Approximation of stability radii for large-scale dissipative Hamiltonian systems

Nicat Aliyev[1] · Volker Mehrmann[2] 🄳 · Emre Mengi[3]

Abstract

A linear time-invariant dissipative Hamiltonian (DH) system $\dot{x} = (J - R)Qx$, with a skew-Hermitian J, a Hermitian positive semidefinite R, and a Hermitian positive definite Q, is always Lyapunov stable and under further weak conditions even asymptotically stable. By exploiting the characterizations from Mehl et al. (SIAM J. Matrix Anal. Appl. **37**(4), 1625–1654, 2016), we focus on the estimation of two stability radii for large-scale DH systems, one with respect to non-Hermitian perturbations of R in the form $R + B\Delta C^H$ for given matrices B, C, and another with respect to Hermitian perturbations in the form $R + B\Delta B^H$, $\Delta = \Delta^H$. We propose subspace frameworks for both stability radii that converge at a superlinear rate in theory. The one for the non-Hermitian stability radius benefits from the DH structure-preserving model order reduction techniques, whereas for the Hermitian stability radius we derive subspaces yielding a Hermite interpolation property between the full and projected problems. With the proposed frameworks, we are able to estimate the two stability radii accurately and efficiently for large-scale systems which include a finite-element model of an industrial disk brake.

Keywords Dissipative hamiltonian system · Robust stability · Stability radius · Eigenvalue optimization · Subspace projection · Structure-preserving subspace framework · Hermite interpolation

Mathematics Subject Classification (2010) 65F15 · 93D09 · 93A15 · 90C26

Extended author information available on the last page of the article.

1 Introduction

Linear time-invariant Dissipative Hamiltonian (DH) systems are of the form

$$\dot{x} = (J - R)Qx, \tag{1.1}$$

where $Q = Q^H \in \mathbb{C}^{n \times n}$ is a Hermitian positive definite matrix denoted by $Q > 0$, whereas $J = -J^H \in \mathbb{C}^{n \times n}$, and the dissipation matrix $R = R^H \in \mathbb{C}^{n \times n}$ is positive semidefinite, denoted by $R \geq 0$. DH systems arise as homogeneous parts of port-Hamiltonian (PH) systems without inputs and outputs.

PH and DH systems play an essential role in many areas of science and engineering, see e.g. [18, 26]. They are automatically Lyapunov stable, i.e., all eigenvalues of $A = (J - R)Q$ are in the closed left half of the complex plane, and any eigenvalue on the imaginary axis is semisimple, see [21]. However, DH systems are not necessarily asymptotically stable, since A may have purely imaginary eigenvalues. Even if a DH system is asymptotically stable, small perturbations that do not preserve the DH structure may cause the system to become unstable so that the solution exhibits exponential growth for some initial conditions. Asymptotic stability in the presence of uncertainty can only be guaranteed when the stability radius of the system is reasonably large, i.e., the system is not only asymptotically stable but also reasonably away from the set of systems with purely imaginary eigenvalues. Hence, an estimation of the stability radius, which involves an optimization problem over all admissible perturbations, is an important ingredient of a proper stability analysis. One concrete real application where the knowledge of the stability radius is invaluable is the avoidance of brake squeals, which is presented next.

1.1 A motivating real world application

Disk brake squeal is a well-known problem in mechanical engineering. It occurs due to self-excited vibrations caused by instability at the pad-rotor interface [1]. In [12], FE models for disk brakes are derived in the form of second order differential equations

$$M\ddot{x} + D(\Omega)\dot{x} + K(\Omega)x = f, \tag{1.2}$$

with large and sparse symmetric coefficient matrices M, $D(\Omega)$, and $K(\Omega) \in \mathbb{R}^{n \times n}$ depending on the rotational speed $\Omega > 0$ such that $M > 0$, $K(\Omega) > 0$ and $D(\Omega) \geq 0$. The function f represents a forcing term or control; hence, it may be set as $f = 0$ for the stability analysis. The incorporation of gyroscopic effects, modeled by a term $G(\Omega)\dot{x}$, with $G(\Omega) = -G(\Omega)^H$, and circulatory effects, modeled by a non-symmetric term Nx, gives rise to a system

$$M\ddot{x} + (D(\Omega) + G(\Omega))\dot{x} + (K(\Omega) + N)x = 0. \tag{1.3}$$

Expressing (1.3) (with $f = 0$) as a first order ODE followed by a straightforward manipulation yields the system

$$\dot{\tilde{z}} = (J - R)Q\tilde{z}, \tag{1.4}$$

where

$$
J = \begin{bmatrix} -G(\Omega) & -K(\Omega) \\ K(\Omega) & 0 \end{bmatrix}, \quad R = \begin{bmatrix} D(\Omega) & N \\ 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} M & 0 \\ 0 & K(\Omega) \end{bmatrix}^{-1},
$$
$$
z = \begin{bmatrix} \dot{x} \\ x \end{bmatrix}, \quad \widetilde{z} = Q^{-1}z. \tag{1.5}
$$

In the absence of the circulatory effects, i.e., when $N = 0$, this is a DH system and, as a result, it is Lyapunov stable and typically it is also asymptotically stable. However, small circulatory effects, i.e., perturbations by a non-symmetric N of small norm, may result in instability. Here, we emphasize that the perturbations due to the matrix N are restricted to the finite element nodes on the brake pad and the associated nodes on the disk. Moreover, the number of such nodes is several orders of magnitude smaller than the system size.

In modern brakes, damping devices (shims) that increase the damping of the brake pad are used to avoid resonances between the disk and the brake pad. The uncertainties and the additional damping due to shims mainly affect the matrices $D(\Omega)$ and N.

Note that for this brake squeal example leading to the DH system in (1.4) and (1.5), the matrix Q^{-1} is sparse, whereas Q is usually not. This however does not cause any difficulty in our approaches; as elaborated on later when we perform numerical experiments, the ideas we introduce throughout the paper can be modified so that they are based on Q^{-1} rather than Q.

1.2 Scope

In this paper, we study the stability radii of large-scale DH systems of the form (1.1) in the presence of perturbations. We typically assume that the DH systems at hand are sparse; however, the approaches we propose may also be applicable to dense systems as long as it is feasible to solve a few linear systems of size equal to the order of the system.

Motivated by the disk brake example, we focus on perturbations in the coefficient matrix R of (1.1). In particular, we consider the following two stability radii for a DH system of the form (1.1) originally introduced in [21], where the notation $i\mathbb{R}$ refers to the imaginary axis in the complex plane, $\Lambda(A)$ to the spectrum of a matrix A, and $\|A\|_2$ to the spectral norm of A.

Definition 1 Consider a DH system of the form (1.1) and suppose that $B \in \mathbb{C}^{n \times m}$ and $C \in \mathbb{C}^{p \times n}$, $n \geq m, p$, are given restriction matrices.

– The non-Hermitian restricted stability radius $r(R; B, C)$ with respect to perturbations of R under the fixed restriction matrices B, C is defined by

$$
r(R; B, C) := \inf\{\|\Delta\|_2 \mid \Delta \in \mathbb{C}^{m \times p}, \ \Lambda\left((J - (R + B\Delta C))Q\right) \cap i\mathbb{R} \neq \emptyset\}.
$$

– The Hermitian restricted stability radius with respect to Hermitian perturbations of R under the fixed restriction matrix B is defined by

$$r^{\text{Herm}}(R; B) := \inf\{\|\Delta\|_2 \mid \Delta = \Delta^H \in \mathbb{C}^{m \times m},$$
$$\Lambda\big((J - (R + B\Delta B^H))Q\big) \cap i\mathbb{R} \neq \emptyset\}. \quad (1.6)$$

The significance of the two stability radii defined above in the context of the disk brake example is discussed next.

Example 1 In the brake-squeal example in Section 1.1, it is of interest to know whether, for given Ω, the norm of the non-symmetric matrix N is tolerable to preserve the asymptotic stability of the system in (1.4). The relevant stability radius for a specified Ω is given by

$$\inf \{\|N\|_2 \mid \Lambda\,(\mathcal{A}(N)) \cap i\mathbb{R} \neq \emptyset\}, \quad (1.7)$$

where

$$\mathcal{A}(N) := \left\{ \begin{bmatrix} -G(\Omega) & -K(\Omega) \\ K(\Omega) & 0 \end{bmatrix} - \left(\begin{bmatrix} D(\Omega) & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & N \\ 0 & 0 \end{bmatrix} \right) \right\} \begin{bmatrix} M & 0 \\ 0 & K(\Omega) \end{bmatrix}^{-1}.$$

The perturbations in N only act on the FE nodes associated with the brake pad. Hence, the stability radius in (1.7) corresponds to the non-Hermitian restricted stability radius $r(R; B, C)$ with the restriction matrices $B = \begin{bmatrix} B_p & 0 \end{bmatrix}^H$ and $C = \begin{bmatrix} 0 & C_p \end{bmatrix}$ with blocks B_p, C_p associated with these finite element nodes. On the other hand, the Hermitian stability radius $r^{\text{Herm}}(R; B)$ is associated with restricted Hermitian perturbations with $B = \begin{bmatrix} B_p & 0 \end{bmatrix}^H, C = B^H$ when the Hermitian matrix $D(\Omega)$ is subject to Hermitian perturbations.

Note that the Hermitian stability radius $r^{\text{Herm}}(R; B)$ is finite (see Theorem 3 below and [21, Corollary 4.10]) and can be arbitrarily small. Additionally, by definition, we have $r(R; B, B^H) \leq r^{\text{Herm}}(R; B)$. Indeed, these two stability radii can differ considerably, see also [21]. Our main concern is the approximation of these two stability radii for large-scale DH systems when $n \gg m, p$. We propose subspace frameworks that replace the large-scale DH system with a small-scale system whose stability radii approximate those of the large-scale system usually well.

Contributions The stability radius under non-Hermitian restricted perturbations $r(R; B, C)$ can be characterized as the reciprocal of the $\mathcal{H}_\infty$ norm of the transfer function of a linear control system, see in particular (2.1) in the next section. It should be noted that, due to the DH structure, the singular value function that needs to be optimized in the $\mathcal{H}_\infty$-norm computation often has many local maximizers.

For the estimation of $r(R; B, C)$, we propose a DH structure–preserving subspace framework that converges at a superlinear rate with respect to the subspace dimension. The idea of the new method is partly inspired by the subspace framework in [2] applicable for the approximation of the $\mathcal{H}_\infty$ norm of an asymptotically stable system, but we form the subspaces in a different way by combining it with DH structure–preserving model order reduction techniques, see [11, 15, 16, 24]. Both the

framework in [2] and the one proposed here seem to converge to local maximizers that are not necessarily global maximizers. However, our numerical experience on synthetic examples indicates that the DH structure–preserving framework converges to the global maximizer more frequently compared with the unstructured framework in [2]. There are other approaches [5–7, 10, 17, 23] for approximating the $\mathcal{H}_\infty$ norm of a large-scale asymptotically stable system, but none of these employ subspace projections.

The Hermitian restricted stability radius $r^{\mathrm{Herm}}(R; B)$ poses more challenges. Unlike the characterization for $r(R; B, C)$, the eigenvalue optimization characterization for $r^{\mathrm{Herm}}(R; B)$ (given in Theorem 3) does not appear to be related to any transfer function. We derive a new subspace projection framework in which Hermite interpolation properties between the original and projected objective eigenvalue functions hold.

Outline The paper is organized as follows. Section 2 proposes subspace frameworks for approximating $r(R; B, C)$, while Section 3 concerns $r^{\mathrm{Herm}}(R; B)$. For the latter, even the solution of the small-scale problems is challenging. Hence, we first discuss how small-scale problems can be solved in Section 3.1. The new structured subspace framework is then presented in Section 3.2. In Sections 2 and 3, we illustrate the new frameworks by means of numerical experiments on the disk brake example and several synthetic examples.

A longer version of this article discussing also perturbations in coefficients other than R, and featuring additional numerical results is available in [4].

2 The non-Hermitian stability radius for large-scale problems

In this section, we study the estimation of the non-Hermitian restricted stability radius $r(R; B, C)$ for large-scale asymptotically stable DH systems. Our approach operates on the characterization [21, Theorem 3.3]

$$
\begin{aligned}
r(R; B, C) \;&=\; 1/\|G\|_{\mathcal{H}_\infty}, \\
\|G\|_{\mathcal{H}_\infty} \;&:=\; \sup_{s \in \mathbb{C}^+} \sigma_{\max}(G(s)) = \max_{\omega \in \mathbb{R}} \sigma_{\max}(G(\mathrm{i}\omega)).
\end{aligned}
\tag{2.1}
$$

for an asymptotically stable DH system of the form (1.1) and restriction matrices $B \in \mathbb{C}^{n \times m}$, $C \in \mathbb{C}^{p \times n}$. Above, we use the notation

$$
G(s) := C Q(s I - (J - R)Q)^{-1} B,
\tag{2.2}
$$

the notation $\sigma_{\max}(G(s))$ for the largest singular value of $G(s)$, and $\mathbb{C}^+ := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 0\}$ for the open right-half of the complex plane. Note that the second equality in the second line in (2.1) follows from the maximum modulus principle, as the matrix $(J - R)Q$ does not have any eigenvalue in $\mathbb{C}^+$.

