumption 3.2 (nondegeneracy). For given real numbers $\delta < 0$ and $\beta > 0$, we have

$$\sigma_{\omega\omega}(\mu_*, \omega_*) \le \delta \quad \text{and} \quad \sigma_{\min}\big(D\big(\mu_*, \mathrm{i}\omega_*\big)\big) \ge \beta,$$

and the subspaces $\mathcal{V}_k, \mathcal{W}_k$ as well as the matrices V_k, W_k satisfy

$$\sigma_{\omega\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu_*, \omega_*) \le \delta \quad \text{and} \quad \sigma_{\min}\big(D^{V_k,W_k}\big(\mu_*, \mathrm{i}\omega_*\big)\big) \ge \beta. \tag{3.1}$$

Our main result is a superlinear convergence result, i.e., there exists a constant C such that

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

By a constant, here and throughout the section, we mean that it may depend only on quantities related to the full problem, and is independent of $\mu^{(k+1)}, \mu^{(k)}, \mu^{(k-1)}$. In particular, it is independent of the subspaces $\mathcal{V}_k, \mathcal{W}_k$ and orthonormal bases for them as long as they satisfy Assumption 3.2.

3.2. Locally defined reduced interpolating functions. The analysis that we present makes use of the interpolation properties between the $\mathcal{H}_\infty$-norm function $\|H[\cdot]\|_{\mathcal{H}_\infty}$, and a counterpart associated with a reduced system. An immediate candidate as a reduced counterpart is $\big\|H^{\mathcal{V}_k,\mathcal{W}_k}[\cdot]\big\|_{\mathcal{L}_\infty}$, but this candidate fails to satisfy the interpolation properties, e.g., even the equalities $\big\|H\big[\mu^{(j)}\big]\big\|_{\mathcal{H}_\infty} = \big\|H^{\mathcal{V}_k,\mathcal{W}_k}\big[\mu^{(j)}\big]\big\|_{\mathcal{L}_\infty}$ for $j = 1, \dots, k$ do not necessarily hold as the definitions of $\mathcal{L}_\infty$-norm functions involve global maximizations over all ω. Instead, we introduce the following reduced functions.

Definition 3.3 (reduced interpolating functions). *We call the function*

$$\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\mu) := \sigma^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big) \tag{3.2}$$

the reduced interpolating function *with respect to the subspaces $\mathcal{V}_k$, $\mathcal{W}_k$, where the function $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}$ is implicitly defined locally around μ_* through the equations*

$$\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu^{(k)}\big) = \omega^{(k)} \quad \textit{and} \quad \sigma_{\omega}^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big) = 0 \quad \forall \mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \tag{3.3}$$

for some $\eta_{\mu,0} > 0$.

Well-posedness of a function $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ as in (3.3), and hence of the function $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ as in (3.2), follows from the next result. For a proof, we refer to Appendix A.2.

Proposition 3.4 (local well-posedness of reduced interpolating functions). *Suppose that Assumptions* 3.1 *and* 3.2 *hold.*

(i) *There exist constants* $\widetilde{\eta}_{\mu,0}, \widetilde{\eta}_{\omega,0} > 0$ *such that both* $\sigma(\mu,\omega)$ *and* $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ *are simple, hence, real analytic, for all* μ *and* ω *in the interior of* $\overline{\mathcal{B}}(\mu_*, \widetilde{\eta}_{\mu,0})$ *and* $\overline{\mathcal{B}}(\omega_*, \widetilde{\eta}_{\omega,0})$.

(ii) *There exists a unique continuous function* $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k} : \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \to \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ *for some constants* $\eta_{\mu,0} \in (0, \widetilde{\eta}_{\mu,0})$, $\eta_{\omega,0} \in (0, \widetilde{\eta}_{\omega,0})$ *that satisfies* (3.3).

(iii) *For* $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, *we have* $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)$ *as the unique stationary point of* $\omega \mapsto \sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ *over all* $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$.

Our approach depends on the interpolation of not only the $\mathcal{L}_\infty$-norm functions, but also their gradients, as well as the approximation of their second derivatives. To this end, we next present a result concerning the smoothness properties of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$, whose proof is given in Appendix A.3.

PROPOSITION 3.5 (uniform boundedness of higher-order derivatives). *Suppose that Assumptions* 3.1 *and* 3.2 *hold.*

(i) *Both* $\|H[\cdot]\|_{\mathcal{H}_\infty}$ *and* $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ *are at least three times continuously differentiable in the interior of* $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, *where* $\eta_{\mu,0}$ *is as in Proposition* 3.4.

(ii) *For every* $\widehat{\eta}_{\mu,0} \in (0, \eta_{\mu,0})$ *there exists a constant* $\gamma > 0$ *such that for all* $\mu \in \overline{\mathcal{B}}(\mu_*, \widehat{\eta}_{\mu,0})$, *we have*

(a) $$\left| \frac{\partial^2 \|H[\mu]\|_{\mathcal{H}_\infty}}{\partial \mu_q \partial \mu_r} \right| \le \gamma \quad \textit{and} \quad \left| \widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}(\mu) \right| \le \gamma, \quad q,\, r = 1, \ldots, \mathsf{d},$$

(b) $$\left| \frac{\partial^3 \|H[\mu]\|_{\mathcal{H}_\infty}}{\partial \mu_q \partial \mu_r \partial \mu_\ell} \right| \le \gamma \quad \textit{and} \quad \left| \widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r\mu_\ell}(\mu) \right| \le \gamma, \quad q,\, r,\, \ell = 1, \ldots, \mathsf{d}.$$

Several interpolation properties between $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ are immediate from Lemmas 2.1 and 2.2. At this point, we especially remark

$$\left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty} = \sigma^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu^{(k)}, \omega^{(k)}\right) = \sigma^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu^{(k)}, \boldsymbol{\omega}(\mu^{(k)})\right) = \widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu^{(k)}\right) \quad \text{and} \tag{3.4}$$

$$\nabla \left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty} = \sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu}\left(\mu^{(k)}, \omega^{(k)}\right) = \sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu}\left(\mu^{(k)}, \boldsymbol{\omega}(\mu^{(k)})\right) = \nabla \widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu^{(k)}\right), \tag{3.5}$$

where the first equality in the first line is due to part (i) of Lemma 2.1, while the first and third equalities in the second line are due to parts (iii) and (iv) of Lemma 2.1, respectively.

3.3. Main superlinear convergence result. We consider $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ as a local model constructed for the minimization of $\|H[\cdot]\|_{\mathcal{H}_\infty}$, analogous to the local quadratic models constructed by quasi-Newton methods for unconstrained optimization. Recall that a quadratic model by a quasi-Newton method interpolates the function to be minimized and its gradient at a given estimate for the minimizer. It then redefines the estimate as the minimizer of the quadratic model function.

Even though $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ is not quadratic, it still interpolates $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and its gradient at $\mu^{(k)}$; see (3.4) and (3.5) above. Moreover, as we shall soon see, under mild assumptions, $\mu^{(k+1)}$ is a stationary point of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$. Recalling that the superlinear convergence is achieved for a quasi-Newton method if the Hessian of the quadratic model converges to the Hessian of the objective function at the minimizer in certain directions, we next relate the Hessians of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ and $\|H[\cdot]\|_{\mathcal{H}_\infty}$.

LEMMA 3.6 (proximity of the Hessians). *Suppose that Assumptions* 3.1 *and* 3.2 *hold. Additionally, assume that* $\nabla^2 \left\| H[\mu_*] \right\|_{\mathcal{H}_\infty}$ *is invertible. There exists a constant*

$\zeta > 0$ *such that the following statements hold for Algorithm* 2.1 *when* $\mathsf{d} = 1$ *and for Algorithm* 2.2:

(i) *We have* $\left\| \nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} - \nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right\|_2 \leq \zeta \| \mu^{(k)} - \mu^{(k-1)} \|_2$.

(ii) *Both* $\nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty}$ *and* $\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$ *are invertible.*

(iii) *We have* $\left\| \left[\nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right]^{-1} - \left[\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)}) \right]^{-1} \right\|_2 \leq \zeta \| \mu^{(k)} - \mu^{(k-1)} \|_2$.

Proof. (i) We focus on Algorithm 2.2 only. The proof for Algorithm 2.1 with $\mathsf{d} = 1$ proceeds similarly by defining $h^{(k)} := \mu^{(k-1)} - \mu^{(k)}$. By part (i) of Proposition 3.5, the functions $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ are three times differentiable in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$. Now suppose, without loss of generality, $\mu^{(k)}$ and $\mu^{(k-1)}$ are 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 $r = 1, \ldots, \mathsf{d}$ and $q = r, \ldots, \mathsf{d}$. (Here we note $\omega^{(k)}, \omega^{(k,rq)} \to \omega_*$ as $\mu^{(k)} \to \mu_*$ due to the assumption that ω_* is the unique global maximizer of $\sigma(\mu_*, \cdot)$; see, for instance, the beginning of the proof of Proposition A.3 in the appendix.)

