Electronic Transactions on Numerical Analysis.
Volume 55, pp. 508–531, 2022.

ISSN 1068–9613.
DOI: 10.1553/etna_vol55s508

ETNA
Kent State University and
Johann Radon Institute (RICAM)

COMPUTATION OF THE NEAREST STRUCTURED MATRIX TRIPLET WITH COMMON NULL SPACE*

NICOLA GUGLIELMI† AND VOLKER MEHRMANN‡

Abstract. We study computational methods for computing the distance to singularity, the distance to the nearest high-index problem, and the distance to instability for linear differential-algebraic systems (DAEs) with dissipative Hamiltonian structure. While for general unstructured DAEs the characterization of these distances is very difficult and partially open, it has been shown in [C. Mehl, V. Mehrmann, and M. Wojtylak, Distance problems for dissipative Hamiltonian systems and related matrix polynomials, Linear Algebra Appl., 623 (2021), pp. 335–366] that for dissipative Hamiltonian systems and related matrix pencils there exist explicit characterizations. We will use these characterizations for the development of computational methods to approximate these distances via methods that follow the flow of a differential equation converging to the smallest perturbation that destroys the property of regularity, index one, or stability.

Key words. dissipative Hamiltonian systems, structured distance to singularity, structured distance to high-index problem, structured distance to instability, low-rank perturbation, differential-algebraic system

AMS subject classifications. 15A18, 15A21, 65K05, 15A22

1. Introduction.

We derive computational methods for determining several distance problems, in particular the distance to singularity, the distance to the nearest high-index problem, and the distance to instability for *linear, time-invariant differential-algebraic systems (DAEs)* with dissipative Hamiltonian (dH) structure. Such dHDAE systems have the form

$$E\dot{x} = (J - R)x + f, \tag{1.1}$$

with constant matrices $E, J, R \in \mathbb{R}^{n\times n}$, where $J = -J^\top$, both $E = E^\top$ and $R = R^\top$ are positive semidefinite (denoted as $E, R \geq 0$), a differentiable state function $x : \mathbb{R} \to \mathbb{R}^n$, and a right-hand side $f : \mathbb{R} \to \mathbb{R}^n$. Here $\dot{x}$ denotes the time derivative of the function x, and $\mathbb{R}^{n\times n}$ denotes the set of real $n \times n$ matrices. The matrix E in (1.1) is associated with the Hessian of the related *Hamiltonian* energy function, which in the case discussed here has the form $\mathcal{H}(x) = \frac{1}{2}x^\top Ex$ and describes the distribution of internal energy in the system, the dissipation matrix R is associated with the loss of energy in the system, and the structure matrix $J = -J^T$ describes the energy flux among the energy storage elements in the system. It is well-known [5, 29] that dHDAEs satisfy a *dissipation inequality* $\mathcal{H}\big(x(t_1)\big) - \mathcal{H}\big(x(t_0)\big) \leq 0$ for $t_1 \geq t_0$.

Systems as in (1.1) arise in all areas of science and engineering [4, 5, 12, 21, 29, 36] as linearization (along stationary solutions), as space discretization, or approximation of physical systems and also in the context of the more general *port-Hamiltonian differential-algebraic equation* systems, which incorporate also inputs and outputs.

*Received May 4, 2021. Accepted April 7, 2022. Published online on May 31, 2022. Recommended by Melina Freitag.
†Gran Sasso Science Institute, Section of Mathematics, Via Crispi 7, I-67010 L' Aquila, Italy (nicola.guglielmi@gssi.it).
‡Institut für Mathematik, MA 4-5, TU Berlin, Str. der 17 Juni 136, D–10623 Berlin, Germany (mehrmann@math.tu-berlin.de).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

EXAMPLE 1.1. Consider as a simple example a linear electrical circuit presented in [29]. This system can be represented as a differential-algebraic system

$$\begin{aligned} L\dot{I} &= -R_L I + V_2 - V_1, \\ C_1\dot{V}_1 &= I - I_G, \\ C_2\dot{V}_2 &= -I - I_R, \\ 0 &= -R_G I_G + V_1 + E_G, \\ 0 &= -R_R I_R + V_2. \end{aligned} \tag{1.2}$$

Here $R_G, R_L, R_R > 0$ represent resistances, $L > 0$ is an inductor, $C_1, C_2 > 0$ are capacitors, and E_G is a controlled voltage source. The system is a dHDAE of the form (1.1) with $x = \begin{bmatrix} I, & V_1, & V_2, & I_G, & I_R \end{bmatrix}^\top$, $E = \operatorname{diag}(L, C_1, C_2, 0, 0)$, $R = \operatorname{diag}(R_L, 0, 0, R_G, R_R)$, and $f = \begin{bmatrix} 0, & 0, & 0, & R_G, & 0 \end{bmatrix}^\top$, where

$$J = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The energy in the system (stored in the inductor and in the two capacitors) leads to the quadratic Hamiltonian

$$\mathcal{H}(I, V_1, V_2) = \frac{1}{2}LI^2 + \frac{1}{2}C_1V_1^2 + \frac{1}{2}C_2V_2^2.$$

This will be an illustrative example for our methodology in Section 7.1. We will discuss the properties of this system and similarly constructed circuit models below.

Model descriptions typically have uncertainties arising from modeling, discretization, or measurement errors. It is therefore important to know whether the model is close (under small perturbations) to a model where some of its nice properties do not hold any longer, and this has been an important research topic recently; see, e.g., [1, 3, 6, 15, 16, 17, 24, 25, 27, 28, 29, 30, 33, 36, 34, 35].

The properties of the system (1.1) are obtained by investigating the spectral properties of the corresponding *dH matrix pencil*

$$L(\lambda) := \lambda E - A := \lambda E - (J - R). \tag{1.3}$$

Well-formulated dHDAE systems of the form (1.1) have many nice properties. To have unique solvability for general right-hand sides f, the pencil $L(\lambda)$ has to be *regular*, i.e., $\det L(\lambda)$ is not identically zero. The system is *stable*, i.e., the finite eigenvalues of $\lambda E - A$ are in the closed left-half complex plane and the eigenvalues on the imaginary axis are semisimple, or in other words have no Jordan blocks of size larger than 1, and the infinite eigenvalues have *index at most two*, i.e., Jordan blocks of size at most 2; see [27, 28]. It should, however, be noted that from a perturbation point of view it would be much better if the system would be asymptotically stable, i.e., all finite eigenvalues are in the open left-half complex plane and the index would be at most one, because then these properties stay invariant under small enough perturbations of the system matrices. It would furthermore be important to characterize how far away (in some norm of the perturbations) the system is from a system where one of these properties is lost.

To answer these questions, this paper is concerned with three nearness problems for the pencil (1.3). The first problem is the *distance to the nearest singular pencil*, i.e., to a pencil $\lambda\tilde{E}-\tilde{A}$ with $\det(\lambda\tilde{E}-\tilde{A})$ identically zero, the *distance to the nearest high-index pencil*, i.e., to a pencil $\lambda\tilde{E}-\tilde{A}$ with Jordan blocks associated to the eigenvalue ∞ of size larger than one, and the *distance to the nearest pencil with eigenvalues on the boundary of the unstable region*, i.e., to a pencil $\lambda\tilde{E}-\tilde{A}$ with purely imaginary eigenvalues. Computing these distances is very difficult for general linear systems (see, for example, [7, 8, 9, 15, 16, 17, 18, 26]). However, if one restricts the perturbations to be structured, i.e., if one considers *structured distances* within the class of dH pencils, then the situation changes completely (see [15, 16, 17, 27, 28]), and one obtains very elegant characterizations that can be used to approximate these distances via (non-convex) optimization approaches.

In contrast to classical optimization approaches, we derive computational methods to approximate these structured distances by following the flow of a differential equation. This approach has been shown to be extremely effective for computing the distance to the nearest singular pencil for general matrix pencils [18], and we will show that this holds even more so in the structured case. The presented procedure is a two-level procedure that minimizes a suitable functional over the set of systems having distance ε from the original system until the functional is annihilated. The parameter distance ε is tuned in the outer iteration, while minimization for a given ε is performed in the inner iteration.

Neither the methods based on non-convex optimization nor the methods based on following a flow are really feasible for large-scale problems, which means problems with tens or hundreds of thousands of unknowns.

To treat the large sparse case, they have to be combined with projections on the sparsity structure and model reduction methods (see [2, 3]), which intertwine the optimization step with model reduction. Here we discuss only the small-scale case, but the combination with interpolation methods can be carried out in an analogous way as in [3].

The paper is organized as follows. In Section 2 we recall a few basic results about linear time-invariant dHDAE systems. In Section 3 we discuss optimization methods that are based on gradient flow computations. Since the cases of even and odd dimension are substantially different, in Section 4 we specialize these methods for the optimization problems associated with the three discussed distance problems for the case when the state dimension is odd, while in Section 6 we discuss the case when the state dimension is even.

It turns out that the minimal distance perturbations are rank-two matrices, so in Section 5, for the odd size case, we discuss the special situation where we restrict the perturbation to be at most of rank two. In Section 7 we briefly discuss the iterative procedure for computing the optimal ε in the upper level of the two level procedure. In all cases, we present numerical examples.

2. Preliminaries. We use the following notation. The set of symmetric (positive semidefinite) matrices in $\mathbb{R}^{n\times n}$ is denoted by Sym^n ($\mathrm{Sym}^n_{\geq 0}$), and the skew-symmetric matrices in $\mathbb{R}^{n\times n}$ by Skew^n. Analogously, we denote the symmetric part of a real matrix A by $\mathrm{Sym}(A)$ and the skew-symmetric part by $\mathrm{Skew}(A)$. By $\|X\|_F$ we denote the Frobenius norm of a (possibly rectangular) matrix X; we extend this norm to matrix tuples $\mathcal{X}=(X_0,\ldots,X_k)$ via $\|\mathcal{X}\|_F=\|[X_0,\ldots,X_k]\|_F$. For $A,B\in\mathbb{C}^{n\times n}$, we denote by

$$\langle A,B\rangle=\mathrm{tr}(B^HA)$$

the Frobenius inner product on $\mathbb{C}^{n\times n}$, where B^H is the conjugate transpose of B. The Euclidian norm in $\mathbb{R}^n$ is denoted by $\|\ \|$. By $\lambda_{\min}(X)$ we denote the smallest eigenvalue of $X\in\mathrm{Sym}^n_{\geq 0}$. The real and imaginary part of a complex matrix $A\in\mathbb{C}^{n\times n}$ are denoted by $\mathrm{Re}(A)$, $\mathrm{Im}(A)$, respectively.