The matrix-valued function $G(s)$ in (2.2) corresponds to the transfer function of the control system

$$
\dot{x} = \widetilde{A}x + \widetilde{B}u, \quad y = \widetilde{C}x,
\tag{2.3}
$$

where we set $\widetilde{A} := (J - R)Q$, $\widetilde{B} := B$ and $\widetilde{C} := CQ$. Moreover, the quantity $\|G\|_{\mathcal{H}_\infty}$ is commonly referred as the $\mathcal{H}_\infty$ norm of the transfer function $G(s)$.

2.1 The subspace framework of [2]

Recently, in [2], a subspace framework for the approximation of the $\mathcal{H}_\infty$ norm of a large-scale asymptotically stable system has been proposed. At iteration k, it requires two subspaces $\mathcal{V}_k$ and $\mathcal{W}_k$ of equal dimension, as well as matrices V_k and W_k, respectively, whose columns form orthonormal bases for these subspaces. The state x of system (2.3) is restricted to $\mathcal{V}_k$, i.e., in (2.3) the state x is replaced by $V_k x_k$, and it is imposed that the residual of the differential part after this restriction is orthogonal to $\mathcal{W}_k$.

This gives rise to the reduced order system

$$\widetilde{E}_k \dot{x}_k = \widetilde{A}_k x_k + \widetilde{B}_k u, \quad y_k = \widetilde{C}_k x_k, \tag{2.4}$$

with $\widetilde{E}_k := W_k^H V_k$, $\widetilde{A}_k := W_k^H \widetilde{A} V_k$, $\widetilde{B}_k := W_k^H \widetilde{B}$, and $\widetilde{C}_k := \widetilde{C} V_k$. Then letting $G_k(s) := \widetilde{C}_k (s\widetilde{E}_k - \widetilde{A}_k)^{-1} \widetilde{B}_k$ be the reduced order transfer function, and assuming that $\widetilde{E}_k$ is invertible, a maximizer $\omega_{k+1} \in \arg\max_{\omega \in \mathbb{R}} \sigma_{\max}(G_k(i\omega))$ can be computed efficiently and reliably if the dimensions of $\mathcal{V}_k$, $\mathcal{W}_k$ are small by employing the globally convergent method in [8, 9]. Once such an ω_{k+1} has been determined, then the subspaces $\mathcal{V}_k$ and $\mathcal{W}_k$ are expanded into larger subspaces $\mathcal{V}_{k+1}$ and $\mathcal{W}_{k+1}$ in such a way that the corresponding reduced transfer function $G_{k+1}(s)$ satisfies the Hermite interpolation conditions

$$\begin{aligned}
\sigma_{\max}(G(i\omega_{k+1})) &= \sigma_{\max}(G_{k+1}(i\omega_{k+1})), \\
\frac{d}{d\omega}\sigma_{\max}(G(i\omega))\bigg|_{\omega=\omega_{k+1}} &= \frac{d}{d\omega}\sigma_{\max}(G_{k+1}(i\omega))\bigg|_{\omega=\omega_{k+1}}.
\end{aligned} \tag{2.5}$$

Denoting the range space of a matrix by A by $\text{Ran}(A)$, it is shown in [2] that the inclusions

$$\text{Ran}((i\omega_{k+1}I - \widetilde{A})^{-1}\widetilde{B}) \subseteq \mathcal{V}_{k+1}, \quad \text{Ran}((\widetilde{C}(i\omega_{k+1}I - \widetilde{A})^{-1})^H) \subseteq \mathcal{W}_{k+1},$$

ensure that the conditions in (2.5) are satisfied. For instance when $m = p$, in practice, we set $V_{k+1} := \text{orth}\left(\begin{bmatrix} V_k & (i\omega_{k+1}I - \widetilde{A})^{-1}\widetilde{B} \end{bmatrix}\right)$, and $W_{k+1} := \text{orth}\left(\begin{bmatrix} W_k & (i\omega_{k+1}I - \widetilde{A})^{-H}\widetilde{C}^H \end{bmatrix}\right)$, where $\text{orth}(M)$ denotes a matrix whose columns form an orthonormal basis for $\text{Ran}(M)$. The definitions of the matrices V_{k+1}, W_{k+1} can be modified in a simple way when $m \neq p$ so that they have equal number of columns; we refer to [2] for the details. The procedure is then repeated with the expanded subspaces $\mathcal{V}_{k+1}, \mathcal{W}_{k+1}$ spanned by the columns of V_{k+1}, W_{k+1}. The sequence $\{\omega_k\}$ always seems to converge to a local maximizer, but a formal proof is open at the moment. In [2], assuming that the sequence $\{\omega_k\}$ converges to a local maximizer, it is proven that this sequence converges at a superlinear rate due to the satisfaction of (2.5) for all k. The solution of the reduced problem $\max_{\omega \in \mathbb{R}} \sigma_{\max}(G_k(i\omega))$ by the method in [8, 9] requires extraction of all eigenvalues of $2d \times 2d$ pencils, where d is the order of the system, i.e., the dimension of the subspaces $\mathcal{V}_k, \mathcal{W}_k$. Assuming that d is small, this is much cheaper than solving

the linear systems $(i\omega_{k+1}I - \widetilde{A})^{-1}\widetilde{B}$, $(i\omega_{k+1}I - \widetilde{A})^{-H}\widetilde{C}^H$ involving the full order system, which is needed to form V_{k+1}, W_{k+1}.

Even though $\widetilde{A} = (J - R)Q$ has the DH structure, this is in general not true for the reduced system $\widetilde{A}_k$. In the next subsection, we modify the procedure of [2] to preserve the DH structure.

2.2 A structure-preserving subspace framework

In this subsection, we introduce an interpolating and DH structure–preserving version of the subspace projection framework. Structure-preserving subspace projection methods in the context of model order reduction of large-scale PH and DH systems have been proposed in [15, 16, 24]. Our approach is inspired by the following interpolation result, see [16, Remark 3]. For this result and the subsequent arguments, we introduce the matrix-valued function

$$T(s) := (J - R)Q - sI \tag{2.6}$$

depending on $s \in \mathbb{C}$ associated with the DH system in (2.3).

Theorem 1 (**Right Tangential Interpolation** [16]) *Let $G(s)$ be the transfer function for the system* (2.3), *and let $G_k(s)$ be the transfer function for the reduced system* (2.4). *For given $\widehat{s} \in \mathbb{C}$, $\widehat{b} \in \mathbb{C}^m$ and a positive integer N, if*

$$T(\widehat{s})^{-\ell}\widetilde{B}\widehat{b} \in \mathcal{V}_k \qquad for \ \ell = 1, \dots, N, \tag{2.7}$$

and W_k is such that $W_k^H V_k = I$, then, denoting the ℓ-th derivatives of $G(s)$ and $G_k(s)$ at the point $\widehat{s}$ with $G^{(\ell)}(\widehat{s})$ and $G_k^{(\ell)}(\widehat{s})$, we have

$$G^{(\ell)}(\widehat{s})\widehat{b} = G_k^{(\ell)}(\widehat{s})\widehat{b} \qquad for \ \ell = 0, \dots, N - 1 \tag{2.8}$$

provided that both $\widehat{s}I - \widetilde{A}$ and $\widehat{s}I - \widetilde{A}_k$ are invertible.

The characterization of $r(R; B, C)$ involves the maximization of the largest singular value of the transfer function $G(s)$ on the imaginary axis. In the following, we make use of Theorem 1 to obtain a reduced order system satisfying the interpolation conditions (2.8) while retaining the structure.

For a given $\widehat{s} \in \mathbb{C}$ and $\widehat{b} \in \mathbb{C}^m$, choose $\mathcal{V}_k$ to satisfy (2.7). Then define

$$W_k := QV_k(V_k^H QV_k)^{-1}, \tag{2.9}$$

so that

$$W_k^H V_k = I \quad \text{and} \quad (W_k V_k^H)^2 = W_k V_k^H, \tag{2.10}$$

i.e., $W_k V_k^H$ is an oblique projector onto $\text{Ran}(QV_k)$. The reduced system matrices $\widetilde{A}_k, \widetilde{B}_k, \widetilde{C}_k$ then satisfy

$$\begin{aligned} \widetilde{A}_k &= W_k^H \widetilde{A} V_k = W_k^H (J - R)QV_k = W_k^H (J - R)W_k V_k^H QV_k \\ &= (J_k - R_k)Q_k, \end{aligned} \tag{2.11}$$

where we define

$$J_k := W_k^H J W_k = -J_k^H, \ \ R_k := W_k^H R W_k = R_k^H \geq 0,$$
$$Q_k := V_k^H Q V_k = Q_k^H > 0, \tag{2.12}$$

and

$$\widetilde{C}_k = \widetilde{C} V_k = C Q V_k = C W_k V_k^H Q V_k = C_k Q_k, \ \ C_k := C W_k,$$
$$\widetilde{B}_k = W_k^H B =: B_k. \tag{2.13}$$

Moreover, $\widetilde{E}_k = W_k^H V_k = I$. Using this construction and Theorem 1, we deduce the following result that indicates particular projection subspaces to achieve Hermite interpolation between the full and reduced order systems in a structure-preserving way.

Theorem 2 *Consider a linear system as in* (2.3) *with the transfer function $G(s)$ in* (2.2). *Furthermore, for a given point $\widehat{s} \in \mathbb{C}$ and a given tangent direction $\widehat{b} \in \mathbb{C}^m$, suppose that V_k is a matrix with orthonormal columns such that*

$$T(\widehat{s})^{-\ell} B\widehat{b} \in Ran(V_k) \quad for \ \ell = 1, \dots, N$$

for some positive integer N, where $T(s)$ is as in (2.6). *Define W_k as in* (2.9), *J_k, R_k, Q_k as in* (2.12), *and B_k, C_k as in* (2.13). *Then the resulting reduced order system*

$$\begin{aligned} \dot{x}_k &= (J_k - R_k)Q_k x_k + B_k u, \\ y_k &= C_k Q_k x_k \end{aligned}$$

is a DH system with the transfer function

$$G_k(s) := C_k Q_k (sI - (J_k - R_k)Q_k)^{-1} B_k \tag{2.14}$$

that satisfies

$$G^{(\ell)}(\hat{s})\hat{b} = G_k^{(\ell)}(\hat{s})\hat{b} \quad for \ \ \ell = 0\dots, N-1,$$

where $G^{(\ell)}(\hat{s})$, $G_k^{(\ell)}(\hat{s})$ denote the ℓ-th derivatives of $G(s)$, $G_k(s)$ at $\hat{s}$.

According to Theorem 2, for a given $\widehat{s} \in \mathbb{C}$, setting

$$V_k := \text{orth}\left(\begin{bmatrix} T(\widehat{s})^{-1}B & T(\widehat{s})^{-2}B \end{bmatrix}\right), \ \ W_k := QV_k(V_k^H Q V_k)^{-1},$$

we obtain $G(\widehat{s}) = G_k(\widehat{s})$ and $G'(\widehat{s}) = G_k'(\widehat{s})$, and thus the Hermite interpolation conditions

$$\sigma_{\max}(G(\widehat{s})) = \sigma_{\max}(G_k(\widehat{s})), \ \ \frac{d}{ds}\sigma_{\max}(G(s))\bigg|_{s=\widehat{s}} = \frac{d}{ds}\sigma_{\max}(G_k(s))\bigg|_{s=\widehat{s}} \tag{2.15}$$

are satisfied. This suggests the greedy subspace framework outlined in Algorithm 1 for the approximation of $r(R; B, C)$.

Algorithm 1 DH structure-preserving subspace method for the estimation of the stability radius $r(R; B, C)$ for large-scale systems.

Require: Matrices $B \in \mathbb{C}^{n \times m}$, $C \in \mathbb{C}^{p \times n}$, $J, R, Q \in \mathbb{C}^{n \times n}$.

Ensure: Sequence $\{\sigma_k\}$ s.t. $1/\sigma_k$ for sufficiently large k is an estimate for $r(R; B, C)$.

1: Choose initial interpolation points $\omega_1, \ldots, \omega_q \in \mathbb{R}$.

2: $V_q \leftarrow \mathrm{orth} \left(\left[\, T(\mathrm{i}\omega_1)^{-1} B \;\; T(\mathrm{i}\omega_1)^{-2} B \; \ldots \; T(\mathrm{i}\omega_q)^{-1} B \;\; T(\mathrm{i}\omega_q)^{-2} B \, \right] \right)$.

3: $W_q \leftarrow Q V_q (V_q^H Q V_q)^{-1}$.