It follows that the functions

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

are continuous and three times differentiable in $(0, 1)$. Additionally, we have

$$\ell(0) = \widetilde{\ell}(0), \quad \ell'(0) = \widetilde{\ell}'(0), \quad \text{and} \quad \ell(1) = \widetilde{\ell}(1). \tag{3.6}$$

The first two of the equalities in (3.6) are immediate from (3.4) and (3.5). To see the last equality in (3.6) at $\alpha = 1$, we observe

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

by Lemma 2.2. Because of the inclusions $\mu^{(k,rq)} \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$ and $\omega^{(k,rq)} \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$, as well as the uniqueness of $\boldsymbol{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k,rq)})$ as the stationary point of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k,rq)}, \cdot)$ over all $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ (see Proposition 3.4), we must have $\boldsymbol{\omega}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k,rq)}) = \omega^{(k,rq)}$. Hence, by employing Lemma 2.2 once again, we deduce

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

Next, by exploiting the interpolation properties in (3.6) in the Taylor expansions

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

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

$$\begin{aligned} &\left[h^{(k)}\right]^2 e_{rq}^{\mathsf{T}} \left[\nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} - \nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu^{(k)}\right)\right] e_{rq} \\ &\qquad = \ell''(0) - \widetilde{\ell}''(0) = \frac{1}{3}\left(\widetilde{\ell}'''(\widetilde{\varepsilon}) - \ell'''(\varepsilon)\right) = \mathcal{O}\left(\left[h^{(k)}\right]^3\right), \end{aligned} \tag{3.7}$$

where the constant hidden in the Landau symbol $\mathcal{O}$ is independent of the subspaces due to part (ii) of Proposition 3.5. By considering particular values of $r = 1, \ldots, \mathsf{d}$ and $q = r, \ldots, \mathsf{d}$ in (3.7), we deduce

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

Once again, the constant hidden in the Landau symbol $\mathcal{O}$ does not depend on the subspaces $\mathcal{V}_k, \mathcal{W}_k$ in the latter equation.

(ii) By the continuity of $\nabla^2 \|H[\cdot]\|_{\mathcal{H}_\infty}$ in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, it is immediate that $\lim_{\mu^{(k)} \rightarrow \mu_*} \nabla^2 \left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty} = \nabla^2 \|H[\mu_*]\|_{\mathcal{H}_\infty}$. Consequently, we suppose, without loss of generality, $\mu^{(k)}$ to be sufficiently close to μ_* so that $\nabla^2 \left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty}$ is invertible. In addition, from part (i), we get

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

implying also the invertibility of $\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu^{(k)}\right)$ under the supposition that $\mu^{(k)}, \mu^{(k-1)}$ are close enough to μ_*.

(iii) This follows from part (i) by employing the adjugate formulas for the inverses of $\nabla^2 \left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty}$ and $\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu^{(k)}\right)$. For details, we refer to [12, Lemma 2.8, part (ii)]. ∎

Now we are ready for the main superlinear convergence result.

THEOREM 3.7 (superlinear convergence to a local minimizer). *Suppose that Assumptions* 3.1 *and* 3.2 *hold. In addition, 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*
- *the function $\sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k+1)}, \cdot)$ has a unique global maximizer, say $\widetilde{\omega}^{(k+1)}$, with $\widetilde{\omega}^{(k+1)} \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$.*

For both Algorithm 2.1 *when* $\mathsf{d} = 1$ *and Algorithm* 2.2*, there exists a constant $C > 0$ such that*

$$\frac{\left\| \mu^{(k+1)} - \mu_* \right\|_2}{\left\| \mu^{(k)} - \mu_* \right\|_2 \max\left\{ \left\| \mu^{(k)} - \mu_* \right\|_2, \left\| \mu^{(k-1)} - \mu_* \right\|_2 \right\}} \leq C. \tag{3.8}$$

Proof. By part (i) of Proposition 3.5, both $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\cdot)$ defined by (3.2) are twice Lipschitz continuously differentiable in the interior of the ball $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$. Additionally, suppose, without loss of generality, $\mu^{(k)}$, $\mu^{(k-1)}$ lie in the interior of $\overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, and $\mu^{(k)}$, $\mu^{(k-1)}$ are close enough to μ_* so that

- $\overline{\mathcal{B}}\left(\mu^{(k)}, h^{(k)}\right) \subset \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$, where $h^{(k)} := \left\| \mu^{(k)} - \mu^{(k-1)} \right\|_2$, as well as
- $\nabla^2 \left\| H\left[\mu^{(k)}\right] \right\|_{\mathcal{H}_\infty}$ and $\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu^{(k)}\right)$ are invertible (see part (ii) of Lemma 3.6).

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

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

which implies

(3.9)
$$0 = \left(\nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right)^{-1} \nabla \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty} + \big(\mu_* - \mu^{(k)}\big) + \left(\nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right)^{-1}$$
$$\times \int_0^1 \left(\nabla^2 \big\| H\big[\mu^{(k)} + t\big(\mu_* - \mu^{(k)}\big)\big] \big\|_{\mathcal{H}_\infty} - \nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right) \big(\mu_* - \mu^{(k)}\big) \mathrm{d}t.$$

Now, by exploiting $\nabla \| H(\mu^{(k)}) \|_{\mathcal{H}_\infty} = \nabla \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}(\mu^{(k)})$ (due to (3.5)), (3.9) can be rearranged as

(3.10)
$$\begin{aligned} 0 = {} & \left(\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k)}\big)\right)^{-1} \nabla \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k)}\big) + \big(\mu_* - \mu^{(k)}\big) \\ & + \left[\left(\nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right)^{-1} - \left(\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k)}\big)\right)^{-1}\right] \nabla \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty} \\ & + \left(\nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right)^{-1} \\ & \quad \times \int_0^1 \left(\nabla^2 \big\| H\big[\mu^{(k)} + t\big(\mu_* - \mu^{(k)}\big)\big] \big\|_{\mathcal{H}_\infty} - \nabla^2 \big\| H\big[\mu^{(k)}\big] \big\|_{\mathcal{H}_\infty}\right) \big(\mu_* - \mu^{(k)}\big) \mathrm{d}t. \end{aligned}$$

Throughout the rest of the proof, by manipulating (3.10), we bound $\|\mu^{(k+1)} - \mu_*\|_2$ from above in terms of $\|\mu^{(k)} - \mu_*\|_2$ and $\|\mu^{(k-1)} - \mu_*\|_2$.

We first focus on the first term on the right-hand side of (3.10). Since $\widetilde{\omega}^{(k+1)} \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ is assumed to be the unique global maximizer of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k+1)}, \cdot\big)$, we obtain $\boldsymbol{\omega}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k+1)}\big) = \widetilde{\omega}^{(k+1)}$ by Proposition 3.4, in particular, by the uniqueness of $\boldsymbol{\omega}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k+1)}\big)$ as the stationary point of $\sigma^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k+1)}, \cdot\big)$ over $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$. It follows that

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

where we use the fact that $\mu^{(k+1)}$ is a minimizer of $\big\| H^{\mathcal{V}_k, \mathcal{W}_k}[\cdot] \big\|_{\mathcal{H}_\infty}$ for the last equality. Moreover, recalling Proposition 3.5(i), a Taylor expansion yields

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

which in turn implies
(3.11)
$$\left(\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k)}\big)\right)^{-1} \nabla \widetilde{\sigma}^{\mathcal{V}_k, \mathcal{W}_k}\big(\mu^{(k)}\big) = \big(\mu^{(k)} - \mu^{(k+1)}\big) + \mathcal{O}\left(\big\|\mu^{(k+1)} - \mu^{(k)}\big\|_2^2\right).$$

As for the second to last terms on the right-hand side of (3.10), by another Taylor expansion and again Proposition 3.5(i),

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

Therefore, by using part (iii) of Lemma 3.6 and part (ii) of Proposition 3.5, we see that

$$
\begin{aligned}
&\left\| \left[\left(\nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right)^{-1} - \left(\nabla^2 \widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\mu^{(k)}) \right)^{-1} \right] \cdot \nabla \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right\|_2 \\
&\quad \le \zeta \| \mu^{(k)} - \mu^{(k-1)} \|_2 \cdot \left\| \nabla \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right\|_2 = \mathcal{O}\left(\| \mu^{(k)} - \mu^{(k-1)} \|_2 \cdot \| \mu^{(k)} - \mu_* \|_2 \right).
\end{aligned}
\tag{3.12}
$$