To characterize the properties of dHDAEs of the form (1.1), we exploit the *Kronecker canonical form* of the associated matrix pencil (1.3); see [14]. If $\mathcal{J}_n(\lambda_0)$ denotes the standard upper-triangular Jordan block of size $n \times n$ associated with an eigenvalue λ_0 and $\mathcal{L}_n$ denotes the standard right Kronecker block of size $n \times (n+1)$, i.e.,

$$\mathcal{L}_n = \lambda_0 \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix},$$

then, for $E, A \in \mathbb{C}^{n\times m}$ there exist nonsingular matrices $S \in \mathbb{C}^{n\times n}$ and $T \in \mathbb{C}^{m\times m}$ that transform the pencil to *Kronecker canonical form*,

$$S(\lambda E - A)T = \operatorname{diag}(\mathcal{L}_{\iota_1}, \dots, \mathcal{L}_{\iota_p}, \mathcal{L}_{\eta_1}^\top, \dots, \mathcal{L}_{\eta_q}^\top, \mathcal{J}_{\rho_1}^{\lambda_1}, \dots, \mathcal{J}_{\rho_r}^{\lambda_r}, \mathcal{N}_{\sigma_1}, \dots, \mathcal{N}_{\sigma_s}),$$

where $p, q, r, s, \iota_1, \dots, \iota_p, \eta_1, \dots, \eta_q, \rho_1, \dots, \rho_r, \sigma_1, \dots, \sigma_s \in \mathbb{N} \cup \{0\}$ and $\lambda_1, \dots, \lambda_r \in \mathbb{C}$, as well as $\mathcal{J}_{\rho_i}^{\lambda_i} = I_{\rho_i} - \mathcal{J}_{\rho_i}(\lambda_i)$, for $i = 1, \dots, r$, and $\mathcal{N}_{\sigma_j} = \mathcal{J}_{\sigma_j}(0) - I_{\sigma_j}$, for $j = 1, \dots, s$.

For real matrices and real transformation matrices S, T, the blocks $\mathcal{J}_{\rho_j}^{\lambda_j}$ with $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ are in *real Jordan canonical form* associated to the corresponding pair of complex conjugate eigenvalues; the other blocks are the same. A real or complex eigenvalue is called *semisimple* if the largest associated Jordan block in the complex Jordan form has size one, and the sizes η_j and ϵ_i are called the *left and right minimal indices* of $\lambda E - A$, respectively. A pencil $\lambda E - A$, is called *regular* if $n = m$ and $\det(\lambda_0 E - A) \neq 0$ for some $\lambda_0 \in \mathbb{C}$, otherwise it is called *singular*; $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ are called the finite eigenvalues of $\lambda E - A$, and $\lambda_0 = \infty$ is an eigenvalue of $\lambda E - A$ if zero is an eigenvalue of $\lambda A - E$. The size of the largest block $\mathcal{N}_{\sigma_j}$ is called the *index* ν of the pencil $\lambda E - A$.

The definition of stability for differential-algebraic systems varies in the literature. We call a pencil $\lambda E - A$ *Lyapunov stable (asymptotically stable)* if it is regular, all finite eigenvalues are in the closed (open) left-half plane, and the ones lying on the imaginary axis (including ∞) are semisimple [11]. Note that pencils with eigenvalues on the imaginary axis or at ∞ are on the boundary of the set of asymptotically stable systems, and those with multiple but semisimple, purely imaginary eigenvalues (including ∞) lie on the boundary of the set of Lyapunov stable pencils.

The following theorem summarizes some results of [27, 28] for real dH pencils; note that some of the results also hold in the complex case.

THEOREM 2.1.
Let $E, R \in \mathrm{Sym}^n_{\geq 0}$ and $J \in \mathrm{Skew}^n$.
Then, the following statements hold for the pencil $L(\lambda) = \lambda E - J + R$.

(i) *If $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $L(\lambda)$ then $\mathrm{Re}(\lambda_0) \leq 0$.*
(ii) *If $\omega \in \mathbb{R}$ and $\lambda_0 = i\omega$ is an eigenvalue of $L(\lambda)$, then λ_0 is semisimple. Moreover, if the columns of $V \in \mathbb{C}^{m\times k}$ form a basis of a regular deflating subspace of $L(\lambda)$ associated with λ_0, then $RQV = 0$.*
(iii) *The index of $L(\lambda)$ is at most two.*
(iv) *All right and left minimal indices of $L(\lambda)$ are zero (if there are any).*
(v) *The pencil $L(\lambda)$ is singular if and only if $\ker J \cap \ker E \cap \ker R \neq \emptyset$.*

Based on Theorem 2.1 in [28], the following distance problems were introduced for dH pencils.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

DEFINITION 2.2. *Let $\mathcal{L}$ denote the class of square $n \times n$ real matrix pencils of the form* (1.3). *Then*

1. *the* structured distance to singularity *is defined as*

$$d^{\mathcal{L}}_{\text{sing}}(L(\lambda)) := \inf\left\{\left\|\Delta_L(\lambda)\right\|_F \;\middle|\; L(\lambda)+\Delta_L(\lambda)\in\mathcal{L} \textit{ and is singular}\right\};$$

2. *the* structured distance to the nearest high-index problem *is defined as*

$$d^{\mathcal{L}}_{\text{hi}}(L(\lambda)) := \inf\left\{\left\|\Delta_L(\lambda)\right\|_F \;\middle|\; L(\lambda)+\Delta_L(\lambda)\in\mathcal{L} \textit{ and is of index} \geq 2\right\};$$

3. *the* structured distance to instability *is defined as*

$$d^{\mathcal{L}}_{\text{inst}}(L(\lambda)) = 0 \qquad \textit{if } L(\lambda) \quad \textit{is unstable,}$$

and otherwise by

$$d^{\mathcal{L}}_{\text{inst}}(L(\lambda)) := \inf\left\{\left\|\Delta_L(\lambda)\right\|_F \;\middle|\; L(\lambda)+\Delta_L(\lambda)\in\mathcal{L} \textit{ has purely imaginary evs.}\right\}.$$

Here $\Delta_L(\lambda) = \lambda\Delta_E - \Delta_J + \Delta_R$, *with* $\Delta_J \in \text{Skew}^{n,n}$, $E+\Delta_E, R+\Delta_R \in \text{Sym}^n_{\geq 0}$, *and* $\left\|[\Delta_J, \Delta_R, \Delta_E]\right\|_F = \left\|[\Delta_L(\lambda)]\right\|_F$.

It has been shown in [28] how these distances can be characterized in terms of the following optimization problems:

$$\begin{aligned} &d^{\mathcal{L}}_{\text{sing}}(\lambda E - J + R) \\ &\quad = \min_{\substack{u\in\mathbb{R}^n \\ \|u\|=1}} \sqrt{2\|Ju\|^2 + 2\left\|(I-uu^\top)Eu\right\|^2 + (u^\top Eu)^2 + 2\left\|(I-uu^\top)Ru\right\|^2 + (u^\top Ru)^2}, \end{aligned}$$

$$\begin{aligned} d^{\mathcal{L}}_{\text{hi}}(\lambda E - J + R) &= d^{\mathcal{L}}_{\text{inst}}(\lambda E - J + R) \\ &= \min_{\substack{u\in\mathbb{R}^n \\ \|u\|=1}} \sqrt{2\left\|(I-uu^\top)Eu\right\|^2 + (u^\top Eu)^2 + 2\left\|(I-uu^\top)Ru\right\|^2 + (u^\top Ru)^2}, \end{aligned}$$

and the following very tight lower and upper bounds are available via the smallest eigenvalues of $-J^2+R^2+E^2$ and R^2+E^2, respectively, i.e.,

$$\sqrt{\lambda_{\min}(-J^2+R^2+E^2)} \leq d^{\mathcal{L}}_{\text{sing}}(\lambda E - J + R) \leq \sqrt{2\lambda_{\min}(-J^2+R^2+E^2)},$$

$$\sqrt{\lambda_{\min}(E^2+R^2)} \leq d^{\mathcal{L}}_{\text{hi}}(\lambda E - J + R) = d^{\mathcal{L}}_{\text{inst}}(\lambda E - J + R) \leq \sqrt{2\lambda_{\min}(E^2+R^2)}.$$

Hence, these distances can be computed by global constrained optimization methods such as [32].

Noting that $d^{\mathcal{L}}_{\text{sing}}$ can be computed by determining the closest structure-preserving triplet to (E, R, J) with a common null vector and $d^{\mathcal{L}}_{\text{hi}}$ can be computed by determining the closest symmetric pair to (E, R) with a common null vector, we proceed in a different way and introduce gradient flow methods to compute the discussed structured distances. This is motivated by our experience in computing the distance to instability and singularity for general matrix pencils [13, 18], where it was shown that gradient flow methods were extremely efficient. In the next section we introduce such gradient methods to compute the discussed structured distances.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

3. ODE-based gradient flow approaches. We consider here the problem of computing the nearest common nullspace of a given triplet(E, R, J), that is, the computation of $d^{\mathcal{L}}_{\text{sing}}\big(\lambda E - J + R\big)$.

In the previous section we have recalled the results of [28] that for dH pencils the structured distance to singularity is characterized by the distance to the nearest common nullspace of three structured matrices and the distance to high index and instability by the distance to the nearest common nullspace of two symmetric positive definite matrices. This means that the last two can be considered subcases of the problem we are addressing.