4: **for** $k = q, \, q + 1, \, \ldots$ **do**

5:     Form G_k as in (2.14) for the choices of J_k, R_k, Q_k in (2.12), and B_k, C_k in (2.13).

6:     $\omega_{k+1} \leftarrow$ a global maximizer of $\sigma_{\max}(G_k(\mathrm{i}\omega))$ over $\omega \in \mathbb{R}$.

7:     $\widehat{V}_{k+1} \leftarrow \left[\, T(\mathrm{i}\omega_{k+1})^{-1} B \;\; T(\mathrm{i}\omega_{k+1})^{-2} B \, \right]$.

8:     $V_{k+1} \leftarrow \mathrm{orth} \left(\left[\, V_k \;\; \widehat{V}_{k+1} \, \right] \right)$.

9:     $W_{k+1} \leftarrow Q V_{k+1} (V_{k+1}^H Q V_{k+1})^{-1}$.

10:     $\sigma_{k+1} \leftarrow \sigma_{\max}(G_{k+1}(\mathrm{i}\omega_{k+1}))$.

11: **end for**

In line 6 of Algorithm 1, at every iteration, the subspace framework computes the $\mathcal{H}_\infty$ norm of a reduced system, in particular it computes the point $\mathrm{i}\omega_*$ on the imaginary axis where this $\mathcal{H}_\infty$ norm is attained. Then the current left and right subspaces are expanded in a way so that the resulting reduced system still has the DH structure and its transfer function Hermite interpolates the original transfer function at $\mathrm{i}\omega_*$. Since the Hermite interpolation conditions (2.15) are satisfied at $\widehat{s} = \mathrm{i}\omega_1, \ldots, \mathrm{i}\omega_k$ at the beginning of iteration k, the analysis in [2] shows a superlinear rate of convergence for the sequence $\{\omega_k\}$, assuming that this sequence converges to a local maximizer. We remark that $\sigma_{k+1} = \sigma_{\max}(G(\mathrm{i}\omega_{k+1}))$ due to line 10 of the algorithm and the Hermite interpolation property. Hence, if the sequence $\{\omega_k\}$ converges to a global maximizer, the reciprocal of the limit of the sequence $\{\sigma_k\}$ yields $r(R; B, C)$.

The most expensive part of Algorithm 1 occurs in lines 2 and 7, where several linear systems have to be solved. If this is done with a direct solver, for each value $\widehat{\omega} \in \mathbb{R}$ one LU factorization of the matrix $T(\mathrm{i}\widehat{\omega})$ suffices. For large n, the computation time is usually dominated by these LU factorizations. In contrast, the reduced problem in line 6 involves a small system; its solution can be obtained efficiently and robustly by the algorithm in [8, 9].

2.3 Numerical experiments

In this subsection, we illustrate the performance of MATLAB implementations of Algorithm 1 and the subspace framework from [2] via several numerical experiments carried out in MATLAB R2017b on a machine with Mac OS 10.12.6, an Intel® Core™ i5-4260U CPU and 4GB RAM.

Algorithm 1 and the subspace framework from [2] are terminated when at least one of the following three conditions is fulfilled:

1. The relative distance between ω_k and ω_{k-1} is less than a prescribed tolerance, i.e., $|\omega_k - \omega_{k-1}| < \varepsilon \cdot |\omega_k|$.
2. Letting $f_k := \max_{\omega \in \mathbb{R}} \sigma_{\max}(G_k(\mathrm{i}\omega))$, two consecutive iterates f_k, f_{k-1} are close enough in a relative sense, i.e., $|f_k - f_{k-1}| < \varepsilon \cdot f_k$.
3. The number of iterations exceeds a specified integer, i.e., $k > k_{\max}$.

In all of the presented numerical examples, we use $\varepsilon = 10^{-12}$ and $k_{\max} = 80$ for both Algorithm 1 and the subspace framework from [2].

Both Algorithms seem to converge locally in our experiments. The initial interpolation points affect whether convergence to a global maximizer occurs. A procedure to choose the initial interpolation points is outlined below. It uses an interval $[l, u]$, ideally containing a global minimizer. A reasonable strategy is to set l, u equal to the approximations of the imaginary parts of the eigenvalues of $(J - R)Q$ with the smallest and largest imaginary parts.

(i)   First, we discretize $[l, u]$ into ρ equally spaced points, say $\omega_{0,1}, \ldots, \omega_{0,\rho}$, including the end-points l, u, where ρ is specified by the user.
(ii) Then we either approximate the eigenvalues $z_1, \ldots, z_\rho$ of $(J - R)Q$ closest to $\mathrm{i}\omega_{0,1}, \ldots, \mathrm{i}\omega_{0,\rho}$ (using `eigs` in MATLAB), respectively, or otherwise set $z_j = \mathrm{i}\omega_{0,j}$ for $j = 1, \ldots, \rho$.
(iii) Next, we permute $z_1, \ldots, z_\rho$ into $z_{j_1}, \ldots, z_{j_\rho}$ where $\{j_1, \ldots, j_\rho\} = \{1, \ldots, \rho\}$ is such that $\sigma_{\max}(G(\mathrm{i}\Im z_{j_1})) \geq \cdots \geq \sigma_{\max}(G(\mathrm{i}\Im z_{j_\rho}))$.
(iv) The interpolation points $\omega_1, \ldots, \omega_\kappa$ employed initially are chosen as the imaginary parts of $z_{j_1}, \ldots, z_{j_\kappa}$, where again $\kappa \leq \rho$ is specified by the user.

For the synthetic examples below, we use $l = -2000$, $u = 0$, $\rho = 25$, $\kappa = 10$, and, in part (ii), $z_1, \ldots, z_\rho$ are set equal to the eigenvalues closest to $\mathrm{i}\omega_1, \ldots, \mathrm{i}\omega_\rho$. For the disk brake example, we use $l = -2 \cdot 10^5$, $u = -1.7 \cdot 10^5$, $\rho = 15$, $\kappa = 15$, and, in part (ii), we set $z_j = \mathrm{i}\omega_{0,j}$ for $j = 1, \ldots, \rho$. The values of l and u employed in these examples are chosen based on rough sketches of the largest singular value functions to be maximized.

2.3.1 Results on synthetic examples

We now present results for 2000 dense linear DH systems with random coefficient matrices J, R, Q of order 500, and the restriction matrices B and C chosen as 500×2 and 2×500 random matrices; MATLAB code generating such random coefficient matrices are openly available.[1]

The relative errors of the values returned by Algorithm 1 and [2, Algorithm 1] for the first 5 random examples are presented in Table 1, where f_{BB} and f_{SF} denote the results returned by the Boyd-Balakrishnan (BB) algorithm [8, 9] and the subspace framework (Algorithm 1 or [2, Algorithm 1]). Note that for the

---

[1]http://home.ku.edu.tr/~emengi/software/DH-stabradii-files/randomDH.m
This function needs to be called as `[J,R,Q,B,C] = randomDH(500,2,2)` to generate such random matrices.

Table 1 Relative errors of the values returned by Algorithm 1 and [2, Algorithm 1] on 5 dense random DH examples of order 500. The number of subspace iterations, run-time (in seconds) are also listed

Ex.	$|f_{SF} - f_{BB}|/f_{BB}$ Alg. 1	[2]	# iterations Alg. 1	[2]	run-time Alg. 1	[2]
1	$2.13 \cdot 10^{-14}$	$1.85 \cdot 10^{-14}$	8	6	8.8	6.8
2	$2.98 \cdot 10^{-15}$	$5.60 \cdot 10^{-2}$	4	4	6.8	6.7
3	$5.88 \cdot 10^{-15}$	$5.51 \cdot 10^{-16}$	6	3	7.2	6.6
4	$3.26 \cdot 10^{-15}$	$1.63 \cdot 10^{-14}$	9	53	7.3	33.3
5	$1.23 \cdot 10^{-14}$	$2.77 \cdot 10^{-15}$	7	15	7.2	7.7

second example, [2, Algorithm 1] returns an estimate with a significant error (highlighted in bold face). The unstructured distance to instability [27] $\Gamma((J - R)Q) := \inf\{\|\Delta\|_2 \mid \exists \Delta \in \mathbb{C}^{n \times n}, \ \Lambda((J - R)Q + \Delta) \cap i\mathbb{R} \neq \emptyset\}$ for all of these 5 examples is larger than $r(R; B, C)$ by ten times or even more.

Considering all 2000 examples, the non-Hermitian stability radii are estimated with relative errors less than the prescribed tolerance 10^{-12} for 1921 of these examples by Algorithm 1 and for 1821 by [2, Algorithm 1]. These numbers correspond to 96.45% and 91.45% accuracy for Algorithms 1 and [2, Algorithm 1], respectively. Algorithms 1 performs 10.92 subspace iterations on average, smaller than the average number of iterations 14.95 by [2, Algorithm 1]. The average run-times on this data set are quite similar, specifically 10.48 and 9.65 s for Algorithms 1 and [2, Algorithm 1], respectively.

Table 2 gives more detailed insight into the performance of the algorithms. It reports the accuracy of the results returned by Algorithm 1 and [2, Algorithm 1], as well as the average number of iterations and the average run-times, when only those examples in the data set for which the rank of R is greater than 14, 20, 30, 40, 50, respectively, are taken into account. (Note that for all examples we have Rank(R) > 14.) We call a result returned by one of the subspace frameworks accurate if the relative error $|f_{SF} - f_{BB}|/f_{BB}$ is less than the prescribed tolerance 10^{-12}. According to

Table 2 The accuracy of Algorithm 1 and [2, Algorithm 1] with respect to Rank(R) on the DH systems of order 500. The second column lists the number of examples in the data set that satisfy the rank constraint, whereas the third column lists the number of examples among these for which the returned result has error less than the prescribed tolerance. The average number of iterations and the average run-times (in seconds) are listed in the fifth and sixth columns

Rank of R >	# Ex.	# accurate Alg. 1	[2]	Accuracy % Alg. 1	[2]	# iter. Alg. 1	[2]	Run-time Alg. 1	[2]
14	2000	1929	1829	96.45%	91.45%	10.92	14.95	10.48	9.65
20	1853	1807	1744	97.52%	94.12%	10.39	14.13	9.55	8.98
30	1619	1595	1562	98.52%	96.48%	9.75	13.08	8.96	8.43
40	1390	1382	1363	99.42%	98.06%	9.20	12.09	8.61	8.08
50	1162	1157	1141	99.57%	98.19%	8.85	11.39	8.47	7.89

the figures in the table, not only the accuracy of the algorithms improves as the rank of R increases, but also the average number of iterations and the average run-time decrease.

2.3.2 The FE model of a disk brake

The only large-scale computation required by Algorithm 1 (after the preprocessing step that yield initial interpolation points) is the solution of the linear systems

$$T(\mathrm{i}\omega)\,X \; = \; B \quad \text{and} \quad T(\mathrm{i}\omega)^2\,Y \; = \; B \tag{2.16}$$

at a given $\omega \in \mathbb{R}$ in lines 2 and 7, where $T(\mathrm{i}\omega)$ is as in (2.6). For the DH system of the disk brake in (1.4) and (1.5), the mass matrix M and the stiffness matrix $K(\Omega)$ and thus the matrix Q^{-1} are sparse, but Q turns out to be dense. Trying to invert Q^{-1} and/or solve a linear system with the coefficient matrix Q is computationally very expensive and would require full matrix storage.

This difficulty can be avoided by exploiting that

$$((J - R)Q - \mathrm{i}\omega I)^{-1} B \;\; = \;\; Q^{-1}((J - R) - \mathrm{i}\omega Q^{-1})^{-1} B.$$

Hence, to compute X, Y as in (2.16), we proceed as follows.

1. We first solve $((J - R) - \mathrm{i}\omega Q^{-1})\widehat{X} = B$ for $\widehat{X}$, and set $X = Q^{-1}\widehat{X}$.
2. Then we solve $((J - R) - \mathrm{i}\omega Q^{-1})\widehat{Y} = X$ for $\widehat{Y}$, and set $Y = Q^{-1}\widehat{Y}$.

A second observation that further speeds up the computations is the particular structure of the coefficient matrix $((J - R) - \mathrm{i}\omega Q^{-1})$ with $N = 0$. Setting $\widetilde{M}(\mathrm{i}\omega; \Omega) := \mathrm{i}\omega M + D(\Omega) + G(\Omega)$, we have

$$(J - R) - \mathrm{i}\omega Q^{-1} \;\; = \;\; \begin{bmatrix} -\widetilde{M}(\mathrm{i}\omega; \Omega) & -K(\Omega) \\ K(\Omega) & -\mathrm{i}\omega K(\Omega) \end{bmatrix}.$$

Hence, to solve $((J - R) - \mathrm{i}\omega Q^{-1})Z = W$ for a given $W = \begin{bmatrix} W_1^T & W_2^T \end{bmatrix}^T$ and the unknown $Z = \begin{bmatrix} Z_1^T & Z_2^T \end{bmatrix}^T$ with W_1, W_2, Z_1, Z_2 having all equal number of rows, we perform a column block permutation and then eliminate the lower left block to obtain

$$\begin{bmatrix} -K(\Omega) & -\widetilde{M}(\mathrm{i}\omega; \Omega) \\ 0 & K(\Omega) + \mathrm{i}\omega\widetilde{M}(\mathrm{i}\omega; \Omega) \end{bmatrix} \begin{bmatrix} Z_2 \\ Z_1 \end{bmatrix} \;\; = \;\; \begin{bmatrix} W_1 \\ W_2 - \mathrm{i}\omega W_1 \end{bmatrix}.$$

At every subspace iteration, the highest costs arise from the computation of the LU factorizations of the sparse matrices $K(\Omega)$ and $K(\Omega) + \mathrm{i}\omega\widetilde{M}(\mathrm{i}\omega; \Omega)$.