Finally, for the last term on the right-hand side of (3.10), we exploit the Lipschitz continuity of $\nabla^2 \|H[\cdot]\|_{\mathcal{H}_\infty}$ near μ_* to deduce

$$
\begin{aligned}
&\Bigg\| \left(\nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right)^{-1} \\
&\quad \times \int_0^1 \left(\nabla^2 \| H[\mu^{(k)} + t(\mu_* - \mu^{(k)})] \|_{\mathcal{H}_\infty} - \nabla^2 \| H[\mu^{(k)}] \|_{\mathcal{H}_\infty} \right) (\mu_* - \mu^{(k)}) \mathrm{d}t \Bigg\|_2 \\
&\qquad = \mathcal{O}\left(\| \mu^{(k)} - \mu_* \|_2^2 \right).
\end{aligned}
\tag{3.13}
$$

Combining (3.10) with (3.11), (3.12), (3.13), and noting

$$
\| \mu^{(k)} - \mu^{(k-1)} \|_2 \le 2 \max \left\{ \| \mu^{(k)} - \mu_* \|_2, \| \mu^{(k-1)} - \mu_* \|_2 \right\},
$$

lead us to

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

for some constants c_1, c_2 from which (3.8) is immediate. ∎

Remark 3.8. 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 $\overline{\mathcal{B}}(\mu_*, \eta)$ (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 $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ in another ball $\overline{\mathcal{B}}(\mu_*, \widetilde{\eta}) \subseteq \overline{\mathcal{B}}(\mu_*, \eta)$, due to the interpolation properties (specifically due to (3.5)), and uniform upper bounds on the derivatives of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ (see, in particular, part (ii) of Proposition 3.5). 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 $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ at $\mu^{(k+1)}$. This means that the rate of convergence analysis above, in particular, the proof of Theorem 3.7, 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) website.[1] Our numerical experiments have been performed on a machine with 4 Intel® Core™ i5-4590 CPUs with 3.30 GHz each and 16 GB RAM using Linux version 4.4.132-53-default and MATLAB version 9.4.0.813654 (R2018a).

4.1. Implementation details and test setup. At each iteration of Algorithm 2.1, the $\mathcal{L}_\infty$-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 nonconvex optimization problem:

- `eigopt`, a MATLAB implementation of the algorithm in [15], which is an adaptation of the algorithm in [7] for eigenvalue optimization. This MATLAB package creates a lower and an upper bound for the optimal value of a given eigenvalue function by employing piecewise 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` [8], 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 $\mathcal{L}_\infty$-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

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

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 $\mathcal{L}_\infty$-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 [5] to compute the $\mathcal{L}_\infty$-norm of small-scale reduced

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

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 $\mathcal{L}_\infty$-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 $\omega_{\max}$ 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 [16]). Our first example is a thermal conduction model in 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 chip producers the capability 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,

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

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 ith 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 $\left[1, 10^4\right] \times \left[1, 10^4\right] \times \left[1, 10^4\right]$.

We report on the results of Algorithm 2.1 applied to 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 [3]). 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 only on 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.

TABLE 4.2
Numerical results for the `anemometer_1p` example.

Setup	n_{iter}	v_*	$\|H[v_*]\|_{\mathcal{H}_\infty}$	Time in s
`eigopt`	6	0	1.32274e–2	32.68
`GRANSO`	6	-6.5276e–14	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
`eigopt`	4	(0.0000, 2.0000, 1.0000e–1)	1.64723e–3	766.06
`GRANSO`	3	(0.0000, 2.0000, 8.3855e–1)	1.64723e–3	40.93

Example 4.3 (scanning electrochemical microscopy (`SECM`); see [9]). SECM 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 Fick's second law which leads to two partial diffusion equations with appropriate boundary conditions. A spatial discretization together with a boundary control then leads to a linear-time invariant system whose transfer function is

$$H[h_1, h_2](s) = C(sE - (h_1A_1 + h_2A_2 - A_3))^{-1}B,$$

where E, A_1, A_2, $A_3 \in \mathbb{R}^{16912\times16912}$, $B \in \mathbb{R}^{16912\times1}$, $C \in \mathbb{R}^{5\times16912}$, and h_1, h_2 are the parameters of the problem. The experiment is performed in the box $\left[1, e^2\right] \times \left[1, e^2\right]$.

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

In all examples, we observe superliner convergence in the final iterations. Specifically, for the `SECM` example, we report the errors when `GRANSO` is used for the subproblems 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 $\mathcal{H}_\infty$-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

TABLE 4.4
Numerical results for the SECM example.

Setup	n_{iter}	$(h_{1,*}, h_{2,*})$	$\Vert H[h_{1,*}, h_{2,*}]\Vert_{\mathcal{H}_\infty}$	Time in s
eigopt	5	(1.0000, 4.1944)	1.85588	180.01
GRANSO	5	(1.0000, 4.2882)	1.85583	20.51

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 $\mathcal{L}_\infty$-norms are listed for the SECM example by using GRANSO for optimization. Here, the short-hands $f^{(k)} := \Vert H^{\mathcal{V}_k,\mathcal{W}_k}[\mu^{(k+1)}]\Vert_{\mathcal{L}_\infty}$ and $f_* := \Vert H[\mu_*]\Vert_{\mathcal{H}_\infty}$ are used.

k	$\mu^{(k+1)}$	$\Vert\mu^{(k+1)} - \mu_*\Vert_2$	$\lvert f^{(k)} - f_*\rvert$
0	(1.000000, 1.208804)	3.08	2.61e–1
1	(1.657869, 7.389056)	3.17	2.72e–4
2	(1.476130, 6.352188)	2.12	1.51e–4
3	(1.000000, 4.288178)	1.55e–9	1.24e–12
4	(1.000000, 4.288178)	< 1e–12	< 1e–12

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 optimality issue with GRANSO, the subspace framework equipped with GRANSO rarely 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 $\mathcal{H}_\infty$-norm minimization providing an optimality certificate available 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 $\mathcal{H}_\infty$-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 $\mathcal{H}_\infty$-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 $\mathsf{n} = 2\mathsf{q}$ with transfer functions of the form

$$H[\mu](s) = C(sI_{\mathsf{n}} - \mu A_1 - A_0)^{-1}B, \tag{4.2}$$

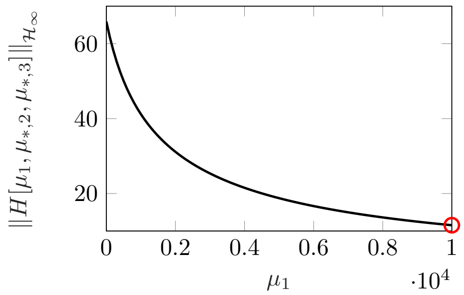
where the matrices A_1, $A_0 \in \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $B \in \mathbb{R}^{\mathsf{n}\times 1}$, $C \in \mathbb{R}^{1\times\mathsf{n}}$ are given by

$$A_1 = \begin{bmatrix} A_{1,1} & & \\ & \ddots & \\ & & A_{1,\mathsf{q}} \end{bmatrix},\ A_0 = \begin{bmatrix} A_{0,1} & & \\ & \ddots & \\ & & A_{0,\mathsf{q}} \end{bmatrix},\ B = \begin{bmatrix} B_1 \\ \vdots \\ B_{\mathsf{q}} \end{bmatrix},\ C = \begin{bmatrix} C_1 & \dots & C_{\mathsf{q}} \end{bmatrix}$$

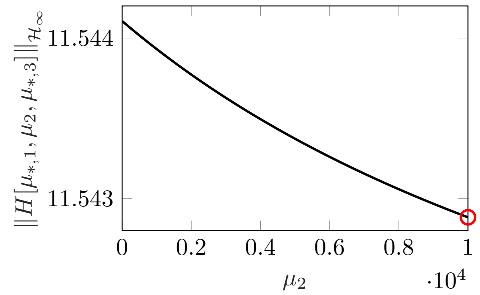
with

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

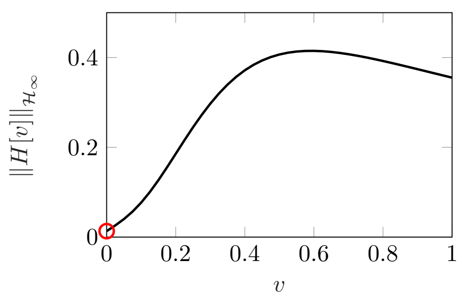
The numbers a_i and b_i are chosen equidistantly in the intervals $[-10^3, -10]$ and $[10, 10^3]$, respectively. The parameter μ is constrained to lie in the interval $[0.02, 1]$.