The perturbation matrices that give the structured distance to singularity can be alternatively expressed by the following optimization problem:

$$\begin{aligned}(\Delta E_*, \Delta R_*, \Delta J_*) = \underset{\Delta E, \Delta R, \Delta J}{\arg\min}\ & \|(\Delta E, \Delta R, \Delta J)\|_F \\ & \text{subj. to } E + \Delta E, R + \Delta R \in \mathrm{Sym}^n_{\geq 0}, \Delta J \in \mathrm{Skew}^n, \\ & \text{and } (E + \Delta E)x = 0, \quad (R + \Delta R)x = 0, \quad (J + \Delta J)x = 0 \\ & \text{for some } x \in \mathbb{R}^n, x \neq 0.\end{aligned} \tag{3.1}$$

Then $d^{\mathcal{L}}_{\text{sing}}\big(\lambda E - J + R\big) = \|(\Delta_*, \Theta_*, \Gamma_*)\|$, and our algorithmic approach is based on the minimization of this functional.

3.1. A two-level minimization. To determine the minimum in (3.1) we use a two-level minimization.

As an inner iteration, for a perturbation size ε, we write

$$\Delta E = \varepsilon\Delta, \quad \Delta R = \varepsilon\Theta, \quad \text{and} \quad \Delta J = \varepsilon\Gamma, \quad \text{with} \quad \|(\Delta, \Theta, \Gamma)\|_F = 1, \tag{3.2}$$

where $\Delta, \Theta \in \mathrm{Sym}^n, \Gamma \in \mathrm{Skew}^n$.

Let us denote

(i) by (λ, x) an eigenvalue/eigenvector pair of $E + \varepsilon\Delta$ associated with the smallest eigenvalue and $\|x\| = 1$;

(ii) by (ν, u) an eigenvalue/eigenvector pair of $R + \varepsilon\Theta$ associated with the smallest eigenvalue and $\|u\| = 1$;

(iii-a) if n is even, by $(\mathbf{i}\mu, w)$ an eigenvalue/eigenvector pair of $J + \varepsilon\Gamma$, with $\mu > 0$ such that $\mathbf{i}\mu$ is the eigenvalue with smallest positive imaginary part and $\|w\| = 1$,

(iii-b) if n is odd, by $(0, w)$ an eigenvalue/eigenvector pair of $J + \varepsilon\Gamma$ (this exists for all Γ).

With this notation, a common kernel vector x of the three matrices E, J, R would have $\lambda = \nu = \mu = 0$ and satisfies $x^T u = x^T w = u^T w = 1$, and we use this to reformulate (3.1) below.

In the inner iteration, for any fixed ε we compute a (local) minimizer of (3.1), which is, however, different for even or odd n.

The case when n is odd. In this case the skew-symmetric matrix always has a zero eigenvalue (with an associated real eigenvector w) so that the only contribution to the optimization is through the alignment of w with x and u.

Let

$$B_1 = \{(\Delta, \Gamma, \Theta) \,:\, \|\,(\Delta, \Gamma, \Theta)\,\|_F = 1\}$$

denote the unit ball of the Frobenius norm in the $(n \times 3n)$-dimensional space. Then, taking into account (3.2), the solution of (3.1) can be expressed in the following equivalent form:

$$\min\{\varepsilon > 0 : f(\varepsilon) = 0\}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with

$$f(\varepsilon) = \min_{(\Delta,\Theta,\Gamma)\in B_1} F_\varepsilon^{\mathrm{od}}(\Delta,\Theta,\Gamma)$$

$$F_\varepsilon^{\mathrm{od}}(\Delta,\Theta,\Gamma) = \frac{1}{2}\Big(\lambda^2+\nu^2+1-(x^\top u)^2+1-(x^\top w)^2\Big). \tag{3.3}$$

In $F_\varepsilon^{\mathrm{od}}(\Delta,\Theta,\Gamma)$, the terms λ^2 and ν^2 are the squares of the smallest eigenvalues of $E+\varepsilon\Delta$ and $R+\varepsilon\Theta$, which is our aim to annihilate. The other two square terms $1-(x^\top u)^2$ and $1-(x^\top w)^2$ address the displacement from collinearity of the relevant eigenvectors of $E+\varepsilon\Delta$, $R+\varepsilon\Theta$, and $J+\varepsilon\Gamma$ that we aim also to anihilate in order to obtain a common null vector of the perturbed triplet.

Clearly $F_\varepsilon^{\mathrm{od}}(\Delta,\Theta,\Gamma)$ is non-negative; moreover if $(\Delta E,\Delta R,\Delta J)=\varepsilon\,(\Delta,\Gamma,\Theta)$ is a solution of (3.1), then the functional is annihilated. The goal thus is to find the minimal value ε such that $f(\varepsilon)=0$, which gives the solution of smallest norm and therefore provides the searched distance.

It is possible to include a further term $1-|u^\top w|^2$ in the functional, which does not change the solution but may have an impact on the conditioning of the problem and hence the numerical performance.

The case when n is even. In this case, when two eigenvalues $\pm\mathbf{i}\mu$ ($\mu>0$) coalesce at 0, they form a semi-simple double eigenvalue, and the associated eigenvectors $w=w_1+\mathbf{i}w_2$ and $\overline{w}=w_1-\mathbf{i}w_2$ form a two-dimensional subspace spanned by the two real vectors w_1 and w_2. These can be assumed to be orthogonal to each other, i.e., $w_1^\top w_2=0$, and have the same norm $1/\sqrt{2}$ so that still $\|w\|=1$. Using w_1,w_2, we define the real orthogonal matrix

$$W=\sqrt{2}\,[w_1,w_2]\,,$$

and in order to satisfy the constraint in (3.1), we require that

$$Wz=x \qquad \text{for some } z\in\mathbb{R}^2.$$

This leads to the minimization of

$$\|Wz-x\| \qquad \text{for some } z\in\mathbb{R}^2.$$

Since W is orthogonal, the solution is $z=W^\top x$, and the part of the functional associated with this constraint takes the form

$$1-x^\top WW^\top x = 1-2\,(x^\top w_1)^2-2\,(x^\top w_2)^2,$$

which is positive if x does not lie in the range of W and zero otherwise.

In summary, the functional in the even case is given by

$$F_\varepsilon^{\mathrm{ev}}(\Delta,\Theta,\Gamma)=\frac{1}{2}\Big(\lambda^2+\nu^2+\mu^2+1-(x^\top u)^2+1-2\,(x^\top \mathrm{Re}(w))^2-2\,(x^\top \mathrm{Im}(w))^2\Big) \tag{3.4}$$

with $\|\,(\Delta,\Gamma,\Theta)\,\|_F\le 1$.

Using the functionals (3.3), respectively (3.4), in our approach the local minimizer of

$$\min_{(\Delta,\Gamma,\Theta)\in B_1} F_\varepsilon(\Delta,\Theta,\Gamma)$$

is determined as an equilibrium point of the associated gradient system. Note, however, that in general this may not be a global minimizer.

For the outer iteration we consider a continuous branch, as a function of ε, of the minimizers $(\Delta(\varepsilon), \Gamma(\varepsilon), \Theta(\varepsilon))$ and vary ε iteratively in order to find the smallest solution of the scalar equations

$$f^{\mathrm{ev}}(\varepsilon) := F_{\varepsilon}^{\mathrm{ev}}\left((\Delta(\varepsilon), \Gamma(\varepsilon), \Theta(\varepsilon))\right) = 0$$

or, in the odd case,

$$f^{\mathrm{od}}(\varepsilon) := F_{\varepsilon}^{\mathrm{od}}\left((\Delta(\varepsilon), \Gamma(\varepsilon), \Theta(\varepsilon))\right) = 0,$$

respectively, with respect to ε.

REMARK 3.1. Note that the techniques for the distance to higher index or to a pencil with eigenvalues on the imaginary axis follow directly by setting $J = 0$ and not perturbing it.

3.2. Derivatives of eigenvalues and eigenvectors. The considered minimization is an eigenvalue optimization problem. We will solve this problem by integrating a differential equation with trajectories that follow the gradient descent and satisfy further constraints. To develop such a method, we first recall a classical result (see, e.g., [22]) for the derivative of a simple eigenvalue and an associated eigenvector of a matrix $C(t)$ with respect to variations in a real parameter t of the entries. Here we use the notation $\dot{C}(t) := \frac{d}{dt}C(t)$ to denote the derivative with respect to t.

LEMMA 3.2 ([22, Section II.1.1]). *Consider a continuously differentiable matrix-valued function $C(t) : \mathbb{R} \to \mathbb{R}^{n\times n}$, with $C(t)$ normal (i.e., $C(t)C(t)^\top = C(t)^\top C(t)$ for all t). Let $\lambda(t)$ be a simple eigenvalue of $C(t)$ for all t, and let $x(t)$ with $\|x(t)\| = 1$ be the associated (right and left) eigenvector. Then $\lambda(t)$ is differentiable with*

$$\dot{\lambda}(t) = x(t)^H \dot{C}(t) x(t) = \langle x(t)x(t)^H, \dot{C}(t)\rangle,$$

where we recall that $\langle\cdot,\cdot\rangle$ denotes the Frobenius inner product.