We have applied Algorithm 1 to estimate the non-Hermitian stability radius $r(R; B, B^T)$ for a DH system of the form (1.4) and (1.5) resulting from a FE brake model with $N = 0$. In this example, $G(\Omega), K(\Omega), D(\Omega), M \in \mathbb{R}^{4669 \times 4669}$ yielding $J, R, Q \in \mathbb{R}^{9338 \times 9338}$, whereas $B \in \mathbb{R}^{9338 \times 3}$.

The plot of the computed estimate values for $r(R; B, B^T)$ with respect to the rotation speed $\Omega \in [2.5, 1800]$ is presented on the left in Fig. 1. It seems that $r(R; B, B^T)$ as a function of Ω exhibits nonsmoothness at several points, but such points in the figure are usually local or global maximizers. Two notable maximizers where non-smoothness is present are near $\Omega = 1120$ and $\Omega = 1590$. Denoting the former of
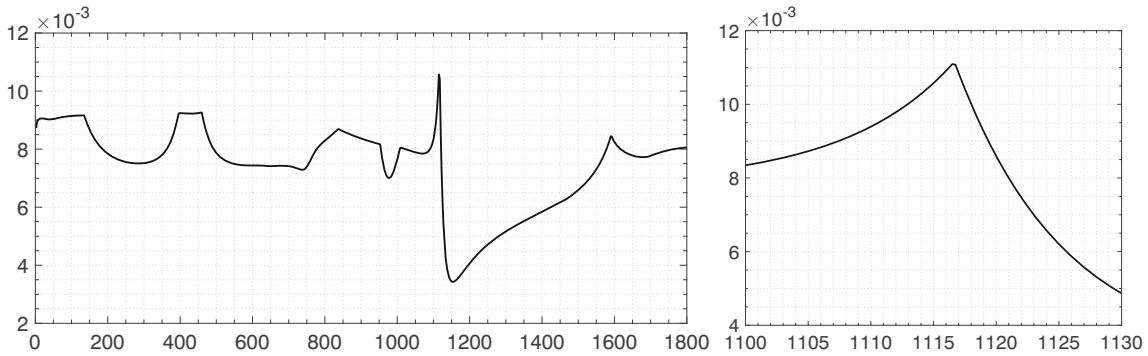
Fig. 1 (Left) Plot of the approximation to the stability radius $r(R; B, B^T)$ by Algorithm for the DH system specified by (1.4) and (1.5) with $N = 0$ as it varies with the rotation speed $\Omega \in [2.5, 1800]$. (Right) Plot is zoomed for $\Omega \in [1100, 1130]$

these maximizers with Ω_*, the plot of $r(R; B, B^T)$ for $\Omega \in [1100, 1130]$ containing Ω_* is zoomed on the right-hand plot in Fig. 1. The nonsmoothness here is due to the fact that $\sigma_{\max}(B^T Q(i\omega I - (J - R)Q)^{-1}B)$ as a function of ω has multiple global maximizers for the particular parameter value Ω_*. A similar remark applies to several other points where nonsmoothess is present.

The estimated values for $r(R; B, B^T)$ are listed in Table 3 for some values of Ω. For comparison purposes, we also estimate $r(R; B, B^T)$ by employing [2, Algorithm 1]. (This requires the solution of $T(i\omega)^H Y = C^H$ instead of the solution of $T(i\omega)^2 Y = B$ in (2.16). It is straightforward to extend the approach discussed above for the efficient solution of $T(i\omega)X = B$ to this setting.) The two algorithms return the same values of $r(R; B, B^T)$ up to prescribed tolerances for all Ω values in the table. In all cases, 2 subspace iterations are sufficient to achieve the prescribed accuracy for Algorithm 1. Both algorithms in the end replace the full problem of order 9338 with considerably smaller reduced systems, in particular of order 102 for Algorithm 1.

Table 3 Estimated values for the stability radii $r(R; B, B^T)$ in the second column by Algorithm 1 and [2, Algorithm 1] for a few Ω values for the system originating from the FE brake model. The other columns display ω_* corresponding to $\arg\max_\omega \sigma_{\max}(B^T Q(i\omega I - (J - R)Q)^{-1}B)$, the number of subspace iterations, total run-time (in seconds) and subspace dimension at termination

	$r(R; B, B^T)$		# iter.		Run-time		Dimension	
Ω	Alg. 1 & [2]	ω_*	Alg. 1	[2]	Alg. 1	[2]	Alg. 1	[2]
5	0.00893	$-1.741 \cdot 10^5$	2	10	30.9	46.8	102	75
50	0.00904	$-1.741 \cdot 10^5$	2	13	31.4	55.1	102	84
300	0.00752	$-1.742 \cdot 10^5$	2	3	31.5	31.2	102	54
1100	0.00834	$-1.789 \cdot 10^5$	2	7	30.8	39.7	102	66
1200	0.00407	$-1.742 \cdot 10^5$	2	8	30.9	41.9	102	69

3 Estimation of the hermitian restricted stability radius

In this section, we focus on Hermitian structured perturbations in the dissipation matrix, in particular approximation of $r^{\text{Herm}}(R; B)$ as in (1.6). The algorithms we propose exploit the following characterization of $r^{\text{Herm}}(R; B)$ established in [21, Theorem 4.9], which is originally presented in terms of an orthonormal basis $U(\lambda)$ for the kernel of $(I - BB^+)T(\lambda)$, where $T(\lambda)$ is as in (2.6) and B^+ denotes the Moore-Penrose pseudoinverse of B. However, it turns out that $U(\lambda)$ does not have to be orthonormal, rather the theorem can be stated in terms of any basis for the kernel of $(I - BB^+)T(\lambda)$. In our version, given in Theorem 3 below, we employ a particular basis that simplifies the formulas and facilitates the computation.

Theorem 3 *For an asymptotically stable DH system of the form* (1.1)*, and a restriction matrix $B \in \mathbb{C}^{n \times m}$ of full column rank, let*

1. $T(\lambda)$ *in* (2.6) *at a given $\lambda \in \mathbb{C}$ be such that $T(\lambda)$ is invertible,*
2. $L(\lambda)$ *be a lower triangular Cholesky factor of*

$$\widetilde{H}_0(\lambda) := (B^H Q T(\lambda)^{-1} B)^H (B^H Q T(\lambda)^{-1} B), \qquad (3.1)$$

   i.e., $L(\lambda)$ is a lower triangular matrix satisfying $\widetilde{H}_0(\lambda) = L(\lambda)L(\lambda)^H$,
3. $H_0(\lambda) := L(\lambda)^{-1}L(\lambda)^{-H}$,
4. $H_1(\lambda) := \text{i}(\widetilde{H}_1(\lambda) - \widetilde{H}_1(\lambda)^H)$, *with* $\widetilde{H}_1(\lambda) := L(\lambda)^{-1}(B^H Q T(\lambda)^{-1} B)^H L(\lambda)^{-H}$.

Then $r^{\text{Herm}}(R; B)$ is finite, and given by

$$r^{\text{Herm}}(R; B) = \left\{ \inf_{\omega \in \mathbb{R}} \sup_{t \in \mathbb{R}} \lambda_{\min}(H_0(\text{i}\omega) + t H_1(\text{i}\omega)) \right\}^{1/2},$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of its Hermitian matrix argument, and the inner supremum is finite if and only if $H_1(\text{i}\omega)$ is indefinite.

We first describe a numerical technique for small-scale problems in Section 3.1 and then develop a subspace framework in Section 3.2, both based on the eigenvalue optimization characterization of $r^{\text{Herm}}(R; B)$ in Theorem 3. Note that in this section, B is always assumed to be of full column rank.

3.1 An algorithm for small-scale problems

The eigenvalue optimization characterization of $r^{\text{Herm}}(R; B)$ is a min-max problem, where the inner maximization problem is concave; indeed, it can alternatively be expressed as a semidefinite program (SDP). Formally, for a given $\omega \in \mathbb{R}$, and $H_0(\text{i}\omega), H_1(\text{i}\omega)$ representing the Hermitian matrices defined in Theorem 3, let us consider

$$
\begin{aligned}
\eta^{\text{Herm}}(R; B, \text{i}\omega) &:= \sup_{t \in \mathbb{R}} \lambda_{\min}(H_0(\text{i}\omega) + t H_1(\text{i}\omega)) \\
&= \sup\{z \mid z, t \in \mathbb{R} \text{ s.t. } H_0(\text{i}\omega) + t H_1(\text{i}\omega) - zI \geq 0\} \quad (3.2)
\end{aligned}
$$

so that

$$r^{\mathrm{Herm}}(R;\,B) = \sqrt{\inf_{\omega \in \mathbb{R}} \eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)}. \tag{3.3}$$

In the following subsections, we solve the minimization problem inside the square root in (3.3).

3.1.1 Inner maximization problems

The characterization in the second line of (3.2) is a linear convex SDP. Some of the most widely used techniques to solve a linear convex SDP are different forms of interior-point methods. Implementations of some of these by the interior-point methods available through the package `cvx` [13, 14].

An alternative is to employ the software package `eigopt` [22] for the eigenvalue optimization problem in the first line of (3.2). This software package is based on an approach that approximates the eigenvalue function with a piecewise quadratic function that lies globally above the eigenvalue function. It requires a global upper bound for the second derivative of the eigenvalue function, which can be chosen as any slightly positive real number (e.g., 10^{-6}) for the concave eigenvalue function in the first line of (3.2).

In our experience, `eigopt` performs the computation of $\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)$ significantly faster than `cvx`. The only downside is that an interval containing the optimal t for (3.2) must be supplied to `eigopt`, whereas such an interval is not needed by `cvx`. In practice, we employ `eigopt` due to its efficiency.

3.1.2 Outer minimization problems

The minimum of $\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)$ with respect to $\omega \in \mathbb{R}$ yields the square of the radius $r^{\mathrm{Herm}}(R;\,B)$, and it turns out that the minimizing $\omega \in \mathbb{R}$ corresponds to the point $\mathrm{i}\omega$ that first becomes an eigenvalue on the imaginary axis under the smallest perturbation possible [21, Theorem 4.9]. This is a nonconvex optimization problem. Indeed, the objective $\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)$ may even blow up at some ω.

We again resort to `eigopt` for the minimization of $\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)$. For the sake of completeness, a formal description is provided in Algorithm 2 below, where we use the abbreviations

$$\eta_j := \eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega_j) \quad \text{and} \quad \eta_j' := \left. \frac{d}{d\omega}\, \eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega) \right|_{\omega=\omega_j}.$$

Every convergent subsequence of the sequence $\{\omega_k\}$ in Algorithm 2 is guaranteed to converge to a global minimizer of $\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega)$ provided γ is chosen sufficiently small. Moreover, by continuity, it also follows that $\sqrt{\eta^{\mathrm{Herm}}(R;\,B,\mathrm{i}\omega_k)}$ in the limit approaches $r^{\mathrm{Herm}}(R;\,B)$ for all γ values sufficiently small. In our experiments, modestly small values of γ, such as $\gamma = -20$, seem to lead to convergence.

Algorithm 2 Small-scale computation of $r^{\mathrm{Herm}}(R; B)$.

Require: Matrices $B \in \mathbb{C}^{n \times m}$, $J, R, Q \in \mathbb{C}^{n \times n}$, a negative real number γ, and a closed interval $\widetilde{\Omega} \subseteq \mathbb{R}$ that contains the global minimizer of $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ over $\omega \in \mathbb{R}$.

Ensure: Sequence $\{\eta_k\}$ s.t. $\lim_{k \to \infty} \sqrt{\eta_k} = r^{\mathrm{Herm}}(R; B)$.

1: $\omega_0 \leftarrow$ an initial point in $\widetilde{\Omega}$.
2: **for** $k = 0, 1, \dots$ **do**
3:     Compute $\eta_k := \eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega_k)$ and $\eta'_k := \frac{d}{d\omega} \eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)\big|_{\omega=\omega_k}$.
4:     $Q_k(\omega) \leftarrow \max \{ q_j(\omega) \mid j = 0, \dots, k \}$, where
$$q_j(\omega) := \eta_j + \eta'_j(\omega - \omega_j) + (\gamma/2)(\omega - \omega_j)^2.$$
5:     $\omega_{k+1} \leftarrow \arg\min_{\omega \in \widetilde{\Omega}} Q_k(\omega)$.
6: **end for**

At step k of the algorithm, $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ and its derivative need to be computed at ω_k. For instance, one can rely on one of the two approaches (`cvx` or `eigopt`) in Section 3.1.1 for the computation of $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega_k)$, and employ a finite difference formula to approximate its derivative, which would require an additional computation of $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ at an ω close to ω_k.