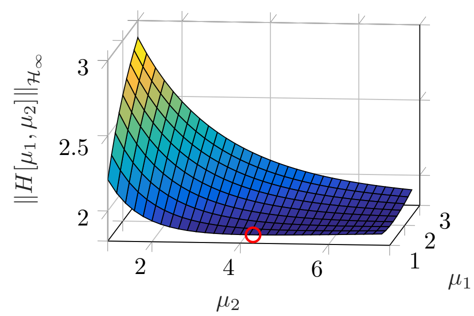
(a) Graph of $\mu_1 \mapsto \|H[\mu_1, \mu_{*,2}, \mu_{*,3}]\|_{\mathcal{H}_\infty}$ for the `T2DAL_BCI` example.



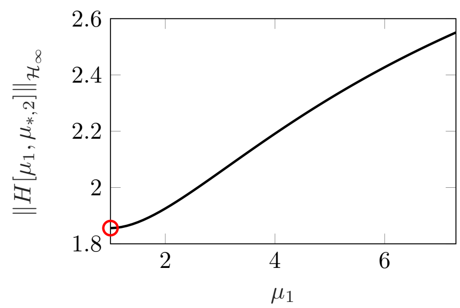
(b) Graph of $\mu_2 \mapsto \|H[\mu_{*,1}, \mu_2, \mu_{*,3}]\|_{\mathcal{H}_\infty}$ for the `T2DAL_BCI` example.



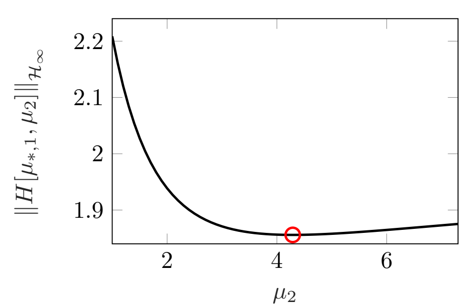
(c) Graph of $v \mapsto \|H[v]\|_{\mathcal{H}_\infty}$ for the `anemometer_1p` example.



(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 $\mathcal{H}_\infty$-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 jth component of μ_* for $j = 1$, 2, or* 3.

We perform our experiments on this synthetic example for several values of $\mathtt{n}$ varying in $10^2, \ldots, 10^6$. For smaller values of $\mathtt{n}$, a comparison of Algorithm 2.1 and the MATLAB package `eigopt` (for the unreduced problems) is provided in Table 4.6. This table indicates that, with or without reduction, we retrieve exactly the same optimal $\mathcal{H}_\infty$-norm values up to the prescribed tolerance $\varepsilon_2 = 10^{-6}$, yet the proposed subspace framework leads to speedups on the order of 10^3; indeed, the ratios of the runtimes increase quickly with respect to $\mathtt{n}$.

Larger examples are considered in Table 4.7, but only using the proposed subspace framework. It does not seem possible to solve these larger $\mathcal{H}_\infty$-norm minimization

TABLE 4.6

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 μ_* by Algorithm 2.1 and* `eigopt`*, the corresponding minimal $\mathcal{H}_\infty$-norms, as well as the runtimes in seconds (s). The optimal $\mathcal{H}_\infty$-norm values returned by Algorithm 2.1 are the same as those returned by* `eigopt` *at least up to six decimal digits. Note that we have set $\gamma = -1000$ in* `eigopt` *for the unreduced problems—otherwise, the runtimes would be even higher.*

		μ_*		$\|H[\mu_*]\|_{\mathcal{H}_\infty}$		Runtime in s	
n	n_{iter}	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

TABLE 4.7

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

problems in a reasonable time without employing reductions. Even the examples of order 10^6 can be solved very fast. All examples can be solved with just two to four iterations. 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 provided γ is chosen sufficiently small. We observe in practice that when the reduced problems are solved by an algorithm that converges only to a local minimizer, convergence to $\mu = 1$, a locally (but not globally) 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 small-scale computation of the $\mathcal{L}_\infty$-norm in the reduced problems 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 always converge.

Finally, the progress of the subspace framework is displayed in Figure 4.2 on this synthetic example for $\mathsf{n} = 500$. After one subspace iteration, the $\mathcal{L}_\infty$-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 $\mathcal{L}_\infty$-norm functions for the reduced and original problems around the minimizer, except for a thin peak that occurs in the reduced problem. The progress in the iterates is further summarized in Table 4.8.

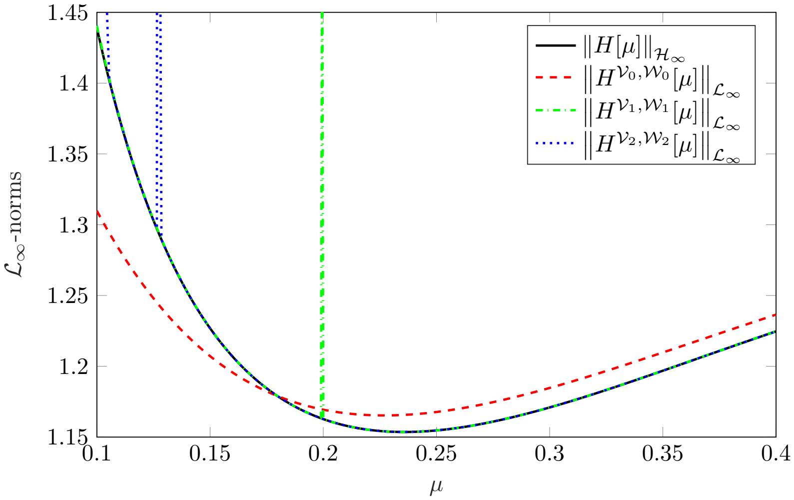


FIG. 4.2. *The plots of the full function $\|H[\cdot]\|_{\mathcal{H}_\infty}$, as well as the reduced functions $\|H^{\mathcal{V}_0,\mathcal{W}_0}[\cdot]\|_{\mathcal{L}_\infty}$, $\|H^{\mathcal{V}_1,\mathcal{W}_1}[\cdot]\|_{\mathcal{L}_\infty}$, and $\|H^{\mathcal{V}_2,\mathcal{W}_2}[\cdot]\|_{\mathcal{L}_\infty}$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

TABLE 4.8

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

k	$\mu^{(k+1)}$	$\lvert\mu^{(k+1)} - \mu_*\rvert_2$	$\lvert f^{(k)} - f_*\rvert$
0	0.226862	8.86e–1	4.80e–4
1	0.235710	7.85e–6	3.12e–10
2	0.235718	$<$ 1e–12	$<$ 1e–12

5. Concluding remarks. In this work, we have developed new subspace restriction techniques to minimize the $\mathcal{H}_\infty$-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, and 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 $\mathcal{H}_\infty$-controllers for large-scale systems partly feasible. A fully feasible method to design optimal $\mathcal{H}_\infty$-controllers for large-scale systems should also take stability considerations into account. We intend to address stability issues in the future.

Appendix A. Derivation and analyses of reduced interpolating functions. The purpose of this section is twofold. First, in section A.2, we show the well-posedness of the reduced interpolating function $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ defined by (3.2) and (3.3). Then we establish uniform upper bounds on the higher-order derivatives of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ independent of the subspaces in section A.3.

Our approaches in sections A.2 and A.3 rely on uniform bounds on the derivatives of the full and reduced singular value functions, which are proven next in section A.1.

A.1. Uniform boundedness of the derivatives of the singular value functions. The main result of this subsection (Lemma A.2) concerns the existence of a positive real number that bounds the second and third derivatives of $\sigma(\cdot,\cdot)$ and $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ in absolute value from above for all choices of subspaces $\mathcal{V}_k, \mathcal{W}_k$.

We start with an auxiliary result that establishes the uniform Lipschitz continuity of $H^{\mathcal{V}_k,\mathcal{W}_k}[\cdot](\mathrm{i}\cdot)$, $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$, and $\sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ independent of the subspaces.