For $A \in \mathrm{Sym}^n$ consider a perturbation matrix $\varepsilon\Delta(t) \in \mathrm{Sym}^n$ that depends on a real parameter t. By Lemma 3.2, for a simple eigenvalue $\lambda(t) \in \mathbb{R}$ of $A + \varepsilon\Delta(t)$ with associated eigenvector $x(t)$, $\|x(t)\| = 1$, we have (omitting the dependence on t)

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\lambda^2 = \varepsilon\,\lambda\, x^\top \dot{\Delta} x = \varepsilon\,\lambda\,\langle xx^\top, \dot{\Delta}\rangle.$$

Similarly, a result needed in the case when n is even, for all t, if $\mathbf{i}\mu(t) \in \mathbf{i}\mathbb{R}$ ($\mu(t) \geq 0$) is a simple eigenvalue of a matrix-valued function $B + \varepsilon\Theta(t) \in \mathrm{Skew}^{n,n}$ with associated eigenvector $w(t)$, $\|w(t)\| = 1$, then we have

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\lambda|^2 = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mu^2 &= \mathrm{Re}\overline{\lambda}\dot{\lambda} = \varepsilon\mathrm{Re}\left(-\mathbf{i}\mu\langle ww^H,\ \dot{\Theta}\rangle\right)\\ &= \varepsilon\mathrm{Re}\left(-\mathbf{i}\mu\left(-\mathbf{i}\langle\mathrm{Im}(ww^H),\ \dot{\Theta}\rangle\right)\right) = -\varepsilon\,\mu\,\langle\mathrm{Im}\left(ww^H\right), \dot{\Theta}\rangle.\end{aligned}$$

To derive the gradient system associated with our optimization problem, we make use of the following definition:

DEFINITION 3.3. *Let $M \in \mathbb{C}^{n\times n}$ be a singular matrix with a simple zero eigenvalue. The* group inverse *(reduced resolvent) of M is the unique matrix G satisfying*

$$MG = GM, \qquad GMG = G, \quad \textit{and} \quad MGM = M.$$

It is well-known (see [31]) that for a singular and normal matrix $M \in \mathbb{C}^{n\times n}$ with simple eigenvalue zero, its group inverse G is equal to the Moore-Penrose pseudoinverse M^+. We have the following lemma:

LEMMA 3.4. [31, Theorem 2] *Consider a sufficiently often differentiable matrix function*

$$C : \mathbb{R} \to \mathbb{C}^{n\times n}.$$

Let $\lambda(t)$ be a simple eigenvalue of $C(t)$ for all t, and let $x(t)$, with $\|x(t)\| = 1$, be the associated right eigenvector function. Moreover, let $M(t) = C(t) - \lambda(t)I$, and let $G(t)$ be the group inverse of $M(t)$. Then $x(t)$ satisfies the system of differential equations

$$\dot{x} = x\, x^H G(t)\dot{M}(t)x - G(t)\dot{M}(t)x.$$

Moreover, if $C(t)$ is pointwise normal, then

$$\dot{x}(t) = -G(t)\dot{M}(t)x(t).$$

After these preparations, in the following sections we determine the associated gradient systems for the functionals (3.3), respectively (3.4).

4. Gradient flow, odd state dimension.

In this section we consider the case when the state dimension is odd, and we construct the gradient system optimization algorithm for the functional (3.3).

4.1. Computation of the gradient.

The functional $F_\varepsilon^{\mathrm{od}}(\Delta, \Theta, \Gamma)$ in (3.3) has several parts. Applying Lemma 3.2 for the perturbations $\varepsilon\Delta(t)$ of E and $\varepsilon\Theta(t)$ of R, the computation of the gradient of the term $\frac{1}{2}\left(\lambda^2 + \nu^2\right)$ is obtained from the expressions

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\lambda^2 = \varepsilon\,\lambda\,\langle xx^\top, \dot{\Delta}\rangle, \qquad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\nu^2 = \varepsilon\,\nu\,\langle uu^\top, \dot{\Theta}\rangle. \tag{4.1}$$

To determine the gradients of $\frac{1}{2}\lambda^2$ with respect to Δ and of $\frac{1}{2}\nu^2$ with respect to Θ (that means, the steepest ascent directions $\dot{\Delta}$ and $\dot{\Theta}$ in (4.1)), which we denote as G_Δ^1 and G_Θ^1, respectively, we directly obtain (recall that $\dot{\Delta}$ and $\dot{\Theta}$ have to be symmetric)

$$G_\Delta^1 \propto \lambda\, xx^\top, \qquad G_\Theta^1 \propto \nu\, uu^\top,$$

were $\propto$ denotes real and positive proportionality (for uniqueness we can normalize them to have unit norm). These represent the gradient terms.

This result is due to the fact that matrices can be considered as vectors with respect to the Frobenius inner product, which is maximized when the two arguments are proportional (i.e., aligned). Note that since $xx^\top$ and $uu^\top$ belong to Sym^n, no orthogonal projection is needed for obtaining $\dot{\Delta}$ and $\dot{\Theta}$.

With this argument in mind, in order to treat the other terms, we observe that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left((x^\top u)^2\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(x^\top uu^\top x\right) = x^\top uu^\top \dot{x} + u^\top xx^\top \dot{u},$$

and thus, making use of Lemma 3.4 and recalling that x and u are real,

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(1-(x^\top u)^2\right) &= \varepsilon\left((x^\top u)u^\top L\dot{\Delta}x + (u^\top x)x^\top N\dot{\Theta}u\right)\\ &= \varepsilon\theta\left(\left\langle L^\top ux^\top, \dot{\Delta}\right\rangle + \left\langle N^\top xu^\top, \dot{\Theta}\right\rangle\right),\end{aligned}$$

where $\theta = x^\top u$, $L = (E+\varepsilon\Delta-\lambda I)^+$, and $N = (R+\varepsilon\Theta-\nu I)^+$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Recalling that $\dot\Delta$ and $\dot\Theta$ have to be symmetric and noting that for $A, B \in \mathbb{R}^{n,n}$ and $A \in \mathrm{Sym}^n$, $\langle A, B\rangle = \langle A, \mathrm{Sym}(B)\rangle$, where $\mathrm{Sym}(A)$ denotes the symmetric part of A, the gradients of $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(1-(x^\top u)^2\right)$ with respect to Δ, say G^2_Δ, and with respect to Θ, say G^2_Θ, are obtained by simply considering the symmetric part of the left terms in the scalar products:

$$G^2_\Delta \propto \theta\,\mathrm{Sym}(L^\top u x^\top), \qquad G^2_\Theta \propto \theta\,\mathrm{Sym}(N^\top x u^\top).$$

Finally, since n is odd, 0 is a (generically) simple eigenvalue of $J + \varepsilon\Gamma$ and the associated eigenvector w can be chosen to be real. Thus, for the last term of (3.3) we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(1-(x^\top w)^2\right) = \varepsilon\eta\left(\left\langle L^\top w x^\top, \dot\Delta\right\rangle + \left\langle P^\top x w^\top, \dot\Gamma\right\rangle\right),$$

where $\eta = x^\top w$ and $P = (J+\varepsilon\Gamma)^+$.

Recalling that $\dot\Gamma$ has to be skew-symmetric and noting that for $A, B \in \mathbb{R}^{n,n}$ and $A \in \mathrm{Skew}^n$, $\langle A, B\rangle = \langle A, \mathrm{Skew}(B)\rangle$, where $\mathrm{Skew}(B)$ denotes the skew-symmetric part of B, the gradients of $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(1-(x^\top w)^2\right)$ with respect to Δ (denoted as G^3_Δ) and with respect to Γ (denoted as G^3_Γ) are given by

$$G^3_\Delta \propto \eta\,\mathrm{Sym}(L^\top w x^\top), \qquad G^3_\Gamma \propto \eta\,\mathrm{Skew}(P^\top x w^\top).$$

4.2. The gradient system of ODEs for the flow in the odd case. In order to compute the steepest descent direction, we minimize the gradient of F_ε and collect the summands involving $\dot\Delta$, $\dot\Theta$, and those involving $\dot\Gamma$. Setting

$$\begin{aligned} p &= \theta L^\top u + \eta L^\top w,\\ q &= \theta N^\top x,\\ r &= \eta P^\top x, \end{aligned} \tag{4.2}$$

and collecting the computed terms $G^{1,2,3}_\Delta, G^{1,2}_\Theta$, and G^3_Γ, we get the so-called *free gradient* of the functional,

$$\mathbf{G} = \left[\mathrm{Sym}\left((\lambda x + p)\, x^\top\right), \mathrm{Sym}\left((\nu u + q)\, u^\top\right), \mathrm{Skew}\left(r\, w^\top\right)\right] := \left[\mathbf{G}_E, \mathbf{G}_R, \mathbf{G}_J\right], \tag{4.3}$$

where "free" emphasizes the fact that we do not impose norm preservation of the perturbation (Δ, Θ, Γ).

However, since we want to impose a norm constraint on the perturbation (Δ, Θ, Γ), we need the following result.

LEMMA 4.1 (Direction of steepest admissible ascent). *Let* $\mathbf{G} = (\mathbf{G}_E, \mathbf{G}_R, \mathbf{G}_J) \in \mathbb{R}^{n,3n}$, $\mathbf{M} = (\Delta, \Theta, \Gamma) \in \mathbb{R}^{n,3n}$, $\mathbf{Z} \in \mathbb{R}^{n,3n}$. *A solution of the optimization problem*

$$\mathbf{Z}_\star = \underset{\|\mathbf{Z}\|_F = 1,\ \langle \mathbf{M}, \mathbf{Z}\rangle = 0}{\arg\min} \langle \mathbf{G}, \mathbf{Z}\rangle$$

is given by

$$\mu \mathbf{Z}_\star = -\mathbf{G} + \varrho\,\mathbf{M}, \qquad \varrho = \langle \mathbf{M}, \mathbf{G}\rangle = \left(\langle \Delta, \mathbf{G}_E\rangle + \langle \Theta, \mathbf{G}_R\rangle + \langle \Gamma, \mathbf{G}_J\rangle\right), \tag{4.4}$$

where μ is the Frobenius norm of the matrix on the right-hand side. The solution is unique if $\mathbf{G}$ *is not a multiple of* $\mathbf{M}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. For real vectors $m, g, z \in \mathbb{R}^d$, the problem

$$\underset{\|z\|=1,\ \langle m,z\rangle=0}{\arg\min} \quad \langle g, z\rangle$$

is solved by choosing as z the orthogonal projection of g onto m.