3.2 An interpolation-based subspace framework for large-scale problems

In this section, we consider the Hermitian restricted stability radius for large DH systems. The characterization in Theorem 3 is in terms of the matrix-valued functions $H_0(\mathrm{i}\omega)$ and $H_1(\mathrm{i}\omega)$, which are of small size provided that B has few columns. The large-scale nature of this characterization is hidden in the matrix-valued function $T(\mathrm{i}\omega) = (J - R)Q - \mathrm{i}\omega I$, because the computation of $H_0(\mathrm{i}\omega)$ and $H_1(\mathrm{i}\omega)$ requires the solution of the linear system $T(\mathrm{i}\omega_k)Z = B$.

3.2.1 Derivation of the reduced problems

To cope with the large-scale setting, we benefit from structure-preserving two-sided projections similar to those described in Section 2.2. In particular, for a given matrix V_k with orthonormal columns, we define W_k as in (2.9), J_k, R_k, Q_k as in (2.12), and B_k as in (2.13), respectively. Again, $\mathcal{V}_k, \mathcal{W}_k$ refer to the right, left subspaces spanned by the columns of V_k, W_k.

However, we no longer have a tool such as Theorem 1 to establish interpolation results, because there is no apparent transfer function, as there is indeed no apparent linear PH system that can be tied to the eigenvalue optimization characterization. The following simple observation turns out to be very useful. The subsequent arguments make use of the matrix-valued function

$$T_k(\lambda) := (J_k - R_k)Q_k - \lambda I \tag{3.4}$$

for $\lambda \in \mathbb{C}$ associated with the reduced problems.

Lemma 1 *Consider a DH model* (1.1), *and a reduced model* $\dot{x}_k = (J_k - R_k)Q_k x_k$ *with coefficients as in* (2.12) *defined in terms of a given V_k and W_k as in* (2.9). *Letting $T(\lambda)$ and $T_k(\lambda)$ be as in* (2.6) *and* (3.4), *respectively, we then have $T_k(\lambda) = W_k^H T(\lambda) V_k$ for all $\lambda \in \mathbb{C}$.*

Proof From the definitions of J_k, R_k, Q_k we obtain

$$
\begin{aligned}
T_k(\lambda) &= (W_k^H(J-R)W_k)V_k^H Q V_k - \lambda I \\
&= W_k^H(J-R)Q V_k - \lambda W_k^H V_k \\
&= W_k^H((J-R)Q - \lambda I)V_k = W_k^H T(\lambda)V_k,
\end{aligned}
$$

where we have employed (2.10) in the second equality. $\qquad\square$

In the following, we develop a subspace framework in which the Hermite interpolation properties between the original problem defined in terms of $T(\lambda)$ and the reduced problem defined in terms of $T_k(\lambda)$ hold. For this, we first show an auxiliary interpolation result between $B^H Q T(\lambda)^{-1} B$ and its reduced counterpart.

Lemma 2 *Consider a DH model* (1.1), *and a reduced model* $\dot{x}_k = (J_k - R_k)Q_k x_k$ *with coefficients as in* (2.12) *defined in terms of a given V_k and W_k as in* (2.9). *Furthermore, let $T(\lambda)$ and $T_k(\lambda)$ be as in* (2.6) *and* (3.4), *respectively. For a given $\widehat{\lambda} \in \mathbb{C}$ such that $T(\widehat{\lambda})$ and $T_k(\widehat{\lambda})$ are invertible, the following assertions hold:*

(i) *If $\mathrm{Ran}(T(\widehat{\lambda})^{-1}B) \subseteq \mathcal{V}_k$, then $B^H Q T(\widehat{\lambda})^{-1}B = B_k^H Q_k T_k(\widehat{\lambda})^{-1}B_k$.*

(ii) *Additionally, if $\mathrm{Ran}(T(\widehat{\lambda})^{-2}B) \subseteq \mathcal{V}_k$ and the orthonormal basis V_k for $\mathcal{V}_k$ is of the form $V_k = \begin{bmatrix} \widetilde{V}_k & \widehat{V}_k \end{bmatrix}$, where the columns of $\widetilde{V}_k$ form an orthonormal basis for $\mathrm{Ran}(T(\widehat{\lambda})^{-1}B)$, then $B^H Q T(\widehat{\lambda})^{-2}B = B_k^H Q_k T_k(\widehat{\lambda})^{-2}B_k$.*

Proof **(i)** If $\mathrm{Ran}(T(\widehat{\lambda})^{-1}B) \subseteq \mathcal{V}_k$, then

$$
\begin{aligned}
B^H Q T(\widehat{\lambda})^{-1}B = B^H Q V_k V_k^H T(\widehat{\lambda})^{-1}B &= B^H W_k V_k^H Q V_k V_k^H T(\widehat{\lambda})^{-1}B \\
&= B_k^H Q_k V_k^H T(\widehat{\lambda})^{-1}B,
\end{aligned}
$$
(3.5)

where the first equality is due to the fact that $V_k V_k^H$ is the orthogonal projector onto $\mathcal{V}_k$, and the second follows, since $W_k V_k^H$ is a projector onto $\mathrm{Ran}(QV_k)$.

We complete the proof by showing that $V_k^H T(\widehat{\lambda})^{-1}B = T_k(\widehat{\lambda})^{-1}B_k$. To this end, let $Z := T(\widehat{\lambda})^{-1}B$, and Z_k be such that $V_k Z_k = Z$. (There exists a unique Z_k with this property, because $\mathrm{Ran}(Z) \subseteq \mathcal{V}_k$.) Then $T(\widehat{\lambda})Z = B$ implies that $T(\widehat{\lambda})V_k Z_k = B$, and thus $W_k^H T(\widehat{\lambda})V_k Z_k = W_k^H B = B_k$. Hence, by Lemma 1, we see that $Z_k = T_k(\widehat{\lambda})^{-1}B_k$, which in turn implies

$$
V_k^H T(\widehat{\lambda})^{-1}B = V_k^H Z = V_k^H(V_k Z_k) = T_k(\widehat{\lambda})^{-1}B_k. \qquad (3.6)
$$

(ii) Following the steps at the beginning of the proof of part (i), in particular from a line of reasoning similar to that in (3.5), we have

$$
B^H Q T(\widehat{\lambda})^{-2}B = B_k^H Q_k V_k^H T(\widehat{\lambda})^{-2}B.
$$

It remains to show that $V_k^H T(\widehat{\lambda})^{-2} B = T_k(\widehat{\lambda})^{-2} B_k$. To this end, we exploit

$$V_k^H T(\widehat{\lambda})^{-2} B = (V_k^H T(\widehat{\lambda})^{-1} \widetilde{V}_k)(\widetilde{V}_k^H T(\widehat{\lambda})^{-1} B). \tag{3.7}$$

Now define $Z := T(\widehat{\lambda})^{-1} \widetilde{V}_k$ and Z_k such that $V_k Z_k = Z$ (once again such a Z_k exists uniquely, because $\mathrm{Ran}(Z) \subseteq \mathcal{V}_k$) so that $T(\widehat{\lambda}) V_k Z_k = \widetilde{V}_k$. Then, due to the property $W_k^H V_k = I_{\widetilde{k}}$, we have $W_k^H T(\widehat{\lambda}) V_k Z_k = W_k^H \widetilde{V}_k = I_{\widetilde{k},m}$, where $I_{\widetilde{k},m}$ is the matrix consisting of the first m columns of the $\widetilde{k} \times \widetilde{k}$ identity matrix $I_{\widetilde{k}}$ and $\widetilde{k} := \dim \mathcal{V}_k \geq m$. By Lemma 1, this implies that $Z_k = T_k(\widehat{\lambda})^{-1} I_{\widetilde{k},m}$, so, for the term inside the first parenthesis on the right-hand side of (3.7), we have

$$V_k^H T(\widehat{\lambda})^{-1} \widetilde{V}_k = V_k^H Z = V_k^H (V_k Z_k) = T_k(\widehat{\lambda})^{-1} I_{\widetilde{k},m}. \tag{3.8}$$

For the term inside the second parenthesis on the right-hand side of (3.7), we make use of the following observation:

$$V_k^H T(\widehat{\lambda})^{-1} B = \begin{bmatrix} \widetilde{V}_k^H \\ \widehat{V}_k^H \end{bmatrix} T(\widehat{\lambda})^{-1} B = \begin{bmatrix} \widetilde{V}_k^H T(\widehat{\lambda})^{-1} B \\ 0 \end{bmatrix}, \tag{3.9}$$

where the last equality follows, since the columns of $\widetilde{V}_k$ form an orthonormal basis for $\mathrm{Ran}(T(\widehat{\lambda})^{-1} B)$. Inserting these in (3.7), we obtain

$$\begin{aligned}
V_k^H T_k(\widehat{\lambda})^{-2} B_k &= (V_k^H T(\widehat{\lambda})^{-1} \widetilde{V}_k)(\widetilde{V}_k^H T(\widehat{\lambda})^{-1} B) \\
&= T_k(\widehat{\lambda})^{-1} I_{\widetilde{k},m} (\widetilde{V}_k^H T(\widehat{\lambda})^{-1} B) \\
&= T_k(\widehat{\lambda})^{-1} I_{\widetilde{k}} (V_k^H T(\widehat{\lambda})^{-1} B) \\
&= T_k(\widehat{\lambda})^{-1} I_{\widetilde{k}} (T_k(\widehat{\lambda})^{-1} B_k) = T_k(\widehat{\lambda})^{-2} B_k,
\end{aligned}$$

where in the second and third equality we exploit (3.8) and (3.9), respectively, and in the fourth equality we employ the equivalence in (3.6).

$\square$

Our reduced problems are expressed in terms of the reduced versions of $H_0(\lambda)$, $L(\lambda)$, and $H_1(\lambda)$ defined via

$$H_{k,0}(\lambda) := L_k(\lambda)^{-1} L_k(\lambda)^{-H}, \tag{3.10}$$

with $L_k(\lambda)$ denoting a lower triangular Cholesky factor of

$$\widetilde{H}_{k,0}(\lambda) := (B_k^H Q_k T_k(\lambda)^{-1} B_k)^H (B_k^H Q_k T_k(\lambda)^{-1} B_k), \tag{3.11}$$

and $H_{k,1}(\lambda) := \mathrm{i}(\widetilde{H}_{k,1}(\lambda) - \widetilde{H}_{k,1}(\lambda)^H)$ with

$$\widetilde{H}_{k,1}(\lambda) := L_k(\lambda)^{-1} (B_k^H Q_k T_k(\lambda)^{-1} B_k)^H L_k(\lambda)^{-H}. \tag{3.12}$$

To ensure the uniqueness of $L(\lambda)$ and $L_k(\lambda)$, we define them as the Cholesky factors of $\widetilde{H}_0(\lambda)$ and $\widetilde{H}_{k,0}(\lambda)$ with real and positive entries along the diagonal. We emphasize here that the matrix functions $H_0(\lambda)$, $L(\lambda)$, $H_1(\lambda)$ and their reduced counterparts are of the same size. However, the reduced matrix-valued functions are defined in terms of $T_k(\lambda)$, which is small if the dimensions of $\mathcal{V}_k$, $\mathcal{W}_k$ are small, in contrast to $T(\lambda)$ in the definitions of $H_0(\lambda)$, $L(\lambda)$, $H_1(\lambda)$.

Our goal is to come up with a reduced counterpart of $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$, that Hermite-interpolates the full function at prescribed points. For a given V_k, and W_k as in (2.9), we introduce

$$\eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\omega) := \sup_{t \in \mathbb{R}} \lambda_{\min}(H_{k,0}(\mathrm{i}\omega) + t H_{k,1}(\mathrm{i}\omega)).$$

Next, we establish that the quantity $\eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ is independent of the choice of the basis for the subspace $\mathcal{V}_k$. For this result, we introduce the notation $\eta_{V_k, W_k}^{\mathrm{Herm}}(R; B, \mathrm{i}\omega))$ to emphasize the particular choices of the bases V_k, W_k used in the definition of $\eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$. Similarly, we indicate the choices of the bases used in the definitions of J_k, R_k, Q_k, and B_k with the help of the notations J_{W_k}, R_{W_k}, Q_{V_k}, and B_{W_k}.