LEMMA A.1 (uniform Lipschitz continuity). *Suppose that Assumption* 3.2 *holds. Then there exist constants* η_μ, η_ω, γ *such that*

(i) $$\left\|H^{\mathcal{V}_k,\mathcal{W}_k}[\widetilde{\mu}](\mathrm{i}\omega) - H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)\right\|_2 \le \gamma\left\|\widetilde{\mu}-\mu\right\|_2 \quad \forall\, \widetilde{\mu},\, \mu \in \overline{\mathcal{B}}(\mu_*,\eta_\mu)\ \forall\, \omega\in\overline{\mathcal{B}}(\omega_*,\eta_\omega);$$

(ii) $$\left\|H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\widetilde{\omega}) - H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)\right\|_2 \le \gamma\left|\widetilde{\omega}-\omega\right| \quad \forall\, \mu \in \overline{\mathcal{B}}(\mu_*,\eta_\mu)\ \forall\, \widetilde{\omega},\, \omega\in\overline{\mathcal{B}}(\omega_*,\eta_\omega);$$

(iii) $$\begin{aligned}&\left|\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\widetilde{\mu},\omega) - \sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)\right| \le \gamma\left\|\widetilde{\mu}-\mu\right\|_2,\\ &\left|\sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\widetilde{\mu},\omega) - \sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)\right| \le \gamma\left\|\widetilde{\mu}-\mu\right\|_2\\ &\qquad \forall\, \widetilde{\mu},\, \mu \in \overline{\mathcal{B}}(\mu_*,\eta_\mu)\ \forall\, \omega\in\overline{\mathcal{B}}(\omega_*,\eta_\omega);\end{aligned}$$

(iv) $$\begin{aligned}&\left|\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\widetilde{\omega}) - \sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)\right| \le \gamma\left|\widetilde{\omega}-\omega\right|,\\ &\left|\sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\widetilde{\omega}) - \sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)\right| \le \gamma\left|\widetilde{\omega}-\omega\right|\\ &\qquad \forall\, \mu \in \overline{\mathcal{B}}(\mu_*,\eta_\mu)\ \forall\, \widetilde{\omega},\, \omega\in\overline{\mathcal{B}}(\omega_*,\eta_\omega).\end{aligned}$$

Proof. Letting V_k, W_k denote matrices with columns forming orthonormal bases for $\mathcal{V}_k$, $\mathcal{W}_k$, by Weyl's theorem [10, Theorem 4.3.1] for every $\mu\in\underline{\Omega}$ and $\omega\in\mathbb{R}$ we have

$$\begin{aligned}\left|\sigma_{\min}\big(D^{V_k,W_k}(\mu,\mathrm{i}\omega)\big) - \sigma_{\min}\big(D^{V_k,W_k}(\mu_*,\mathrm{i}\omega_*)\big)\right| &\le \left\|D^{V_k,W_k}(\mu,\mathrm{i}\omega) - D^{V_k,W_k}(\mu_*,\mathrm{i}\omega_*)\right\|_2\\ &= \|W_k^*(D(\mu,\mathrm{i}\omega)-D(\mu_*,\mathrm{i}\omega_*))V_k\|_2\\ &\le \|D(\mu,\mathrm{i}\omega)-D(\mu_*,\mathrm{i}\omega_*)\|_2\\ &\le \nu(\|\mu-\mu_*\|_2 + |\omega-\omega_*|)\end{aligned}$$

for some constant $\nu>0$, where the last inequality is due to the fact that $(\mu,\omega)\mapsto D(\mu,\mathrm{i}\omega)$ is continuously differentiable in a neighborhood of $\underline{\Omega}\times\mathbb{R}$. This uniform Lipschitz continuity property of $(\mu,\omega)\mapsto\sigma_{\min}\big(D^{V_k,W_k}(\mu,\mathrm{i}\omega)\big)$, combined with $\sigma_{\min}\big(D^{V_k,W_k}(\mu_*,\mathrm{i}\omega_*)\big)\ge\beta$, implies the existence of constants $\eta_\mu>0$, $\eta_\omega>0$ such that

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

It follows that $(\mu,\omega)\mapsto H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)$ 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

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

for $j = 1, \ldots, \mathsf{d}$. Setting

$$M'_{D,j} := \max\left\{ \left\| \frac{\partial D(\mu, \mathrm{i}\omega)}{\partial \mu_j} \right\|_2 \;\middle|\; \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega) \right\},$$

$$M'_{C,j} := \max\left\{ \left\| \frac{\partial C(\mu)}{\partial \mu_j} \right\|_2 \;\middle|\; \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\},\ M_C := \max\left\{ \|C(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\},$$

$$M'_{B,j} := \max\left\{ \left\| \frac{\partial B(\mu)}{\partial \mu_j} \right\|_2 \;\middle|\; \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\},\ M_B := \max\left\{ \|B(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu) \right\},$$

and exploiting

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

as well as $\sigma_{\min}\big(D^{V_k, W_k}(\mu, \mathrm{i}\omega)\big) \geq \beta/2$, we deduce from (A.1) that

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

for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$, $\omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$ and $j = 1, \ldots, \mathsf{d}$. This implies

$$\left| \left[\partial H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega)/\partial \mu_j \right]_{k\ell} \right| \leq M_j$$

for $k = 1, \ldots, \mathsf{p}$, $\ell = 1, \ldots, \mathsf{m}$. Setting $M := \max\{M_j \mid j = 1, \ldots, \mathsf{d}\}$, for every $\widetilde{\mu}, \mu \in \overline{\mathcal{B}}(\mu_*, \eta_\mu), \omega \in \overline{\mathcal{B}}(\omega_*, \eta_\omega)$, by the mean value theorem we obtain

$$\begin{aligned} \left| \left[H^{\mathcal{V}_k, \mathcal{W}_k}\left[\widetilde{\mu}\right](\mathrm{i}\omega) \right]_{k\ell} - \left[H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right]_{k\ell} \right| &\leq \left| \nabla_\mu [H^{\mathcal{V}_k, \mathcal{W}_k}[\widehat{\mu}](\mathrm{i}\omega)]_{k\ell}^{\mathsf{T}} (\widetilde{\mu} - \mu) \right| \\ &\leq \sum_{j=1}^{\mathsf{d}} M_j \left| \widetilde{\mu}_j - \mu_j \right| \leq \mathsf{d} M \left\| \widetilde{\mu} - \mu \right\|_2 \end{aligned}$$

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

$$\left\| H^{\mathcal{V}_k, \mathcal{W}_k}\left[\widetilde{\mu}\right](\mathrm{i}\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right\|_2 \leq \sqrt{\mathsf{pm}}\mathsf{d} M \left\| \widetilde{\mu} - \mu \right\|_2,$$

as desired.

(ii) A similar proof as in part (i) applies but now by differentiating the function $(\mu, \omega) \mapsto H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega)$ with respect to ω instead of μ_j.

(iii) By Weyl's theorem [10, Theorem 4.3.1] and part (i) we have

$$\left| \sigma^{\mathcal{V}_k, \mathcal{W}_k}\left(\widetilde{\mu}, \omega\right) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \right| \leq \left\| H^{\mathcal{V}_k, \mathcal{W}_k}\left[\widetilde{\mu}\right](\mathrm{i}\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right\|_2 \leq \gamma \left\| \widetilde{\mu} - \mu \right\|_2,$$
$$\left| \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}\left(\widetilde{\mu}, \omega\right) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \right| \leq \left\| H^{\mathcal{V}_k, \mathcal{W}_k}\left[\widetilde{\mu}\right](\mathrm{i}\omega) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right\|_2 \leq \gamma \left\| \widetilde{\mu} - \mu \right\|_2$$

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

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

$$\left| \sigma^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu, \widetilde{\omega}\right) - \sigma^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \right| \leq \left\| H^{\mathcal{V}_k, \mathcal{W}_k}[\mu]\left(\mathrm{i}\widetilde{\omega}\right) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right\|_2 \leq \gamma \left| \widetilde{\omega} - \omega \right|,$$
$$\left| \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}\left(\mu, \widetilde{\omega}\right) - \sigma_2^{\mathcal{V}_k, \mathcal{W}_k}(\mu, \omega) \right| \leq \left\| H^{\mathcal{V}_k, \mathcal{W}_k}[\mu]\left(\mathrm{i}\widetilde{\omega}\right) - H^{\mathcal{V}_k, \mathcal{W}_k}[\mu](\mathrm{i}\omega) \right\|_2 \leq \gamma \left| \widetilde{\omega} - \omega \right|$$

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

Now we are ready to assert uniform upper bounds on the derivatives of the largest singular value function for the reduced problem in the next lemma. Its proof is inspired from [12, Proposition 2.9].

LEMMA A.2. *Suppose that Assumptions* 3.1 *and* 3.2 *hold.*

(i) *The singular value functions* $\sigma(\mu,\omega)$ *and* $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ *both are simple, hence, real analytic, for all* μ *and* ω *in the interior of* $\overline{\mathcal{B}}(\mu_*,\widetilde{\eta}_\mu)$ *and the interior of* $\overline{\mathcal{B}}(\omega_*,\widetilde{\eta}_\omega)$*, respectively, for some constants* $\widetilde{\eta}_\mu,\widetilde{\eta}_\omega > 0$.