The result thus follows on noting that the function to minimize is a real inner product on $\mathbb{R}^{n,3n}$ and the real inner product with a given vector (which here is a matrix) is minimized over a subspace by orthogonally projecting the vector onto that subspace. The expression in (4.4) is the orthogonal projection of L to the tangent space at $\mathbf{M}$ of the manifold of matrices of unit Frobenius norm. ∎

Lemma 4.1 and formula (4.3) lead us to consider the system of differential equations for the matrices Δ, Θ, and Γ:

$$\begin{aligned}\dot{\Delta} &= -\mathrm{Sym}\left((\lambda x + p)\ x^\top\right) + \varrho\Delta,\\ \dot{\Theta} &= -\mathrm{Sym}\left((\nu u + q)\ u^\top\right) + \varrho\Theta,\\ \dot{\Gamma} &= -\mathrm{Skew}\left(r\, w^\top\right) + \varrho\Gamma,\end{aligned} \tag{4.5}$$

where, for $X \in \mathbb{R}^{n,n}$, $\mathrm{Sym}(X) = \frac{X+X^\top}{2}$, $\mathrm{Skew}(X) = \frac{X-X^\top}{2}$, and

$$\varrho = \left(\langle \Delta, \mathrm{Sym}\left((\lambda x + p)\, x^\top\right)\rangle + \langle \Theta, \mathrm{Sym}\left((\nu u + q)\, u^\top\right)\rangle + \langle \Gamma, \mathrm{Skew}\left(r\, w^\top\right)\rangle\right)$$

is used to ensure the norm conservation, i.e., $\langle(\dot{\Delta}, \dot{\Theta}, \dot{\Gamma}), (\Delta, \Theta, \Gamma)\rangle = 0$.

REMARK 4.2. Neglecting the last differential equation in (4.5) and setting $\Gamma \equiv 0$ and $\mathbf{G}_J = 0$ allows us to treat the other distance problems mentioned in the introduction, which involve the search of a common null space of two symmetric positive definite matrices.

The following result is a consequence of the gradient system property of the system of ODEs (4.5).

THEOREM 4.3. *Let $(\Delta(t), \Theta(t), \Gamma(t))$ of unit Frobenius norm satisfy the differential equation (4.5). If $\lambda(t)$ is a simple eigenvalue of $E + \varepsilon\Delta(t)$, then*

$$\frac{d}{dt} F_\varepsilon^{\mathrm{od}}\left(\Delta(t), \Theta(t), \Gamma(t)\right) \leq 0. \tag{4.6}$$

Proof. The result follows directly by the fact that (4.5) is a constrained gradient system. ∎

In this way we have preserved the symmetry of E, R and the skew-symmetry of J. It may happen that along the solution trajectory of (4.5), due to the projection on the matrix manifolds, the smallest eigenvalue ν of $R + \varepsilon\Theta$ and/or the smallest eigenvalue λ of $E + \varepsilon\Delta$ become negative. In this case the perturbed system is not a dissipative Hamiltonian system any longer. This is in general not an issue for the optimization algorithm, since the dynamical gradient system leads to eigenvalues ν, λ, with $|\nu|$ as small as possible, for a given ε, and thus drives them to zero when $\varepsilon = \varepsilon^*$, so that in the limiting situation also the positive semidefiniteness of $E + \varepsilon\Delta$ and $R + \varepsilon\Theta$ holds.

4.3. Stationary points of (4.5) and low-rank property. In this section we discuss the existence of stationary points of the solution trajectory of (4.5).

LEMMA 4.4. *Let ε be fixed, and let $F_\varepsilon^{\mathrm{od}}(\Delta, \Theta, \Gamma) > 0$. Let λ be a simple eigenvalue of $E + \varepsilon\Delta$ with associate normalized eigenvector x, let ν be a simple eigenvalue of $R + \varepsilon\Theta$ with associate normalized eigenvector u, and let 0 be a simple eigenvalue of $J + \varepsilon\Gamma$ with associate*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

normalized eigenvector w. Then, in the generic situation, i.e., if $\lambda, \nu \neq 0$, $\theta = x^\top u \neq 0$, and $\eta = x^\top w \in (0,1)$, we have

$$\lambda x + p \neq 0, \qquad \nu u + q \neq 0, \qquad \textit{and} \qquad r \neq 0.$$

Proof. The proofs for the three cases are similar.

(i) Exploiting the property that $Lx = 0$ (see, e.g., [31][Section 2]), we obtain that $x^\top p = 0$. If we had $x^\top (\lambda x + p) = 0$, then this would imply $\lambda x^\top x = 0$, and thus, since $\lambda \neq 0$, we get a contradiction since $\|x\| = 1$.

(ii) Exploiting the property $Nu = 0$, we obtain that $u^\top q = 0$. If we had $u^\top (\nu u + q) = 0$, then we would get $\nu u^\top u = 0$, and again we have a contradiction.

(iii) Having assumed $\eta = x^\top w \neq 1$, we have that x and w are not aligned. As a consequence $\eta P^\top x \neq 0$. ∎

Using Lemma 4.4, we have the following characterization of stationary points.

THEOREM 4.5. *Let $(\Delta(t), \Theta(t), \Gamma(t))$ of unit Frobenius norm satisfy the differential equation (4.5). Moreover, suppose that for all t*

$$F_\varepsilon^{\mathrm{od}}(\Delta(t), \Theta(t), \Gamma(t)) > 0$$

and that $0 \neq \lambda(t) \in \mathbb{R}$ is a simple eigenvalue of $E + \varepsilon\Delta(t)$ with normalized eigenvector $x(t)$, that $0 \neq \nu(t) \in \mathbb{R}$ is a simple eigenvalue of $R + \varepsilon\Theta(t)$ with associated eigenvector $u(t)$, and that $J + \varepsilon\Gamma(t)$ has a null vector $w(t)$.

Then the following are equivalent (here we omit the argument t):

(1) $\dfrac{d}{dt} F_\varepsilon^{\mathrm{od}}(\Delta, \Theta, \Gamma) = 0$;

(2) $\dot{\Delta} = 0$, $\dot{\Theta} = 0$, $\dot{\Gamma} = 0$;

(3) *Δ is a multiple of the rank-two matrix $\mathrm{Sym}\left((\lambda x + p)\ x^\top\right)$; Θ is a multiple of the rank-two matrix $\mathrm{Sym}\left((\nu u + q)\ u^\top\right)$; Γ is a multiple of the rank-two matrix $\mathrm{Skew}\left(r\, w^\top\right)$ with p, q, r given by (4.2).*

Proof. The proof follows directly by equating to zero the right-hand side of (4.5) and by Lemma 4.4, which prevents the matrices to be zero. ∎

We have also the following extremality property:

THEOREM 4.6. *Consider the functional (3.3), and suppose that $F_\varepsilon^{\mathrm{od}}(\Delta, \Theta, \Gamma) > 0$. Let $\Delta_* \in \mathrm{Sym}^n, \Theta_* \in \mathrm{Sym}^n$, and $\Gamma_* \in \mathrm{Skew}^n$, with $\|(\Delta_*, \Theta_*, \Gamma_*)\|_F = 1$. Let $0 \neq \lambda_* \in \mathbb{R}$ be a simple eigenvalue of $E + \varepsilon\Delta_*$ with associated eigenvector x, let $0 \neq \nu_* \in \mathbb{R}$ be a simple eigenvalue of $R + \varepsilon\Theta_*$ with associated eigenvector u, and let $J + \varepsilon\Gamma_*$ have a null vector w. Then the following are equivalent:*

(i) *Every differentiable path $(\Delta(t), \Theta(t), \Gamma(t))$ (for small $t \geq 0$) with the properties that $\|(\Delta(t), \Theta(t), \Gamma(t))\|_F \leq 1$, that both $\lambda(t)$ and $\nu(t)$ are simple eigenvalues of $E + \varepsilon\Delta(t)$ and $R + \varepsilon\Theta(t)$ with associated eigenvectors $x(t)$ and $u(t)$, respectively, and for which $w(t)$ is the null vector of $J + \varepsilon\Gamma(t)$, so that $\Delta(0) = \Delta_*$, $\Theta(0) = \Theta_*$, $\Gamma(0) = \Gamma_*$, satisfies*

$$\frac{d}{dt} F_\varepsilon^{\mathrm{od}}(\Delta(t), \Theta(t), \Gamma(t)) \geq 0.$$

(ii) *The matrix Δ_* is a multiple of the rank-two matrix $\mathrm{Sym}\left((\lambda x + p)\ x^\top\right)$, Θ_* is a multiple of the rank-two matrix $\mathrm{Sym}\left((\nu u + q)\ u^\top\right)$, and Γ_* is a multiple of the rank-two matrix $\mathrm{Skew}\left(r\, w^\top\right)$ with p, q, r given by (4.2).*

Proof. Lemma 4.4 ensures that $\lambda x + p \neq 0$. Assume that (i) does not hold. Then there exists a path $(\Delta(t), \Theta(t), \Gamma(t))$ through $(\Delta_*, \Theta_*, \Gamma_*)$ with $\frac{d}{dt} F_\varepsilon^{\rm od}(\Delta(t), \Theta(t), \Gamma(t))\big|_{t=0} < 0$. The steepest descent gradient property shows that also the solution path of (4.5) passing through $(\Delta_*, \Theta_*, \Gamma_*)$ is such a path. Hence $(\Delta_*, \Theta_*, \Gamma_*)$ is not a stationary point of (4.5), and Theorem 4.5 then yields that (ii) does not hold.

Conversely, if

$$(\Delta_*, \Theta_*, \Gamma_*) \not\propto \left(\mathrm{Sym}\left((\lambda x + p)\; x^\top\right), \mathrm{Sym}\left((\nu u + q)\; u^\top\right), \mathrm{Skew}\left(r\, w^\top\right)\right),$$

then $(\Delta_*, \Theta_*, \Gamma_*)$ is not a stationary point of (4.5), and Theorems 4.5 and 4.3 yield that $\frac{d}{dt} F_\varepsilon^{\rm od}(\Delta(t), \Theta(t), \Gamma(t))\big|_{t=0} < 0$ along the solution path of (4.5). ◻

In the next section we illustrate the properties of the optimization procedure with a numerical example.