Lemma 3 *Let the columns of V_k and $\widetilde{V}_k$ form orthonormal bases for the subspace $\mathcal{V}_k$, and let $W_k := Q V_k (V_k^H Q V_k)^{-1}$, $\widetilde{W}_k := Q \widetilde{V}_k (\widetilde{V}_k^H Q \widetilde{V}_k)^{-1}$. Then*

$$\eta_{V_k, W_k}^{Herm}(R; B, \mathrm{i}\omega) = \eta_{\widetilde{V}_k, \widetilde{W}_k}^{Herm}(R; B, \mathrm{i}\omega)$$

for all $\omega \in \mathbb{R}$.

Proof The proof makes use of the characterization

$$\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega) = \inf\{\|\Delta\|_2 \mid \Delta = \Delta^H, \ \mathrm{i}\omega \in \Lambda\big((J - (R + B\Delta B^H))Q\big)\}^2 \quad (3.13)$$

shown in [21, Theorem 4.9].

Now since the columns of V_k and $\widetilde{V}_k$ form bases for the same space, there exists a unitary matrix P such that $\widetilde{V}_k = V_k P$. Furthermore,

$$\widetilde{W}_k = Q(V_k P)(P^H V_k^H Q V_k P)^{-1} = Q V_k P P^H (V_k^H Q V_k)^{-1} P = W_k P.$$

The assertion then follows from (3.13) and the following set of equivalences:

$$\begin{aligned}
\mathrm{i}\omega \in \Lambda\big((J_{W_k} - R_{W_k})Q_{V_k} - (B_{W_k}\Delta B_{W_k}^H)Q_{V_k}\big) &\iff \\
\det(W_k^H\big((J - R)Q - B\Delta B^H Q - \mathrm{i}\omega I\big)V_k) = 0 &\iff \\
\det(\widetilde{W}_k^H\big((J - R)Q - B\Delta B^H Q - \mathrm{i}\omega I\big)\widetilde{V}_k) = 0 &\iff \\
\mathrm{i}\omega \in \Lambda\big((J_{\widetilde{W}_k} - R_{\widetilde{W}_k})Q_{\widetilde{V}_k} - (B_{\widetilde{W}_k}\Delta B_{\widetilde{W}_k}^H)Q_{\widetilde{V}_k}\big).
\end{aligned}$$

Above, the first and third equivalences are due to the identities in (2.10), whereas the second equivalence is due to $\widetilde{V}_k = V_k P$ and $\widetilde{W}_k = W_k P$. $\square$

We now prove our main interpolation result.

Theorem 4 (Hermite Interpolation of Hermitian Backward Errors) *Consider a DH model (1.1), and a reduced model $\dot{x}_k = (J_k - R_k)Q_k x_k$ with coefficients as in (2.12) defined in terms of a given V_k and W_k as in (2.9). Furthermore, let $T(\lambda)$ and $T_k(\lambda)$ be as in (2.6) and (3.4), respectively. For a given $\widehat{\omega} \in \mathbb{R}$, suppose that the subspace $\mathcal{V}_k := Ran(V_k)$ is such that*

$$Ran(T(\mathrm{i}\widehat{\omega})^{-1}B) \subseteq \mathcal{V}_k, \quad Ran(T(\mathrm{i}\widehat{\omega})^{-2}B) \subseteq \mathcal{V}_k. \quad (3.14)$$

(i) *The quantity $\eta^{Herm}(R; B, \mathrm{i}\widehat{\omega})$ is finite if and only if $\eta_k^{Herm}(R; B, \mathrm{i}\widehat{\omega})$ is finite. If $\eta^{Herm}(R; B, \mathrm{i}\widehat{\omega})$ is finite, then*

$$\eta^{Herm}(R; B, \mathrm{i}\widehat{\omega}) = \eta_k^{Herm}(R; B, \mathrm{i}\widehat{\omega}). \tag{3.15}$$

(ii) *Moreover, if $\eta^{Herm}(R; B, \mathrm{i}\omega)$ and $\eta_k^{Herm}(R; B, \mathrm{i}\omega)$ are differentiable at $\widehat{\omega}$, then we have*

$$\left. \frac{d}{d\omega} \eta^{Herm}(R; B, \mathrm{i}\omega) \right|_{\omega=\widehat{\omega}} = \left. \frac{d}{d\omega} \eta_k^{Herm}(R; B, \mathrm{i}\omega) \right|_{\omega=\widehat{\omega}}. \tag{3.16}$$

Proof Without loss of generality, we may assume that the matrix V_k is such that $V_k = \begin{bmatrix} \widetilde{V}_k & \widehat{V}_k \end{bmatrix}$, with the columns of $\widetilde{V}_k$ forming an orthonormal basis for $\mathrm{Ran}(T(\mathrm{i}\widehat{\omega})^{-1}B)$. By Lemma 3, it suffices to prove the claims for such a particular choice of orthonormal basis.

(i) By the definitions of $\widetilde{H}_0(\lambda)$, $\widetilde{H}_{k,0}(\lambda)$ in (3.1), (3.11), respectively, and by part (i) of Lemma 2, we have

$$\begin{aligned} \widetilde{H}_0(\mathrm{i}\widehat{\omega}) &= (B^H Q T(\mathrm{i}\widehat{\omega})^{-1}B)^H (B^H Q T(\mathrm{i}\widehat{\omega})^{-1}B) \\ &= (B_k^H Q_k T_k(\mathrm{i}\widehat{\omega})^{-1}B_k)^H (B_k^H Q_k T_k(\mathrm{i}\widehat{\omega})^{-1}B_k) = \widetilde{H}_{k,0}(\mathrm{i}\widehat{\omega}). \end{aligned}$$

This also implies that $L(\mathrm{i}\widehat{\omega}) = L_k(\mathrm{i}\widehat{\omega})$ due to the uniqueness of the Cholesky factors of $\widetilde{H}_0(\mathrm{i}\widehat{\omega})$, $\widetilde{H}_{k,0}(\mathrm{i}\widehat{\omega})$, so $H_0(\mathrm{i}\widehat{\omega}) = L(\mathrm{i}\widehat{\omega})^{-1}L(\mathrm{i}\widehat{\omega})^{-H} = L_k(\mathrm{i}\widehat{\omega})^{-1}L_k(\mathrm{i}\widehat{\omega})^{-H} = H_{k,0}(\mathrm{i}\widehat{\omega})$. Furthermore, recalling the definitions of $\widetilde{H}_1(\lambda)$, $H_1(\lambda)$ in part 4 of Theorem 3, the definitions of $\widetilde{H}_{k,1}(\lambda)$, $H_{k,1}(\lambda)$ in (3.12) and the line above (3.12), we have

$$\begin{aligned} \widetilde{H}_1(\mathrm{i}\widehat{\omega}) &= L(\mathrm{i}\widehat{\omega})^{-1}(B^H Q T(\mathrm{i}\widehat{\omega})^{-1}B)^H L(\mathrm{i}\widehat{\omega})^{-H} \\ &= L_k(\mathrm{i}\widehat{\omega})^{-1}(B_k^H Q_k T_k(\mathrm{i}\widehat{\omega})^{-1}B_k)^H L_k(\mathrm{i}\widehat{\omega})^{-H} = \widetilde{H}_{k,1}(\mathrm{i}\widehat{\omega}) \end{aligned}$$

and $H_1(\mathrm{i}\widehat{\omega}) = H_{k,1}(\mathrm{i}\widehat{\omega})$. From the latter equality and Theorem 3, we deduce that $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\widehat{\omega})$ is finite if and only if $\eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\widehat{\omega})$ is finite. Additionally, if $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\widehat{\omega})$ is finite, then

$$\begin{aligned} \eta^{\mathrm{Herm}}(R; B, \mathrm{i}\widehat{\omega}) &= \max_{t\in\mathbb{R}} \lambda_{\min}(H_0(\mathrm{i}\widehat{\omega}) + t H_1(\mathrm{i}\widehat{\omega})) \\ &= \max_{t\in\mathbb{R}} \lambda_{\min}(H_{k,0}(\mathrm{i}\widehat{\omega}) + t H_{k,1}(\mathrm{i}\widehat{\omega})) = \eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\widehat{\omega}), \end{aligned}$$

completing the proof of (3.15).

(ii) Now let us suppose that $\eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ and $\eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\omega)$ are differentiable at $\widehat{\omega}$. To prove the interpolation property in the derivatives, we benefit from analytical expressions. It follows from the standard perturbation theory of simple eigenvalues (e.g., see [20]) that

$$\begin{aligned} \left. \frac{d}{d\omega} \eta^{\mathrm{Herm}}(R; B, \mathrm{i}\omega) \right|_{\omega=\widehat{\omega}} &= v^H \left(\left. \frac{d}{d\omega} \left(H_0(\mathrm{i}\omega) + \widehat{t} H_1(\mathrm{i}\omega) \right) \right|_{\omega=\widehat{\omega}} \right) v, \\ \left. \frac{d}{d\omega} \eta_k^{\mathrm{Herm}}(R; B, \mathrm{i}\omega) \right|_{\omega=\widehat{\omega}} &= v^H \left(\left. \frac{d}{d\omega} \left(H_{k,0}(\mathrm{i}\omega) + \widehat{t} H_{k,1}(\mathrm{i}\omega) \right) \right|_{\omega=\widehat{\omega}} \right) v, \end{aligned} \tag{3.17}$$

where $\widehat{t}$ is such that

$$\widehat{t} \in \arg\max_t \lambda_{\min}(H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega})) = \arg\max_t \lambda_{\min}(H_{k,0}(i\widehat{\omega}) + tH_{k,1}(i\widehat{\omega}))$$

and v is a unit eigenvector corresponding to $\lambda_{\min}(H_0(i\widehat{\omega}) + \widehat{t}H_1(i\widehat{\omega})) = \lambda_{\min}(H_{k,0}(i\widehat{\omega}) + \widehat{t}H_{k,1}(i\widehat{\omega}))$. Thus, it suffices to prove

$$\frac{d}{d\omega}H_0(i\omega)\Big|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}H_{k,0}(i\omega)\Big|_{\omega=\widehat{\omega}}, \qquad \frac{d}{d\omega}H_1(i\omega)\Big|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}H_{k,1}(i\omega)\Big|_{\omega=\widehat{\omega}}.$$

We first establish the interpolation properties between the derivatives of $H_0(i\omega)$ and $H_{k,0}(i\omega)$ at $\omega = \widehat{\omega}$. To this end, observe that parts (i) and (ii) of Lemma 2 imply that

$$\begin{aligned}
\frac{d}{d\omega}\widetilde{H}_0(i\omega)\Big|_{\omega=\widehat{\omega}} &= i(B^H Q T(i\widehat{\omega})^{-1} B)^H (B^H Q T(i\widehat{\omega})^{-2} B) \\
&- i(B^H Q T(i\widehat{\omega})^{-2} B)^H (B^H Q T(i\widehat{\omega})^{-1} B) \\
&= i(B_k^H Q_k T_k(i\widehat{\omega})^{-1} B_k)^H (B_k^H Q_k T_k(i\widehat{\omega})^{-2} B_k) \\
&- i(B_k^H Q_k T_k(i\widehat{\omega})^{-2} B_k)^H (B_k^H Q_k T_k(i\widehat{\omega})^{-1} B_k) = \frac{d}{d\omega}\widetilde{H}_{k,0}(i\omega)\Big|_{\omega=\widehat{\omega}}.
\end{aligned} \tag{3.18}$$

Next, for a given ω, consider the Cholesky factorizations $\widetilde{H}_0(i\omega) = L(i\omega)L(i\omega)^H$ and $\widetilde{H}_{k,0}(i\omega) = L_k(i\omega)L_k(i\omega)^H$. By differentiating these two equations and exploiting (3.18), we have]selectfont

$$\begin{aligned}
\left\{ L(i\omega)\left(\tfrac{d}{d\omega}L(i\omega)\right)^H + \left(\tfrac{d}{d\omega}L(i\omega)\right)L(i\omega)^H \right\}\Big|_{\omega=\widehat{\omega}} &= \\
\left\{ L_k(i\omega)\left(\tfrac{d}{d\omega}L_k(i\omega)\right)^H + \left(\tfrac{d}{d\omega}L_k(i\omega)\right)L_k(i\omega)^H \right\}\Big|_{\omega=\widehat{\omega}} &= \\
\left\{ L(i\omega)\left(\tfrac{d}{d\omega}L_k(i\omega)\right)^H + \left(\tfrac{d}{d\omega}L_k(i\omega)\right)L(i\omega)^H \right\}\Big|_{\omega=\widehat{\omega}}, &
\end{aligned}$$

where we have used $L(i\widehat{\omega}) = L_k(i\widehat{\omega})$, established in part (i), in the second equality. Thus, both $\frac{d}{d\omega}L(i\omega)|_{\omega=\widehat{\omega}}$ and $\frac{d}{d\omega}L_k(i\omega)|_{\omega=\widehat{\omega}}$ are lower triangular solutions of the matrix equation

$$\frac{d}{d\omega}\widetilde{H}_0(i\omega)\Big|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}\widetilde{H}_{k,0}(i\omega)\Big|_{\omega=\widehat{\omega}} = L(i\widehat{\omega})X^H + XL(i\widehat{\omega})^H,$$

which has a unique lower triangular solution with real entries along the diagonal, so $\frac{d}{d\omega}L(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}L_k(i\omega)|_{\omega=\widehat{\omega}}$. Furthermore, it follows from the definitions of $H_0(i\omega)$ and $H_{k,0}(i\omega)$ in part 3 of Theorem 3 and (3.10) that $L(i\omega)H_0(i\omega)L(i\omega)^H = I = L_k(i\omega)H_{k,0}(i\omega)L_k(i\omega)^H$, which we differentiate at $\omega = \widehat{\omega}$, and exploit $L(i\widehat{\omega}) = L_k(i\widehat{\omega})$, $\frac{d}{d\omega}L(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}L_k(i\omega)|_{\omega=\widehat{\omega}}$, $H_0(i\widehat{\omega}) = H_{k,0}(i\widehat{\omega})$ to deduce that the equality $\frac{d}{d\omega}H_0(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}H_{k,0}(i\omega)|_{\omega=\widehat{\omega}}$ holds.