(ii) *There exist constants* $U>0$, $\eta_\mu\in(0,\widetilde{\eta}_\mu)$, $\eta_\omega\in(0,\widetilde{\eta}_\omega)$ *such that*

$$\left|\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\chi_1}(\mu,\omega)\right|\le U,\quad \left|\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\chi_1\chi_2}(\mu,\omega)\right|\le U,\quad \left|\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\chi_1\chi_2\chi_3}(\mu,\omega)\right|\le U$$
$$\forall\,\mu\in\overline{\mathcal{B}}(\mu_*,\eta_\mu)\ \forall\,\omega\in\overline{\mathcal{B}}(\omega_*,\eta_\omega)$$

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

Proof. (i) By the continuity of $(\mu,\omega)\mapsto\sigma_{\min}(D(\mu,\mathrm{i}\omega))$, there exists a neighborhood $\widetilde{\mathcal{N}}$ of (μ_*,ω_*) such that $\sigma_{\min}(D(\mu,\mathrm{i}\omega))\ge\beta/2$ for all $(\mu,\omega)\in\widetilde{\mathcal{N}}$. Consequently, the mapping $(\mu,\omega)\mapsto H[\mu](\mathrm{i}\omega)$ is continuously differentiable and $\sigma(\cdot,\cdot)$, $\sigma_2(\cdot,\cdot)$ are continuous in $\widetilde{\mathcal{N}}$. The continuity of $\sigma(\cdot,\cdot)$, $\sigma_2(\cdot,\cdot)$, combined with Assumption 3.1 (in particular the assumption that $\sigma(\mu_*,\omega_*)$ is simple), implies that $\sigma(\mu,\omega)$ remains a simple singular value of $H[\mu](\mathrm{i}\omega)$, hence, it is bounded away from zero, in a neighborhood $\mathcal{N}\subseteq\widetilde{\mathcal{N}}$ of (μ_*,ω_*). Formally,

$$\sigma(\mu,\omega)-\sigma_2(\mu,\omega)\ge\widehat{\varepsilon}\quad\forall(\mu,\omega)\in\mathcal{N}\tag{A.2}$$

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

Moreover, by employing the interpolation properties

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

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 A.1), there exists a region $\overline{\mathcal{B}}(\mu_*,\widetilde{\eta}_\mu)\times\overline{\mathcal{B}}(\omega_*,\widetilde{\eta}_\omega)\subseteq\mathcal{N}$ in which $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ is simple, hence, also a positive singular value of $H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)$. More precisely, we have

$$\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)-\sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)\ge\varepsilon\quad\forall(\mu,\omega)\in\overline{\mathcal{B}}(\mu_*,\widetilde{\eta}_\mu)\times\overline{\mathcal{B}}(\omega_*,\widetilde{\eta}_\omega)\tag{A.3}$$

for some constants $\varepsilon\in(0,\widehat{\varepsilon})$, $\widetilde{\eta}_\mu>0$, $\widetilde{\eta}_\omega>0$. Note that here $\|\mu^{(k)}-\mu_*\|_2\ll\widehat{\varepsilon}$ is assumed. It follows from the simplicity of $\sigma(\cdot,\cdot)$ and $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ in $\overline{\mathcal{B}}(\mu_*,\widetilde{\eta}_\mu)\times\overline{\mathcal{B}}(\omega_*,\widetilde{\eta}_\omega)$ that these singular value functions are real analytic in the interior of this region.

(ii) Let us 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 mapping

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

for $(\mu,s)\in\mathbb{C}^{\mathsf{d}}\times\mathbb{C}$ near $(\mu_*,\mathrm{i}\omega_*)$, where

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

and

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

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 $M^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)$ for real μ and real ω, as indeed $H_*^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega) = \left\{H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)\right\}^*$. These Hermiticity properties are lost when we replace f_j, g_j, h_j, k_j with their analytic continuations $\widehat{f}_j$, $\widehat{g}_j$, $\widehat{h}_j$, $\widehat{k}_j$ or if we choose $s \notin \mathrm{i}\mathbb{R} := \{\mathrm{i}\omega \mid \omega \in \mathbb{R}\}$. 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 $\widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}$, $\widehat{H}_*^{\mathcal{V}_k,\mathcal{W}_k}$, $\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}$. As the subsequent arguments are for these complex continuations, in the rest of the proof

$$\overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\eta) := \{\mu \in \mathbb{C}^{\mathsf{d}} \mid \|\mu - \mu_*\|_2 \le \eta\} \quad \text{and} \quad \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\eta) := \{s \in \mathbb{C} \mid |s - \mathrm{i}\omega_*| \le \eta\}$$

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 $(\mu,\omega) \mapsto H^{\mathcal{V}_k,\mathcal{W}_k}[\mu](\mathrm{i}\omega)$ established in parts (i) and (ii) of Lemma A.1 extend to its complex counterpart. In particular, there exist constants γ, $\widehat{\eta}_\mu$, $\widehat{\eta}_\omega$ such that

$$\begin{aligned}
&\left\|\widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}\left[\widetilde{\mu}\right]\left(\widetilde{s}\right) - \widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)\right\|_2\\
&\quad\le \left\|\widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}\left[\widetilde{\mu}\right]\left(\widetilde{s}\right) - \widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}[\mu]\left(\widetilde{s}\right)\right\|_2 + \left\|\widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}[\mu]\left(\widetilde{s}\right) - \widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)\right\|_2\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \le \gamma\left(\left\|\widetilde{\mu} - \mu\right\|_2 + \left|\widetilde{s} - s\right|\right)
\end{aligned}$$

for all $\widetilde{\mu}$, $\mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\widehat{\eta}_\mu) \subset \mathbb{C}^{\mathsf{d}}$, and for all $\widetilde{s}$, $s \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\widehat{\eta}_\omega) \subset \mathbb{C}$. An analogous uniform Lipschitz continuity assertion also holds for $\widehat{H}_*^{\mathcal{V}_k,\mathcal{W}_k}$. Consequently, there exist constants γ, $\widehat{\eta}_\mu$, $\widehat{\eta}_\omega$ such that

$$\begin{aligned}
&\left\|\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}\left[\widetilde{\mu}\right]\left(\widetilde{s}\right) - \widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)\right\|_2 \le \gamma\left(\left\|\widetilde{\mu} - \mu\right\|_2 + \left|\widetilde{s} - s\right|\right)\\
&\qquad\qquad\qquad \forall\, \widetilde{\mu},\, \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\widehat{\eta}_\mu) \subset \mathbb{C}^{\mathsf{d}} \quad \forall\, \widetilde{s},\, s \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\widehat{\eta}_\omega) \subset \mathbb{C}.
\end{aligned} \tag{A.4}$$

Now, for $(\mu,s) \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\widehat{\eta}_\omega) \times \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\widehat{\eta}_\omega)$, consider the eigenvalue $\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\mu,s)$ of the analytic extension $\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)$ (i.e., $\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ is the extension of the eigenvalue $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ of $M^{\mathcal{V}_k,\mathcal{W}_k}[\cdot](\mathrm{i}\cdot)$). This eigenvalue function is no longer real valued, since $\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)$ is not necessarily a Hermitian matrix. However, by (A.3) and (A.4), as well as Theorem 5.1 in [18, Chapter 4], there exist constants $\eta_{\mu,m} \le \min\left\{\widetilde{\eta}_\mu, \widehat{\eta}_\mu\right\}$ and $\eta_{\omega,m} \le \min\left\{\widetilde{\eta}_\omega, \widehat{\eta}_\omega\right\}$ such that the eigenvalue $\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\mu,s)$ remains simple for all $\mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\eta_{\mu,m})$ and all $s \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\eta_{\omega,m})$. We remark that $\eta_{\mu,m}$ and $\eta_{\omega,m}$ are independent of $\mathcal{V}_k$, $\mathcal{W}_k$. 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 $\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ in the interior of $\overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\eta_{\mu,m}) \times \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\eta_{\omega,m})$, for a given $\widetilde{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*,\eta_\mu)$ and $\widetilde{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*,\eta_\omega/2)$, by Cauchy's integral formula we have

$$\widehat{\sigma}_s^{\mathcal{V}_k,\mathcal{W}_k}(\widetilde{\mu},\widetilde{s}) = \frac{1}{2\pi\mathrm{i}} \oint_{\left|s-\widetilde{s}\right|=\eta_\omega/2} \frac{\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\widetilde{\mu},s)}{(s-\widetilde{s})^2}\mathrm{d}s. \tag{A.5}$$