4.4. A numerical example. Let $n = 5$ and consider the randomly generated matrices

$$E = \begin{bmatrix} 0.15 & 0.02 & -0.04 & 0.02 & -0.04 \\ 0.02 & 0.22 & 0 & -0.01 & -0.03 \\ -0.04 & 0 & 0.11 & -0.07 & -0.04 \\ 0.02 & -0.01 & -0.07 & 0.01 & 0.10 \\ -0.04 & -0.03 & -0.04 & 0.10 & 0.39 \end{bmatrix},$$

$$R = \begin{bmatrix} 0.49 & -0.13 & 0.05 & -0.15 & 0.11 \\ -0.13 & 0.23 & -0.05 & -0.10 & -0.19 \\ 0.05 & -0.05 & 0.48 & -0.06 & 0.02 \\ -0.15 & -0.10 & -0.06 & 0.55 & 0.16 \\ 0.11 & -0.19 & 0.02 & 0.16 & 0.48 \end{bmatrix},$$

$$J = \begin{bmatrix} 0 & -0.27 & -0.03 & -0.01 & 0.21 \\ 0.27 & 0 & -0.15 & 0.03 & 0.11 \\ 0.03 & 0.15 & 0 & 0.07 & -0.07 \\ 0.01 & -0.03 & -0.07 & 0 & 0.05 \\ -0.21 & -0.11 & 0.07 & -0.05 & 0 \end{bmatrix}.$$

Running the two level iteration with an initial value of the functional $F_0^{\rm od}(\cdot, \cdot, \cdot) = 0.9181$, we find a perturbation at a distance (rounded to four digits) $\varepsilon^* = 0.3568$ with a common null space given by the vector

$$c = \begin{bmatrix} 0.2195 \\ -0.6664 \\ -0.0639 \\ 0.3187 \\ -0.6341 \end{bmatrix}$$

and the computed perturbations are given by

$$\Delta E = \varepsilon^* \Delta = \begin{bmatrix} -0.0385 & 0.0912 & 0.0089 & 0.0251 & 0.1408 \\ 0.0912 & -0.2114 & -0.0218 & -0.0903 & -0.3482 \\ 0.0089 & -0.0218 & -0.0009 & 0.0019 & -0.0293 \\ 0.0251 & -0.0903 & 0.0019 & 0.1628 & -0.0123 \\ 0.1408 & -0.3482 & -0.0293 & -0.0123 & -0.4801 \end{bmatrix},$$

$$\Delta R = \varepsilon^* \Theta = \begin{bmatrix} -0.0689 & 0.1166 & 0.0148 & -0.1044 & 0.1242 \\ 0.1166 & -0.1164 & -0.0307 & 0.1766 & -0.1433 \\ 0.0148 & -0.0307 & -0.0049 & 0.0208 & -0.0319 \\ -0.1044 & 0.1766 & 0.0208 & -0.1592 & 0.1884 \\ 0.1242 & -0.1433 & -0.0319 & 0.1884 & -0.1679 \end{bmatrix},$$

$$\Delta J = \varepsilon^* \Gamma = \begin{bmatrix} 0 & 0.0887 & -0.0118 & -0.0474 & 0.0852 \\ -0.0887 & 0 & 0.0496 & 0.0003 & 0.0027 \\ 0.0118 & -0.0496 & 0 & 0.0257 & -0.0488 \\ 0.0474 & -0.0003 & -0.0257 & 0 & -0.0030 \\ -0.0852 & -0.0027 & 0.0488 & 0.0030 & 0 \end{bmatrix}.$$

Considering the same problem—but with the aim of studying its distance from a highest index problem—we set $\Delta J \equiv 0$ and optimize with respect to ΔE and ΔR. In this way we find $\varepsilon^* = 0.2854$ with

$$\Delta E = \begin{bmatrix} -0.0392 & -0.0779 & 0.0020 & -0.0156 & 0.0080 \\ -0.0779 & -0.1449 & -0.0091 & -0.0440 & 0.0125 \\ 0.0020 & -0.0091 & 0.0306 & 0.0478 & -0.0010 \\ -0.0156 & -0.0440 & 0.0478 & 0.0824 & -0.0046 \\ 0.0080 & 0.0125 & -0.0010 & -0.0046 & 0.0015 \end{bmatrix},$$

$$\Delta R = \begin{bmatrix} -0.0208 & -0.0214 & -0.0282 & -0.0456 & 0.0059 \\ -0.0214 & -0.0048 & -0.0389 & -0.0624 & 0.0167 \\ -0.0282 & -0.0389 & -0.0249 & -0.0448 & 0.0051 \\ -0.0456 & -0.0624 & -0.0448 & -0.0777 & 0.0072 \\ 0.0059 & 0.0167 & 0.0051 & 0.0072 & 0.0062 \end{bmatrix}.$$

4.5. Sparsity preservation. If the matrices E, R, and J have a given sparsity pattern, then we may include as a constraint that the perturbations do not alter the sparsity structure. In terms of the Frobenius norm it is immediate to obtain the constrained gradient system. Denoting by Π_E, Π_R, and Π_J, respectively, projections onto the manifold of sparse matrices with the given sparsity pattern (and symmetric/skew-symmetric structure) of E, R, and J, then we get

$$\begin{aligned} \dot{\Delta} &= -\Pi_E \mathrm{Sym}\left((\lambda x + p)\ x^\top\right) + \varrho \Delta, \\ \dot{\Theta} &= -\Pi_R \mathrm{Sym}\left((\nu u + q)\ u^\top\right) + \varrho \Theta, \\ \dot{\Gamma} &= -\Pi_J \mathrm{Skew}\left(r\, w^\top\right) + \varrho \Gamma, \end{aligned}$$

where

$$\varrho = \left(\langle \Delta, \Pi_E \mathrm{Sym}\left((\lambda x + p)\, x^\top\right)\rangle + \langle \Theta, \Pi_R \mathrm{Sym}\left((\nu u + q)\, u^\top\right)\rangle + \langle \Gamma, \Pi_J \mathrm{Skew}\left(r\, w^\top\right)\rangle\right).$$

Note that the orthogonal projections Π_E and Π_R commute with Sym, and similarly Π_J commutes with Skew.

5. Rank-two optimization. Theorem 4.5 motivates to search for a differential equation on the manifold of rank-two symmetric/skew-symmetric matrices which still leads to a gradient system for $F_\varepsilon^{\mathrm{od}}$, but in addition requires the derivatives of the matrices Δ, Θ, Γ lying in the respective tangent spaces.

Let $\mathcal{M}_2^{n,n} := \{X \in \mathbb{R}^{n\times n} : \mathrm{rank}(X) = 2\}$. Then we restrict the perturbations to the matrix manifolds

$$\Delta, \Theta \in \mathcal{M}_2^{\mathrm{Sym}^n}, \qquad \Gamma \in \mathcal{M}_2^{\mathrm{Skew}^n},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $\mathcal{M}_2^{\mathrm{Sym}^n} = \mathcal{M}_2^{n,n} \cap \mathrm{Sym}^n$ and $\mathcal{M}_2^{\mathrm{Skew}^n} = \mathcal{M}_2^{n,n} \cap \mathrm{Skew}^n$.

Following [23], every real symmetric rank-two matrix X of dimension $n \times n$ can be written in the form

$$X = USU^{\mathrm{T}}, \tag{5.1}$$

where $U \in \mathbb{R}^{n\times 2}$ has orthonormal columns, i.e., $U^{\mathrm{T}}U = I_2$ and $S \in \mathrm{Sym}^2$. Here we will not assume that S is diagonal. Note that the representation (5.1) is not unique; indeed replacing U by $\widetilde{U} = UU_1$ with orthogonal $U_1 \in \mathbb{R}^{2\times 2}$ and correspondingly S by $\widetilde{S} = U_1^{\mathrm{T}} SU_1$, yields the same matrix $X = USU^{\mathrm{T}} = \widetilde{U}\widetilde{S}\widetilde{U}^{\mathrm{T}}$.

As a compensation for the non-uniqueness in the decomposition (5.1), we will use a unique decomposition in the tangent space. Let $\mathcal{V}_{n,2}$ denote the Stiefel manifold of real $n \times 2$ matrices with orthonormal columns. The tangent space at $U \in \mathcal{V}_{n,2}$ is given by

$$T_U\mathcal{V}_{n,2} = \{\dot{U} \in \mathbb{R}^{n\times 2} : \ \dot{U}^{\mathrm{T}}U + U^{\mathrm{T}}\dot{U} = 0\} = \{\dot{U} \in \mathbb{R}^{n\times 2} : \ U^{\mathrm{T}}\dot{U} \text{ is skew-symmetric}\}.$$

Following [23], every tangent matrix $\dot{X} \in T_X\mathcal{M}_2^{\mathrm{Sym}^n}$ is of the form

$$\dot{X} = \dot{U}SU^{\mathrm{T}} + U\dot{S}U^{\mathrm{T}} + US\dot{U}^{\mathrm{T}}, \tag{5.2}$$

where $\dot{S} \in \mathrm{Sym}^2$, $\dot{U} \in T_U\mathcal{V}_{n,2}$, and $\dot{S}, \dot{U}$ are uniquely determined by $\dot{X}$ and U, S, if we impose the orthogonality condition

$$U^{\mathrm{T}}\dot{U} = 0. \tag{5.3}$$

We note the following lemma adapted from [23].

LEMMA 5.1. *The orthogonal projection onto the tangent space $T_X\mathcal{M}_2^{\mathrm{Sym}^n}$ at $X = USU^{\mathrm{T}} \in \mathcal{M}_2^{\mathrm{Sym}^n}$ is given by*

$$P_X^{\mathrm{Sym}}(Z) = Z - (I - UU^{\mathrm{T}})Z(I - UU^{\mathrm{T}})$$

for $Z \in \mathrm{Sym}^n$. Analogous results hold for $Y \in \mathrm{Skew}^n$, this time with $S \in \mathrm{Skew}^2$.