Finally, by exploiting $L(i\omega)\widetilde{H}_1(i\omega)L(i\omega)^H = (B^H Q T(i\omega)^{-1}B)^H$, which follows from the definition of $H_1(i\omega)$, $H_{k,1}(i\omega)$ in part 4 of Theorem 3, we

obtain

$$
\begin{aligned}
\frac{d}{d\omega}\left(L(i\omega)\widetilde{H}_1(i\omega)L(i\omega)^H\right)\Big|_{\omega=\widehat{\omega}} &= -i(B^H Q T(i\widehat{\omega})^{-2}B)^H \\
&= -i(B_k^H Q_k T_k(i\widehat{\omega})^{-2}B_k)^H \\
&= \frac{d}{d\omega}\left(L_k(i\omega)\widetilde{H}_{k,1}(i\omega)L_k(i\omega)^H\right)\Big|_{\omega=\widehat{\omega}},
\end{aligned}
$$

where we exploit $L(i\widehat{\omega}) = L_k(i\widehat{\omega})$, $\frac{d}{d\omega}L(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}L_k(i\omega)|_{\omega=\widehat{\omega}}$, $\widetilde{H}_1(i\widehat{\omega}) = \widetilde{H}_{k,1}(i\widehat{\omega})$ to deduce $\frac{d}{d\omega}\widetilde{H}_1(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}\widetilde{H}_{k,1}(i\omega)|_{\omega=\widehat{\omega}}$. It immediately follows that $\frac{d}{d\omega}H_1(i\omega)|_{\omega=\widehat{\omega}} = \frac{d}{d\omega}H_{k,1}(i\omega)|_{\omega=\widehat{\omega}}$, and the proof of (3.16) is complete.

$\square$

Remark 1 The function $\eta^{\mathrm{Herm}}(R; B, i\omega)$ is differentiable at $\widehat{\omega}$ whenever

1. $\eta^{\mathrm{Herm}}(R; B, i\widehat{\omega})$ is finite, equivalently $H_1(i\widehat{\omega})$ is indefinite,
2. the global minimum of $\lambda_{\min}(H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega}))$ is attained at a unique t, and
3. $\lambda_{\min}(H_0(i\widehat{\omega}) + \widehat{t}H_1(i\widehat{\omega}))$ is simple, where $\widehat{t} := \arg\max_{t\in\mathbb{R}}\lambda_{\min}(H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega}))$.

These conditions guarantee also the differentiability of $\eta_k^{\mathrm{Herm}}(R; B, i\omega)$ at $\widehat{\omega}$ provided that the subspace inclusion in (3.14) holds. This latter differentiability property is due to $H_0(i\widehat{\omega}) = H_{k,0}(i\widehat{\omega})$ and $H_1(i\widehat{\omega}) = H_{k,1}(i\widehat{\omega})$ from part (i) of Theorem 4.

Additionally, when the concave function $g(t) := \lambda_{\min}(H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega}))$ attains its maximum, the maximizer is nearly always unique. The function $g(t)$ is the minimum of m real analytic functions [19, 25], each corresponding to an eigenvalue of $H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega})$. If $g(t)$ does not have a unique maximizer, then at least one of these real analytic functions must be constant and equal to $\eta^{\mathrm{Herm}}(R; B, i\widehat{\omega}) = \max_t g(t)$. Thus, a simple sufficient condition that ensures the uniqueness of the maximizer is that $H_1(i\widehat{\omega})$ has full rank, in which case all eigenvalues of $H_0(i\widehat{\omega}) + tH_1(i\widehat{\omega})$ blow up, either to ∞ or $-\infty$, as $t \to \infty$ implying that each of the real analytic functions is non-constant.

3.2.2 The subspace framework for $r^{\mathrm{Herm}}(R; B)$

The Hermite interpolation result of Theorem 4 immediately suggests the subspace framework in Algorithm 3 for the approximation of $r^{\mathrm{Herm}}(R; B)$. This resembles the structure-preserving subspace framework to estimate the non-Hermitian stability radius $r(R; B, C)$, especially in the way the subspaces $\mathcal{V}_k, \mathcal{W}_k$ are built. At every iteration, a reduced problem is solved in line 5 as discussed in Section 3.1, specifically by employing Algorithm 2. Letting $\widehat{\omega}$ be a global minimizer of the reduced problem, the subspaces are expanded so that the original function $\eta^{\mathrm{Herm}}(R; B, i\omega)$ is Hermite interpolated by its reduced counterpart at $\omega = \widehat{\omega}$.

Assuming that the sequence $\{\omega_k\}$ converges to a minimizer ω_* of the function $\eta^{\mathrm{Herm}}(R; B, i\omega)$ such that

1. $H_1(i\omega_*)$ is indefinite with full rank, and
2. $\lambda_{\min}(H_0(i\omega_*) + t_* H_1(i\omega_*))$ is simple, with $t_* := \arg\max_{t \in \mathbb{R}} \lambda_{\min}(H_0(i\omega_*) + t H_1(i\omega_*))$,

the analysis in [3] shows that the sequence $\{\omega_k\}$ converges at a superlinear rate. As discussed in Remark 1, the two conditions above ensure the differentiability of $\eta^{\mathrm{Herm}}(R; B, i\omega)$ and $\eta_k^{\mathrm{Herm}}(R; B, i\omega)$ for all large k in a neighborhood of ω_*. This differentiability property is essential for the convergence analysis in [3].

Algorithm 3 Subspace method for large-scale estimation of the hermitian stability radius $r^{\mathrm{Herm}}(R; B)$.

Require: Matrices $B \in \mathbb{C}^{n \times m}$, $J, R, Q \in \mathbb{C}^{n \times n}$.
Ensure: Sequence $\{\eta_k\}$ s.t. $\sqrt{\eta_k}$ for sufficiently large k is an estimate for $r^{\mathrm{Herm}}(R; B)$.
1: Choose the initial interpolation points $\omega_1, \ldots, \omega_q \in \mathbb{R}$.
2: $V_q \leftarrow \mathrm{orth}\left(\left[\, T(i\omega_1)^{-1}B \;\; T(i\omega_1)^{-2}B \;\; \ldots \;\; T(i\omega_q)^{-1}B \;\; T(i\omega_q)^{-2}B \,\right]\right)$.
3: $W_q \leftarrow QV_q(V_q^H Q V_q)^{-1}$.
4: **for** $k = q, q+1, \ldots$ **do**
5:     $\omega_{k+1} \leftarrow$ a global minimizer of $\eta_k^{\mathrm{Herm}}(R; B, i\omega)$ over $\omega \in \mathbb{R}$.
6:     $\widehat{V}_{k+1} \leftarrow \left[\, T(i\omega_{k+1})^{-1}B \;\; T(i\omega_{k+1})^{-2}B \,\right]$.
7:     $V_{k+1} \leftarrow \mathrm{orth}\left(\left[\, V_k \;\; \widehat{V}_{k+1} \,\right]\right)$.
8:     $W_{k+1} \leftarrow QV_{k+1}(V_{k+1}^H Q V_{k+1})^{-1}$.
9:     $\eta_{k+1} \leftarrow \eta_{k+1}^{\mathrm{Herm}}(R; B, i\omega_{k+1})$.
10: **end for**

3.3 Numerical experiments

In this section, we present numerical results for the estimation of the Hermitian stability radius. Most of the examples are large-scale, and, hence, involve the application of the proposed subspace framework (Algorithm 3). However, recall that Algorithm 3 uses Algorithm 2 for solving reduced subproblems. For Algorithm 3, we use the same stopping criteria as in Section 2.3, but now $f_k := \min_{\omega \in \mathbb{R}} \eta_k^{\mathrm{Herm}}(R; B, i\omega)$, and we set $k_{\max} = 80$ and $\varepsilon = 10^{-6}$. The larger value for ε is due to the fact that we compute the derivatives of $\eta_k^{\mathrm{Herm}}(R; B, i\omega)$ with respect to ω by means of forward differences, which yield approximate values accurate up to about half of the double precision.

We form the initial subspaces nearly as described in Section 2.3 for the non-Hermitian stability radius. The only difference is that, in part (iii), the points $z_1, \ldots, z_\rho$ are permuted into $z_{j_1}, \ldots, z_{j_\rho}$ so that $\eta^{\mathrm{Herm}}(R; B, i\Im z_{j_1}) \le \cdots \le \eta^{\mathrm{Herm}}(R; B, i\Im z_{j_\rho})$. For the synthetic examples in Section 3.3.1, the parameter values $l = -200$, $u = 0$, $\rho = \kappa = 5$ are used, and $z_1, \ldots, z_\rho$ are set equal to the eigenvalues of $(J - R)Q$ closest to $i\omega_1, \ldots, i\omega_\rho$. As for the disk brake example in Section 3.3.2, we use $l = -2 \cdot 10^5$, $u = -1.7 \cdot 10^5$, $\rho = \kappa = 15$, as well as $z_j = i\omega_j$ for $j = 1, \ldots, \rho$.
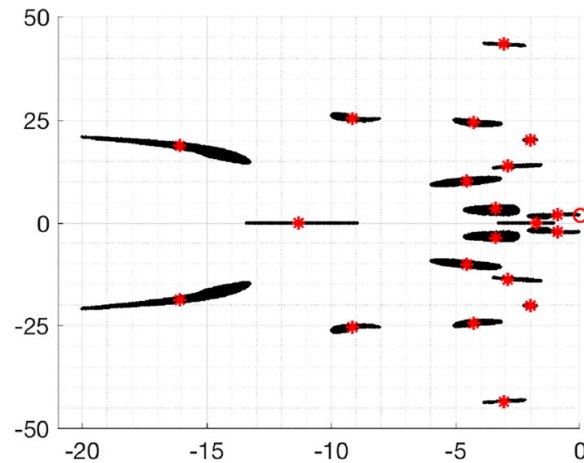
Fig. 2 A DH system of order 20 with random system matrices. Eigenvalues of $(J - R)Q$ are marked with red asterisks, whereas the black regions represent the eigenvalues of all matrices of the form $(J - (R + B\Delta B^H))Q$ for 100000 randomly chosen real symmetric Δ matrices with $\|\Delta\|_2 = r^{\text{Herm}}(R; B)$. The red circle marks $1.9794i$, where 1.9794 is the computed global minimizer of $\eta^{\text{Herm}}(R; B, i\omega)$ over $\omega \in \mathbb{R}$.

3.3.1 Synthetic examples

We first present a numerical result for a small dense random example, where $J, R, Q \in \mathbb{R}^{20 \times 20}$ and $B \in \mathbb{R}^{20 \times 2}$. Application of Algorithm 2 to this example yields $r^{\text{Herm}}(R; B) = 0.0501$, and the point that is first reached on the imaginary axis under the smallest perturbation is $1.9794i$. Figure 2 illustrates that there exist real symmetric matrices Δ with $\|\Delta\|_2 = 0.0501$ such that $1.9794i$ is contained (nearly) in the spectrum of $(J - (R + B\Delta B^H))Q$.

Next we experiment with larger dense random matrices, specifically with random $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 2}$. The results of Algorithm 3 are reported in Tables 4 and 5. For $n = 1000$, as shown in Table 4, the direct application of Algorithm 2 and the subspace framework (Algorithm 3) return exactly the same values for $r^{\text{Herm}}(R; B)$ up to the prescribed tolerances, yet the subspace framework already needs less computing time than Algorithm 2. Listed in this table is also the unstructured distance to instability $\Gamma((J - R)Q)$ defined in Section 2.3.1. It appears that $\Gamma((J - R)Q)$ overestimates $r^{\text{Herm}}(R; B)$ considerably.