We claim that the term $\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\widetilde{\mu},s)$ inside the integral in modulus is uniformly bounded from above. To this end, as $\left|\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}\left(\widetilde{\mu},s\right)\right| \le \left\|\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}\left[\widetilde{\mu}\right](s)\right\|_2$, it suffices to show the uniform boundedness of $\left\|\widehat{M}^{\mathcal{V}_k,\mathcal{W}_k}[\widetilde{\mu}](s)\right\|_2$. Letting $\beta := \sigma_{\min}(D(\mu_*, \mathrm{i}\omega_*))$ and following the arguments at the beginning of the proof of Lemma A.1, there exists a neighborhood $\widehat{\mathcal{N}} \subset \mathbb{C}^{\mathsf{d}} \times \mathbb{C}$ of $(\mu_*, \mathrm{i}\omega_*)$ such that $\sigma_{\min}\left(D^{V_k,W_k}(\mu,s)\right) \ge \beta/2$ for all $(\mu,s) \in \widehat{\mathcal{N}}$. Without loss of generality, we assume $\widehat{\mathcal{N}} = \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu) \times \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*, \eta_\omega)$ (as we can choose η_μ and η_ω as small as we wish). Hence,

$$\left\|\widehat{H}^{\mathcal{V}_k,\mathcal{W}_k}[\mu](s)\right\|_2 \le 2\frac{\left\|C^{V_k}(\mu)\right\|_2\left\|B^{W_k}(\mu)\right\|_2}{\beta} \le 2\frac{M_{\mathbb{C},C}M_{\mathbb{C},B}}{\beta} \quad \forall\, \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\ \forall\, s \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*, \eta_\omega),$$

where

$$M_{\mathbb{C},C} := \max\left\{\|C(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\right\}, M_{\mathbb{C},B} := \max\left\{\|B(\mu)\|_2 \mid \mu \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\right\}.$$

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 $s \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*, \eta_\omega)$, which gives rise to

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

We deduce from (A.5) that

$$\text{(A.6)} \quad \left|\widehat{\sigma}_s^{\mathcal{V}_k,\mathcal{W}_k}\left(\widetilde{\mu},\widetilde{s}\right)\right| \le \frac{1}{2\pi}\left\{\max_{\left|s-\widetilde{s}\right|=\eta_\omega/2}\left|\widehat{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}\left(\widetilde{\mu},s\right)\right|\right\}\frac{1}{(\eta_\omega/2)^2}(2\pi\eta_\omega/2) \le \frac{2M_{\mathbb{C}}}{\eta_\omega} \quad \forall\, \widetilde{\mu} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mu_*, \eta_\mu)\ \forall\, \widetilde{s} \in \overline{\mathcal{B}}_{\mathbb{C}}(\mathrm{i}\omega_*, \eta_\omega/2),$$

hence, also $\left|\sigma_\omega^{\mathcal{V}_k,\mathcal{W}_k}\left(\widetilde{\mu},\widetilde{\omega}\right)\right| \le 2M_{\mathbb{C}}/\eta_\omega$ for all $\widetilde{\mu} \in \overline{\mathcal{B}}(\mu_*, \eta_\mu)$ and all $\widetilde{\omega} \in \overline{\mathcal{B}}(\omega_*, \eta_\omega/2)$.

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

$$\text{(A.7)} \qquad \widehat{\sigma}_{s\mu_1}^{\mathcal{V}_k,\mathcal{W}_k}\left(\widehat{\mu},\widehat{s}\right) = \frac{1}{2\pi\mathrm{i}}\oint_{\mathcal{C}}\frac{\widehat{\sigma}_s^{\mathcal{V}_k,\mathcal{W}_k}\left(\mu,\widehat{s}\right)}{\left(\mu_1-\widehat{\mu}_1\right)^2}\mathrm{d}\mu_1,$$

where the contour integral is over

$$\mathcal{C} := \left\{\mu \in \mathbb{C}^{\mathsf{d}} \;\middle|\; \left|\mu_1-\widehat{\mu}_1\right| = \eta_\mu/2,\ \mu_j = \widehat{\mu}_j,\ j = 2,\ldots,\mathsf{d}\right\}.$$

Taking the modulus of both sides in (A.7) and using (A.6) yield

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

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

A.2. Well-posedness of the reduced interpolating functions. We deduce the well-posedness of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ by analyzing the dependence of the maximizer of the mapping $\omega \mapsto \sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ on μ. To this end, the next result draws the important conclusion that the maximizer of $\omega \mapsto \sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\omega)$ can be expressed as a smooth function of μ in a uniform neighborhood of μ_* independent of the subspaces.

PROPOSITION A.3. *Under Assumptions* 3.1 *and* 3.2*, the following assertions hold for some constants* $\eta_{\mu,0}$, $\eta_{\omega,0}$, $\varepsilon > 0$*:*

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

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

Furthermore, $\sigma_{\omega\omega}(\mu, \boldsymbol{\omega}(\mu)) \leq \delta/2$ *for all* $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$*, where* $\delta < 0$ *is as in Assumption* 3.2*.*

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

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

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

(iii) *We have*

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

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

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

(iv) *We have*

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

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

Proof. As argued in the opening of the proof of Lemma A.2, we have

$$\sigma(\mu, \omega) - \sigma_2(\mu, \omega) \geq \widehat{\varepsilon} \quad \forall (\mu, \omega) \in \overline{\mathcal{B}}(\mu_*, \widehat{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widehat{\eta}_\omega) \tag{A.8}$$

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

$$\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu, \omega) - \sigma_2^{\mathcal{V}_k,\mathcal{W}_k}(\mu, \omega) \geq \varepsilon \quad \forall (\mu, \omega) \in \overline{\mathcal{B}}(\mu_*, \widetilde{\eta}_\mu) \times \overline{\mathcal{B}}(\omega_*, \widetilde{\eta}_\omega) \tag{A.9}$$

for some constants $\varepsilon \in (0, \widehat{\varepsilon})$, $\widetilde{\eta}_\mu > 0$, $\widetilde{\eta}_\omega > 0$.

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 A.2, there exists a constant $\widetilde{\delta} > 0$ such that

$$\left|\sigma_{\omega\omega\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu, \omega)\right| \leq \widetilde{\delta} \tag{A.10}$$

holds uniformly for all (μ, ω) in a neighborhood of (μ_*, ω_*), where $\widetilde{\delta}$ and the neighborhood are independent of the subspaces $\mathcal{V}_k, \mathcal{W}_k$ as long as they satisfy Assumption 3.2. 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 Assumption 3.2 and (A.10), 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_*). Additionally, we must have $\omega^{(k)} \to \omega_*$ as $\mu^{(k)} \to \mu_*$ due to

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

as well as the uniqueness of ω_* as the maximizer of $\sigma(\mu_*, \cdot)$ and the continuity of $\sigma(\cdot,\cdot)$. Hence, consider $\mu^{(k)}$ sufficiently close to μ_* so that $\big(\mu^{(k)}, \omega^{(k)}\big) \in \mathcal{N}$, in particular, $\sigma_{\omega\omega}^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu^{(k)}, \omega^{(k)}\big) \leq \delta/2 < 0$. Now the assertion again follows from the implicit function theorem. The uniformity of the radii $\eta_{\mu,0}$, $\eta_{\omega,0}$ over the subspaces follows from the uniform upper bound $\delta/2$ on the second derivatives.

(iii) Assume $\eta_{\mu,0} \in (0, \widehat{\eta}_\mu)$, $\eta_{\omega,0} \in (0, \widehat{\eta}_\omega)$ without loss of generality, where $\widehat{\eta}_\mu$, $\widehat{\eta}_\omega$ are as in (A.8). But then for $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \subset \overline{\mathcal{B}}(\mu_*, \widehat{\eta}_\mu)$, we have $\boldsymbol{\omega}(\mu) \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0}) \subset \overline{\mathcal{B}}(\omega_*, \widehat{\eta}_\omega)$. Hence, (A.8) implies $\sigma(\mu, \boldsymbol{\omega}(\mu)) - \sigma_2(\mu, \boldsymbol{\omega}(\mu)) \geq \widehat{\varepsilon} > \varepsilon$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0})$.

To show that $\boldsymbol{\omega}(\mu)$ 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 A.2, there exists a neighborhood $\widetilde{\mathcal{N}}$ of (μ_*, ω_*) where the transfer function $(\mu,\omega) \mapsto H[\mu](\mathrm{i}\omega)$ is continuously differentiable. As a result, the largest singular value function $\sigma(\cdot,\cdot)$ is Lipschitz continuous, say with the Lipschitz constant ζ over $\widetilde{\mathcal{N}}$ 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 [14, 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 $\widetilde{\boldsymbol{\omega}}(\mu)$ 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 $\sigma_\omega(\mu, \widetilde{\boldsymbol{\omega}}(\mu)) = 0$. The fact that $\boldsymbol{\omega}(\mu)$ as in part (i) is the unique point in $\overline{\mathcal{B}}(\omega_*, \eta_{\omega,0})$ satisfying $\sigma_\omega(\mu, \boldsymbol{\omega}(\mu)) = 0$ is implied by the implicit function theorem. Hence, we must have $\widetilde{\boldsymbol{\omega}}(\mu) = \boldsymbol{\omega}(\mu)$, so $\boldsymbol{\omega}(\mu)$ is the unique global maximizer of $\sigma(\mu,\cdot)$.