5.1. A differential equation for rank-two matrices. To derive the differential equation in the rank-two case, we replace in (4.5) the right-hand sides by the orthogonal projections to $T_\Delta\mathcal{M}_2^{\mathrm{Sym}^n}$, $T_\Theta\mathcal{M}_2^{\mathrm{Sym}^n}$, and $T_\Gamma\mathcal{M}_2^{\mathrm{Skew}^n}$, respectively, so that solutions starting with rank-two initial values will retain rank-two for all t. This gives the differential equations

$$\begin{aligned}
\dot{\Delta} &= -P_\Delta^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\lambda x + p)\ x^\top\right) + \varrho\Delta\Big),\\
\dot{\Theta} &= -P_\Theta^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\nu u + q)\ u^\top\right) + \varrho\Theta\Big),\\
\dot{\Gamma} &= -P_\Gamma^{\mathrm{Skew}}\Big(\mathrm{Skew}\left(r\, w^\top\right) + \varrho\Gamma\Big),
\end{aligned}$$

where again p, q, and r are defined by (4.2) and

$$\begin{aligned}
\varrho = {}&\Big\langle \Delta, P_\Delta^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\lambda x + p)\ x^\top\right)\Big)\Big\rangle +\\
&+ \Big\langle \Theta, P_\Theta^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\nu u + q)\ u^\top\right)\Big)\Big\rangle + \Big\langle \Gamma, P_\Gamma^{\mathrm{Skew}}\Big(\mathrm{Skew}\left(r\, w^\top\right)\Big)\Big\rangle.
\end{aligned}$$

Since for $X \in \mathcal{M}_2^{\mathrm{Sym}^n}$ and $Z \in \mathrm{Sym}^n$, we have $P_X^{\mathrm{Sym}}(X) = X$ and $\langle X, Z\rangle = \langle X, P_X^{\mathrm{Sym}}(Z)\rangle$, (and analogous properties hold for $X \in \mathcal{M}_2^{\mathrm{Skew}^n}$ and $Y \in \mathcal{M}_2^{\mathrm{Skew}^n}$), the system of differential equations can be rewritten as

$$\dot{\Delta} = -P_\Delta^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\lambda x + p)\ x^\top\right)\Big) + \varrho\Delta,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$
\begin{aligned}
\dot{\Theta} &= -P_{\Theta}^{\mathrm{Sym}}\Big(\mathrm{Sym}\left((\nu u + q)\; u^{\top}\right)\Big) + \varrho\Theta, \\
\dot{\Gamma} &= -P_{\Gamma}^{\mathrm{Skew}}\Big(\mathrm{Skew}\left(r\, w^{\top}\right)\Big) + \varrho\Gamma,
\end{aligned} \tag{5.4}
$$

with

$$
\begin{aligned}
\varrho = \Big(&\Big\langle \Delta, P_{\Delta}^{\mathrm{Sym}}\left(\mathrm{Sym}\left((\lambda x + p)\, x^{\top}\right)\right)\Big\rangle + \Big\langle \Theta, P_{\Theta}^{\mathrm{Sym}}\left(\mathrm{Sym}\left((\nu u + q)\, u^{\top}\right)\right)\Big\rangle \\
&+ \left\langle \Gamma, P_{\Gamma}^{\mathrm{Skew}}\left(\mathrm{Skew}\left(r\, w^{\top}\right)\right)\right\rangle \Big).
\end{aligned}
$$

This system differs from (4.5) in that the free gradient terms are replaced by their orthogonal projections on the rank-two manifold of the corresponding structure.

To obtain the differential equation in a form that uses the factors in $X = USU^{\mathrm{T}}$ rather than the full $n \times n$ matrix X, we use the following result, whose proof is similar to that given in [23, Prop. 2.1].

LEMMA 5.2. *For $X = USU^{\mathrm{T}} \in \mathcal{M}_2^{\mathrm{Sym}^n}$ with nonsingular $S \in \mathrm{Sym}^2$ and with $U \in \mathbb{R}^{n\times 2}$ having orthonormal columns, the equation $\dot{X} = P_X^{\mathrm{Sym}}(Z)$ with Z symmetric is equivalent to $\dot{X} = \dot{U}SU^{\mathrm{T}} + U\dot{S}U^{\mathrm{T}} + US\dot{U}^{\mathrm{T}}$, where*

$$
\dot{S} = U^{\mathrm{T}} Z U, \qquad \dot{U} = (I - UU^{\mathrm{T}}) Z U S^{-1}.
$$

An analogous statement holds for $Y \in \mathcal{M}_2^{\mathrm{Skew}^n}$ and Z skew-symmetric.

With the ansatz $Z = -\lambda x x^{\top} - \frac{1}{2} p x^{\top} + \frac{1}{2} x p^{\top}$ and introducing for S_1, U_1 the quantities $g_1 = U_1^{\top} x \in \mathbb{R}^2$ and $h_1 = U_1^{\top} p \in \mathbb{R}^2$, this yields that the differential equation (5.4) for $\Delta = U_1 S_1 U_1^T$ is equivalent to the following system (5.5) of differential equations. From (5.2) and (5.3), premultiplying by $U^{\top}$ and post-multiplying by U_1 we obtain $\dot{S}_1$; then post-multiplying by U_1 we obtain $\dot{U}_1$.

$$
\begin{aligned}
\dot{S}_1 &= -\lambda g_1 g_1^{\top} - \frac{1}{2}\left(g_1 h_1^{\top} + h_1 g_1^{\top}\right) + \varrho S_1, \\
\dot{U}_1 &= \left(-\lambda x g_1^{\top} - \frac{1}{2}\left(x h_1^{\top} + p g_1^{\top}\right) + U_1 \left(\lambda g_1 g_1^{\top} + \frac{1}{2}\left(g_1 h_1^{\top} + h_1 g_1^{\top}\right)\right)\right) S_1^{-1}.
\end{aligned} \tag{5.5}
$$

Similarly, for $\Theta = U_2 S_2 U_2^T$, setting $g_2 = U_2^{\top} u \in \mathbb{R}^2, h_2 = U_2^{\top} q \in \mathbb{R}^2$, we obtain the system of differential equations

$$
\begin{aligned}
\dot{S}_2 &= -\nu g_2 g_2^{\top} - \frac{1}{2}\left(g_2 h_2^{\top} + h_2 g_2^{\top}\right) + \varrho S_2, \\
\dot{U}_2 &= \left(-\nu u g_2^{\top} - \frac{1}{2}\left(u h_2^{\top} + q g_2^{\top}\right) + U_2 \left(\lambda g_2 g_2^{\top} + \frac{1}{2}\left(g_2 h_2^{\top} + h_2 g_2^{\top}\right)\right)\right) S_2^{-1}.
\end{aligned} \tag{5.6}
$$

Finally, for $\Gamma = U_3 S_3 U_3^T$ (with $S_3 \in \mathrm{Skew}^2$), setting $g_3 = U_3^{\top} w \in \mathbb{R}^2, h_3 = U_3^{\top} r \in \mathbb{R}^2$, we obtain the system of differential equations

$$
\begin{aligned}
\dot{S}_3 &= \frac{1}{2}\left(-g_3 h_3^{\top} + h_3 g_3^{\top}\right) + \varrho S_3, \\
\dot{U}_3 &= \left(\frac{1}{2}\left(-w h_3^{\top} + r g_3^{\top}\right) + U_3 \left(\frac{1}{2}\left(g_3 h_3^{\top} - h_3 g_3^{\top}\right)\right)\right) S_3^{-1}.
\end{aligned} \tag{5.7}
$$

Having established differential equations for rank-two factors, in the next section we discuss the monotonicity of the functional.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

5.2. Monotonicity of the functional. We have the following monotonicity result, which establishes that (5.4) is a gradient system for $F_\varepsilon^{\rm od}(\Delta(t),\Theta(t),\Gamma(t))$.

THEOREM 5.3. *Let* $\Delta(t),\Theta(t) \in \mathcal{M}_2^{{\rm Sym}^n}, \Gamma(t) \in \mathcal{M}_2^{{\rm Skew}^n}$ *satisfy the differential equation* (5.4)*, and suppose that*

$$F_\varepsilon^{\rm od}(\Delta(t),\Theta(t),\Gamma(t)) > 0.$$

If $\lambda(t)$ *is a simple eigenvalue of* $E+\varepsilon\Delta(t)$*,* $\mu(t)$ *is a simple eigenvalue of* $R+\varepsilon\Theta(t)$*, and* 0 *a simple eigenvalue of* $J+\varepsilon\Gamma$*, then*

$$\frac{d}{dt}F_\varepsilon^{\rm od}(\Delta(t),\Theta(t),\Gamma(t)) < 0.$$

Proof. We note that

$$\begin{aligned}
&\frac{1}{2\varepsilon}\frac{\mathrm{d}}{\mathrm{dt}}F_\varepsilon^{\rm od}(\Delta,\Theta,\Gamma)\\
&\quad = \lambda\,\langle xx^\top,\dot\Delta\rangle + \nu\,\langle uu^\top,\dot\Theta\rangle + \Big(\Big\langle px^\top,\dot\Delta\Big\rangle + \Big\langle qu^\top,\dot\Theta\Big\rangle + \Big\langle rw^\top,\dot\Gamma\Big\rangle\Big)\\
&\quad = -\Big\langle\,(\lambda x+p)\;x^\top, -P_\Delta^{\rm Sym}\Big({\rm Sym}\left((\lambda x+p)\;x^\top\right)\Big)\Big\rangle + \varrho\langle(\lambda x+p)\;x^\top,\Delta\rangle\\
&\qquad\qquad -\Big\langle\,(\nu u+q)\;u^\top, -P_\Theta^{\rm Sym}\Big({\rm Sym}\left((\nu u+q)\;u^\top\right)\Big)\Big\rangle + \varrho\langle(\nu u+q)\;u^\top,\Theta\rangle\\
&\qquad\qquad -\Big\langle r\,w^\top, -P_\Gamma^{\rm Skew}\Big({\rm Skew}\left(r\,w^\top\right)\Big)\Big\rangle + \varrho\langle r\,w^\top,\Theta\rangle\\
&\quad = \varrho^2 - \Big\|P_\Delta^{\rm Sym}\Big({\rm Sym}\left((\lambda x+p)\;x^\top\right)\Big)\Big\|^2 - \Big\|P_\Theta^{\rm Sym}\Big({\rm Sym}\left((\nu u+q)\;u^\top\right)\Big)\Big\|^2\\
&\qquad\qquad - \Big\|P_\Gamma^{\rm Skew}\Big({\rm Skew}\left(r\,w^\top\right)\Big)\Big\|^2,
\end{aligned}$$

with

$$\begin{aligned}
\varrho = {}& \Big\langle\Delta, P_\Delta^{\rm Sym}\Big({\rm Sym}\left((\lambda x+p)\;x^\top\right)\Big)\Big\rangle + \Big\langle\Theta, P_\Theta^{\rm Sym}\Big({\rm Sym}\left((\nu u+q)\;u^\top\right)\Big)\Big\rangle\\
&+\Big\langle\Gamma, P_\Gamma^{\rm Skew}\Big({\rm Skew}\left(r\,w^\top\right)\Big)\Big\rangle.
\end{aligned}$$

Applying the Cauchy-Schwarz inequality proves the assertion. ∎

5.3. Computational approach. For the numerical approximation of the rank-two ODEs, we use Algorithm 1, an adaptation of the method proposed in [10].