Table 4 Performance of Algorithm 3 to estimate $r^{\text{Herm}}(R; B)$ on dense random DH systems of order 1000. The results from Algorithm 2, the non-Hermitian stability radius $r(R; B, B^H)$ and the unstructured distance to instability $\Gamma((J - R)Q)$ are also included. The fifth column contains the number of subspace iterations, the sixth the run-times (in seconds)

#	$r^{\text{Herm}}(R; B)$ Alg. 2 & 3	$r(R; B, B^H)$	$\Gamma((J - R)Q)$	# iter. Alg. 3	Run-time Alg. 2	Alg. 3
1	0.012922	0.011263	2.04	7	2090.0	463.5
2	0.010110	0.010104	0.36	1	1544.1	91.8
3	0.014067	0.012826	3.63	5	2840.2	335.6
4	0.009605	0.007189	0.94	11	2926.5	739.6

Table 5 Performance of Algorithm 3 to estimate $r^{\text{Herm}}(R; B)$ on dense random DH systems of order $n = 2000$ and $n = 4000$. The fifth column contains the number of subspace iterations, and the sixth the run-times (in seconds)

n	#	$r^{\text{Herm}}(R; B)$ Alg. 3	$r(R; B, B^H)$	# iter. Alg. 3	Run-time Alg. 3
2000	1	0.010217	0.005231	6	408.6
2000	2	0.003835	0.003754	13	811.8
2000	3	0.012456	0.010876	5	369.2
2000	4	0.010096	0.010031	3	221.8
4000	1	0.008982	0.008705	1	213.1
4000	2	0.008376	0.006803	3	343.2
4000	3	0.010400	0.008646	2	267.8
4000	4	0.002068	0.002055	11	828.9

For larger systems, Algorithm 2 becomes computationally too expensive, so we do not report the results of Algorithm 2. For the same reason, we also do not report $\Gamma((J - R)Q)$ for larger systems. Rather, we report only the results of Algorithm 3 for $n = 2000$ and $n = 4000$ in Table 5, as well as the non-Hermitian stability radii. We remark that in these experiments the number of subspace iterations to reach the prescribed accuracy is usually small and seems independent of n.

All of these examples involve optimization of highly nonconvex functions. Figure 3 depicts $\eta^{\text{Herm}}(R; B, i\omega)$ as a function of ω with the solid curve for the first example in Table 4 with $n = 1000$. The same figure also depicts the reduced function $\eta_k^{\text{Herm}}(R; B, i\omega)$ with the dashed curve for the same example at termination after 7 subspace iterations. Even though the reduced function $\eta_k^{\text{Herm}}(R; B, i\omega)$ uses projected matrices onto 48 dimensional subspaces, it captures the original function $\eta^{\text{Herm}}(R; B, i\omega)$ remarkably well near the global minimizer $\omega_* = -70.9581$.

Fig. 3 Application of Algorithm 3 to estimate $r^{\text{Herm}}(R; B)$ on a dense random DH system of order 1000. The full and reduced functions $\eta^{\text{Herm}}(R; B, i\omega)$ and $\eta_k^{\text{Herm}}(R; B, i\omega)$ at termination after 7 subspace iterations are plotted with the solid and dashed curves, respectively. The point $(\omega_*, \eta^{\text{Herm}}(R; B, i\omega_*))$ with ω_* denoting the global minimizer of $\eta^{\text{Herm}}(R; B, i\omega)$ is marked with a circle
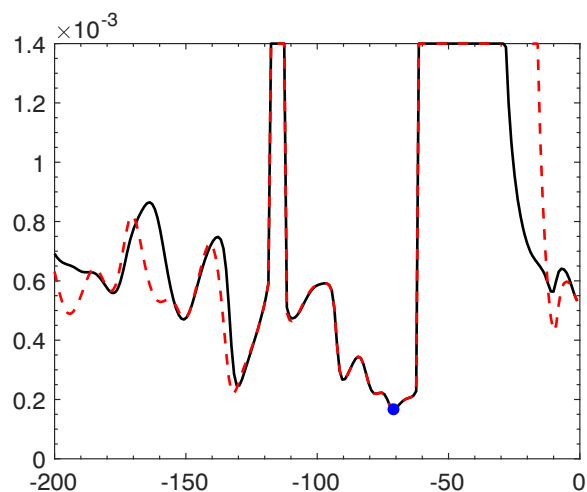
Table 6 Estimated values for the Hermitian stability radii $r^{\text{Herm}}(R; B)$ by Algorithm 3 in the second column for the FE model of a disk brake of order 9338 for a few values of Ω. The column ω_* displays the computed global minimizer of $\eta^{\text{Herm}}(R; B, i\omega)$ by the algorithm, whereas the last columns depicts the total run-time (in seconds) for Algorithm 3

Ω	$r^{\text{Herm}}(R; B)$	$r(R; B, B^T)$	ω_*	Run-time
5	0.01038	0.00893	-1.938×10^5	227.4
50	0.01001	0.00904	-1.938×10^5	222.3
100	0.00988	0.00914	-1.938×10^5	223.9
1100	0.00837	0.00834	-1.789×10^5	228.5
1200	0.00408	0.00407	-1.742×10^5	224.1

3.3.2 FE model of a disk brake

We also applied our implementation of Algorithm 3 to the FE model of the disk brake described in Section 2.3.2. The estimated values for the Hermitian structured radius $r^{\text{Herm}}(R; B)$ along with the associated minimizer ω_* of $\eta^{\text{Herm}}(R; B, i\omega)$, as well as the run-time (in seconds) are listed in Table 6 for a few values of Ω. Only one subspace iteration is performed in these examples and the subspace dimension at termination is 96. It is worth comparing the values of $r^{\text{Herm}}(R; B)$ in this table with those for the non-Hermitian stability radius $r(R; B, B^T)$, also included in the table. As expected, the listed values for the Hermitian structured stability radii are larger. The estimates for the two stability radii differ by significant amounts for smaller values of Ω, but only slightly at larger Ω.

The results show that the FE model of the disk brake is close to instability for all Ω. They indicate that further damping via shims would be helpful to avoid the instabilities due to uncertainties.

4 Concluding remarks

We have proposed subspace frameworks to estimate the stability radii for large-scale dissipative Hamiltonian systems. At every iteration, we apply DH structure-preserving projections to small subspaces, then compute the corresponding stability radii for the resulting reduced system by exploiting its eigenvalue optimization characterization. We expand the subspaces used in the projections so that Hermite interpolation properties between the objective eigenvalue function of the full and the reduced problems are attained at the optimizer of the reduced problem.

The proposed frameworks seem to converge to local optimizers, but a formal proof of this observation is open at the moment. They do converge at a superlinear rate in theory under the local convergence assumption. With the purpose of improving the likelihood of convergence to a global maximizer, we have initiated the subspaces to attain Hermite interpolation at several points on the imaginary axis between the full and initial reduced problems.

MATLAB implementations of the proposed algorithms and subspace frameworks, as well as some of the data are made publicly available.[2]

A related research direction that is currently investigated is the maximization of the stability radii when J, R, Q depend on parameters, e.g., the FE model of the disk brake depends on the rotation speed Ω.

References

1. Akay, A.: Acoustics of friction. J. Acoust. Soc. Am. **111**(4), 1525–1548 (2002)
2. Aliyev, A., Benner, P., Mengi, E., Schwerdtner, P., Voigt, M.: Large-scale computation of $\mathcal{L}_\infty$-norms by a greedy subspace method. SIAM J. Matrix Anal. Appl. **38**(4), 1496–1516 (2017)
3. Aliyev, A., Benner, P., Mengi, E., Voigt, M.: A subspace framework for $\mathcal{H}_\infty$-norm minimization. arXiv:1508.04214v2 [math.NA]. Submitted to SIAM J. Matrix Anal. Appl (2019)
4. Aliyev, N., Mehrmann, V., Mengi, E.: Computation of stability radii for large-scale dissipative hamiltonian systems. Preprint 2018-09, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG. http://www.math.tu-berlin.de/menue/forschung/veroeffentlichungen/preprints_2018 (2018)
5. Benner, P., Lowe, R., Voigt, M.: $\mathcal{L}_\infty$ norm computation for large-scale descriptor systems using structured iterative eigensolvers. Numerical Algebra Control and Optimization **8**(1), 119–133 (2018)
6. Benner, P., Mitchell, T.: Faster and more accurate computation of the H_∞ norm via optimization. SIAM J. Sci. Comput. **40**(5), A3609–A3635 (2018)
7. Benner, P., Voigt, M.: A structured pseudospectral method for $\mathcal{H}_\infty$-norm computation of large-scale descriptor systems. Math. Control Signals Syst. **26**(2), 303–338 (2014)
8. Boyd, S., Balakrishnan, V.: A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. Syst. Control Lett. **15**(1), 1–7 (1990)
9. Bruinsma, N.A., Steinbuch, M.: A fast algorithm to compute the H_∞-norm of a transfer function matrix. Syst. Control Lett. **14**(4), 287–293 (1990)
10. Freitag, M.A., Spence, A., Van Dooren, P.: Calculating the H_∞-norm using the implicit determinant method. Linear Algebra Appl. **35**(2), 619–635 (2014)
11. Gallivan, K., Vandendorpe, A., Van Dooren, P.: Model reduction of mIMO systems via tangential interpolation. SIAM J. Matrix Anal. Appl. **26**(2), 328–349 (2005)
12. Gräbner, N., Mehrmann, V., Quraishi, S., Schröder, C., von Wagner, U.: Numerical methods for parametric model reduction in the simulation of disc brake squeal. Z. Angew. Math. Mech. **96**, 1388–1405 (2016). https://doi.org/10.1002/zamm.201500217

[2] http://home.ku.edu.tr/~emengi/software/DH-stabradii.html

13. Grant, M., Boyd, S.: Graph implementations for nonsmooth convex programs. In: Blondel, V., Boyd, S., Kimura, H. (eds.) Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited (2008)

14. Grant, M., Boyd, S.: CVX: Matlab Software for disciplined convex programming, version 2.1 http://cvxr.com/cvx (2014)

15. Gugercin, S., Polyuga, R.V., Beattie, C., van der Schaft, A.J.: Interpolation-based model reduction for port-Hamiltonian systems. In: Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009, pp. 5362–5369 (2009)

16. Gugercin, S., Polyuga, R.V., Beattie, C., van der Schaft, A.J.: Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems. Automatica **48**(9), 1963–1974 (2012)

17. Guglielmi, N., Gürbüzbalaban, M., Overton, M.L.: Fast approximation of the H_∞-norm via optimization over spectral value sets. SIAM J. Matrix Anal. Appl. **34**(2), 709–737 (2013)

18. Jacob, B., Zwart, H.: Linear Port-Hamiltonian Systems on Infinite-Dimensional Spaces Operator Theory: Advances and Applications, vol. 223. Birkhäuser/Springer Basel AG, Basel (2012)

19. Kato, T.: Perturbation Theory for Linear Operators. Springer, Berlin (1995)

20. Lancaster, P.: On eigenvalues of matrices dependent on a parameter. Numer. Math. **6**, 377–387 (1964)

21. Mehl, C., Mehrmann, V., Sharma, P.: Stability radii for linear Hamiltonian systems with dissipation under structure-preserving perturbations. SIAM J. Matrix Anal. Appl. **37**(4), 1625–1654 (2016). https://doi.org/10.1137/16M1067330

22. Mengi, E., Yildirim, E.A., Kiliç, M.: Numerical optimization of eigenvalues of Hermitian matrix functions. SIAM J. Matrix Anal. Appl. **35**(2), 699–724 (2014). https://doi.org/10.1137/130933472

23. Mitchell, T., Overton, M.L.: Hybrid expansion-contraction: a robust scaleable method for approximating the H_∞ norm. IMA J. Numer. Anal. **36**(3), 985–1014 (2016)

24. Polyuga, R.V., van der Schaft, A.J.: Structure preserving moment matching for port-Hamiltonian systems: Arnoldi and Lanczos. IEEE Trans. Autom. Cont. **56**(6), 1458–1462 (2011)

25. Rellich, F.: Perturbation Theory of Eigenvalue Problems. Notes on Mathematics and its Applications. Gordon and Breach, New York (1969)

26. van der Schaft, A.J., Jeltsema, D.: Port-Hamiltonian systems theory: an introductory overview. Found Trends Syst Control **1**(2-3), 173–378 (2014)

27. Van Loan, C.F.: How near is a stable matrix to an unstable matrix? Contemporary Math. **47**, 465–478 (1985)

Affiliations

Nicat Aliyev[1] · Volker Mehrmann[2] · Emre Mengi[3]

✉ Nicat Aliyev
naliyev@ku.edu.tr

Volker Mehrmann
mehrmann@math.tu-berlin.de

Emre Mengi
emengi@ku.edu.tr

[1]   Institute of Mathematics and Mechanics, Azerbaijan National Academy of Sciences, B.Vahabzade 9, AZ1141, Baku, Azerbaijan

[2]   Institut fur Mathematik, Technische Universität Berlin, MA 4-5. Str. des 17 Juni 136, 0623, Berlin, Germany

[3]   Department of Mathematics, Koç University, Rumeli Feneri Yolu 34450, Sarıyer, Istanbul, Turkey