(iv) We assume without loss of generality that $\eta_{\mu,0} \in (0, \widetilde{\eta}_\mu)$ and $\eta_{\omega,0} \in (0, \widetilde{\eta}_\omega)$. Consequently, $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu) \in \overline{\mathcal{B}}(\omega_*, \eta_{\omega,0}) \subset \overline{\mathcal{B}}(\omega_*, \widetilde{\eta}_\omega)$ for all $\mu \in \overline{\mathcal{B}}(\mu_*, \eta_{\mu,0}) \subset \overline{\mathcal{B}}(\mu_*, \widetilde{\eta}_\mu)$, so (A.9) yields $\sigma^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big) - \sigma_2^{\mathcal{V}_k,\mathcal{W}_k}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big) \geq \varepsilon$ for such μ.

The uniqueness of $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)$ 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 $\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\omega\omega}(\mu,\omega)\leq\delta/2<0$. 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 $\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)$ must also be the unique global maximizer of $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\mu,\cdot)$ in $\overline{\mathcal{B}}(\omega_*,\eta_{\omega,0})$. ∎

Now the well-posedness of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$, that is Proposition 3.4, is a simple corollary of Proposition A.3 combined with Lemma A.2.

Proof of Proposition 3.4. Part (i) is a restatement of part (i) of Lemma A.2, while parts (ii) and (iii) of Proposition 3.4 are immediately implied by parts (ii) and (iv) of Proposition A.3, respectively. ∎

A.3. Derivatives of the $\mathcal{H}_\infty$-norm and the reduced interpolating functions. The smoothness of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ and $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ around μ_* is implied by Proposition A.3. With a little more effort, below we extend the uniform upper bounds on the derivatives of $\sigma^{\mathcal{V}_k,\mathcal{W}_k}(\cdot,\cdot)$ in Lemma A.2 to the derivatives of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$, and deduce Proposition 3.5.

Proof of Proposition 3.5. (i) The assertion that $\|H[\cdot]\|_{\mathcal{H}_\infty}$ is three times continuous differentiable in the interior of $\overline{\mathcal{B}}(\mu_*,\eta_{\mu,0})$ is a simple corollary of parts (i) and (iii) of Proposition A.3, since $\|H[\mu]\|_{\mathcal{H}_\infty}=\sigma(\mu,\boldsymbol{\omega}(\mu))$, where $\sigma(\mu,\boldsymbol{\omega}(\mu))$ is simple and $\boldsymbol{\omega}(\mu)$ is three times continuously differentiable for all μ in the interior of $\overline{\mathcal{B}}(\mu_*,\eta_{\mu,0})$.

Similarly, three times continuous differentiability of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)$ in the interior of $\overline{\mathcal{B}}(\mu_*,\eta_{\mu,0})$ is a corollary of parts (ii) and (iv) of Proposition A.3.

(ii) As for part (a), the first assertion, that is the boundedness of the second derivatives of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ in $\overline{\mathcal{B}}(\mu_*,\widehat{\eta}_{\mu,0})$ is immediate. Let us prove the existence of a constant $\gamma>0$ such that

$$\left|\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}(\mu)\right|\leq\gamma\qquad\forall\,\mu\in\overline{\mathcal{B}}(\mu_*,\widehat{\eta}_{\mu,0})\tag{A.11}$$

for q, $r=1,\ldots,d$ independent of the subspaces. To this end, we first observe

$$\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}(\mu)=\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)+\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_r}(\mu).\tag{A.12}$$

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

$$\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_r}(\mu)=-\frac{\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_r\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)}{\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\omega\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)},$$

which we plug into (A.12) to obtain

$$\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}(\mu)=\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)-\frac{\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_r\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)}{\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\omega\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)}.\tag{A.13}$$

By part (ii) of Proposition A.3, we have $\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\omega\omega}\big(\mu,\boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)\leq\delta/2<0$ for all $\mu\in\overline{\mathcal{B}}(\mu_*,\eta_{\mu,0})$. Moreover, by Lemma A.2, all 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). Hence, we conclude with (A.11) as desired.

As for part (b), the boundedness of the third derivatives of $\|H[\cdot]\|_{\mathcal{H}_\infty}$ in $\overline{\mathcal{B}}(\mu_*, \widehat{\eta}_{\mu,0})$ is also immediate from its three times continuous differentiability. The boundedness of the absolute values of the third derivatives of $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}(\cdot)$ uniformly by a constant independent of the subspaces can be established in a similar way by extending the approach in the previous two paragraphs for the second derivatives. Specifically, by differentiating (A.13) with respect to μ_ℓ, it can be seen that $\widetilde{\sigma}^{\mathcal{V}_k,\mathcal{W}_k}_{\mu_q\mu_r\mu_\ell}(\mu)$ is a ratio, where the expression in the numerator is a sum of products of the second derivatives $\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\chi_1,\chi_2}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)$ and third derivatives $\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\chi_1,\chi_2\chi_3}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)$ for $\chi_1, \chi_2, \chi_3 \in \{\omega, \mu_q, \mu_r, \mu_\ell\}$, while the expression in the denominator is $\sigma^{\mathcal{V}_k,\mathcal{W}_k}_{\omega\omega}\big(\mu, \boldsymbol{\omega}^{\mathcal{V}_k,\mathcal{W}_k}(\mu)\big)^3$. Hence, once again, the conclusion

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

for q, r, $\ell = 1, \ldots, \mathsf{d}$ for some constant γ can be drawn from part (ii) of Proposition A.3 and Lemma A.2. $\square$

Code availability. The MATLAB implementation of our algorithm, test data as well as the computational results are publicly available under the DOI 10.5281/zenodo.3533086.

Acknowledgment. The authors would like to thank the three anonymous referees for their valuable comments, which helped to improve the presentation.

REFERENCES

[1] N. Aliyev, P. Benner, E. Mengi, P. Schwerdtner, and M. Voigt, *Large-scale computation of $\mathcal{L}_\infty$-norms by a greedy subspace method*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1496–1516.

[2] U. Baur, C. Beattie, P. Benner, and S. Gugercin, *Interpolatory projection methods for parameterized model reduction*, SIAM J. Sci. Comput., 33 (2011), pp. 2489–2518.

[3] 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.

[4] C. Beattie and S. Gugercin, *Interpolatory projection methods for structure-preserving model reduction*, Systems Control Lett., 58 (2009), pp. 225–232.

[5] P. Benner, V. Sima, and M. Voigt, *$\mathcal{L}_\infty$-norm computation for continuous-time descriptor systems using structured matrix pencils*, IEEE Trans. Automat. Control, 57 (2012), pp. 233–238.

[6] S. Boyd and V. Balakrishnan, *A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm*, Systems Control Lett., 15 (1990), pp. 1–7.

[7] L. Breiman and A. Cutler, *A deterministic algorithm for global optimization*, Math. Program., 58 (1993), pp. 179–199.

[8] 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.

[9] 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.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[11] A. Johansson, J. Wei, H. Sandberg, K. H. Johansson, and J. Chen, *Optimization of the $\mathcal{H}_\infty$-norm of Dynamic Flow Networks*, in Proceedings of the 2018 Annual American Control Conference, Milwaukee, MN, pp. 1280–1285.

[12] F. Kangal, K. Meerbergen, E. Mengi, and W. Michiels, *A subspace method for large scale eigenvalue optimization*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 48–82.

[13] 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.

[14] E. Mengi, *Large-scale and global maximization of the distance to instability*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1776–1809.
[15] 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.
[16] E. B. Rudnyi and J. G. Korvink, *Boundary condition independent thermal model*, in Dimension Reduction of Large-Scale Systems, Lect. Notes Comput. Sci. Eng. 45, P. Benner, D. C. Sorensen, and V. Mehrmann, eds., Springer, Berlin, 2005, pp. 345–348.
[17] P. Schwerdtner and M. Voigt, *Computation of the $\mathcal{L}_\infty$-norm using rational interpolation*, IFAC-PapersOnLine, 51 (2018), pp. 84–89.
[18] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[19] A. Varga and P. Parillo, *Fast algorithms for solving H_∞-norm minimization problems*, in Proceedings of the 40th IEEE Conference on Decision and Control, Orlando, FL, IEEE, Piscataway, NJ, 2001, pp. 261–266.
[20] 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.
[21] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1996.