Here we solve the system

$$\dot K(t) = F\left(t, K(t)\,U_0^\top\right)\,U_0$$

on the interval $[t_0,t_1]$ by one step of the explicit Euler method.

Similarly, we make use of one Euler step to approximate the solution at time t_1 of the ODE system

$$\dot S(t) = U_1^\top\,F\left(t, U_1 S(t) U_1^\top\right).$$

If n is large, then the memory requirement and the computing time are significantly reduced with respect to the integration of the full-rank ODEs.

Algorithm 1: Low-rank symmetry/skew-symmetry preserving integrator.

Data: Matrix $X_0 = U_0 S_0 U_0^\top$, $F(t,X)$, t_0, t_1, h
Result: Matrix $X_1 = U_1 S_1 U_1^\top$
begin

1. Solve the $n \times 2$ ODE $\dot{K}(t) = F\left(t, K(t)\, U_0^\top\right) U_0,\ K(t_0) = U_0 S_0$.
2. Compute a QR-decomposition $K(t_1) = U_1\, R$.
3. Integrate the 2×2 ODE $\dot{S}(t) = U_1^\top\, F\left(t, U_1 S(t) U_1^\top\right)$ with initial value $S(t_0) = U_1^\top X_0 U_1 = \left(U_1^\top U_0\right) X_0 \left(U_1^\top U_0\right)^\top$.
4. Set $S_1 = S(t_1)/\|S(t_1)\|$,(normalization).
5. Return S_1, U_1

5.4. An illustrative example. Using the example of Section 4.4, integrating equations (5.5), (5.6), and (5.7), we obtain the same distance ε^* and a common null vector c of the same accuracy as when integrating (4.5), i.e.,

$$\varepsilon^* S_1 = \begin{bmatrix} -0.2722 & 0 \\ 0 & 0.0719 \end{bmatrix}, \quad \varepsilon^* S_2 = \begin{bmatrix} -0.2053 & 0 \\ 0 & 0.0205 \end{bmatrix}, \quad \varepsilon^* S_3 = \begin{bmatrix} 0 & 0.0529 \\ -0.0529 & 0 \end{bmatrix},$$

and

$$U_1 = \begin{bmatrix} -0.2291 & 0.0737 \\ 0.5521 & -0.3270 \\ 0.0449 & 0.0453 \\ 0.0682 & 0.9264 \\ 0.7975 & 0.1658 \end{bmatrix}, \quad U_2 = \begin{bmatrix} -0.3685 & -0.3829 \\ 0.4904 & -0.5782 \\ 0.0689 & 0.1841 \\ -0.5685 & -0.5870 \\ 0.5439 & -0.3750 \end{bmatrix}, \quad U_3 = \begin{bmatrix} -0.8360 & 0.2633 \\ -0.1780 & -0.6387 \\ 0.4867 & -0.0431 \\ 0.1176 & 0.3495 \\ -0.1367 & -0.6315 \end{bmatrix}.$$

6. Gradient flow for even state dimension. The derivation of the gradients in the case when the space dimension is even is slightly more complicated since in this case the skew-symmetric matrix J is not guaranteed to have a zero eigenvalue.

6.1. Computation of the gradient of the functional (3.4). Similarly to the odd case we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mu^2 = -\,\varepsilon\,\mu\left\langle \mathrm{Im}\left(ww^H\right), \dot{\Gamma}\right\rangle = \varepsilon\,\mu\,\left\langle\left(\mathrm{Re}(w)\,\mathrm{Im}(w)^\top - \mathrm{Im}(w)\,\mathrm{Re}(w)^\top\right), \dot{\Gamma}\right\rangle.$$

Considering orthogonal projections with respect to the Frobenius inner product onto the respective matrix manifolds Sym^n, Skew^n, we identify the constrained gradient directions of the terms associated to eigenvalues as

$$\dot{\Delta} \propto \lambda\, xx^\top, \qquad \dot{\Theta} \propto \nu\, uu^\top, \qquad \dot{\Gamma} \propto \mu\,\left(\mathrm{Re}(w)\,\mathrm{Im}(w)^\top - \mathrm{Im}(w)\,\mathrm{Re}(w)^\top\right).$$

Different to the odd case we have to consider

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left((x^\top \mathrm{Re}(w))^2\right) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(x^\top \mathrm{Re}(w)\mathrm{Re}(w)^\top x\right) \\ &= x^\top \mathrm{Re}(w)\mathrm{Re}(w)^\top \dot{x} + \mathrm{Re}(w)^\top x x^\top \mathrm{Re}(\dot{w}), \end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and thus

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(-(x^\top \mathrm{Re}(w))^2\right)\\
&\quad = \varepsilon\left((x^\top \mathrm{Re}(w))\mathrm{Re}(w)^\top L\dot{\Delta}x + (\mathrm{Re}(w)^\top x)x^\top\left(\frac{P\dot{\Gamma}w+\overline{P}\dot{\Gamma}\overline{w}}{2}\right)\right)\\
&\quad = \varepsilon\left(\left\langle \eta\, L^\top \mathrm{Re}(w)x^\top, \dot{\Delta}\right\rangle + \left\langle \eta\,\mathrm{Re}(P^H x w^H), \dot{\Gamma}\right\rangle\right),
\end{aligned}$$

where $\eta = x^\top \mathrm{Re}(w)$ and P is the pseudoinverse of $J+\varepsilon\Gamma - \mathbf{i}\mu I$. Analogously,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(-(x^\top \mathrm{Im}(w))^2\right) = \varepsilon\left(\left\langle \zeta\, L^\top \mathrm{Im}(w)x^\top, \dot{\Delta}\right\rangle + \left\langle \zeta\,\mathrm{Im}(P^H x w^H), \dot{\Gamma}\right\rangle\right),$$

where $\zeta = x^\top \mathrm{Im}(w)$. Introduce $w_{\mathrm{r}} := \mathrm{Re}(w)$ and $w_{\mathrm{i}} := \mathrm{Im}(w)$. In order to compute the steepest descent direction, we minimize the gradient of $F_\varepsilon^{\mathrm{ev}}$ and collect the summands involving $\dot{\Delta}$, $\dot{\Theta}$ and those involving $\dot{\Gamma}$. This yields

$$\frac{\mathrm{d}}{\mathrm{d}t}F_\varepsilon^{\mathrm{ev}}(\Delta,\Theta) = \varepsilon\,\langle(\lambda x+p)\ x^\top,\dot{\Delta}\rangle + \varepsilon\,\langle(\nu u+q)\ u^\top,\dot{\Theta}\rangle + \varepsilon\,\langle W+\eta\,\mathrm{Re}(H)+\zeta\,\mathrm{Im}(H),\dot{\Gamma}\rangle,$$

with

$$\begin{aligned}
p &= \theta L^\top u\ + L^\top(\eta w_{\mathrm{r}} + \zeta w_{\mathrm{i}}), \qquad & W &= w_{\mathrm{r}} w_{\mathrm{i}}^\top - w_{\mathrm{i}} w_{\mathrm{r}}^\top,\\
q &= \theta N^\top x, & H &= P^H x w^H.
\end{aligned}$$

Taking into consideration projection with respect to the Frobenius inner product of the vector field onto the manifolds of symmetric and skew-symmetric matrices, this leads to the system of differential equations,

$$\begin{aligned}
\dot{\Delta} &= -\mathrm{Sym}\left((\lambda x+p)\ x^\top\right) + \varrho\Delta,\\
\dot{\Theta} &= -\mathrm{Sym}\left((\nu u+q)\ u^\top\right) + \varrho\Theta,\\
\dot{\Gamma} &= -\mathrm{Skew}\left(W+\eta\,\mathrm{Re}(H)+\zeta\,\mathrm{Im}(H)\right) + \varrho\Gamma,
\end{aligned}$$

where

$$\begin{aligned}
\varrho = \big(&\langle\Delta, \mathrm{Sym}\left((\lambda x+p)\,x^\top\right)\rangle + \langle\Theta, \mathrm{Sym}\left((\nu u+q)\,u^\top\right)\rangle\\
&+\ \langle\Gamma, \mathrm{Skew}\left(W+\eta\,\mathrm{Re}(H)+\zeta\,\mathrm{Im}(H)\right)\rangle\big)
\end{aligned}$$

is again used to ensure the norm conservation. In this way we have again obtained a structured flow with matrices in Sym^n and Skew^n, respectively.

REMARK 6.1. Similarly to the odd case it is possible to derive a rank-two gradient system and obtain a more effective numerical integration.

6.2. A unifying functional. One may also try to construct a unifying function that treats the odd and even dimension cases together. For this we denote by x the eigenvector associated with λ, the smallest eigenvalue of $E+\varepsilon\Delta$, with the goal to make this the common null vector in the end. Introduce the alternative functional

$$\widetilde{F}_\varepsilon(\Delta,\Theta,\Gamma) = \frac{1}{2}\Big(\lambda^2 + \nu^2 + \|\,(R+\varepsilon\Theta)\,x\|_2^2 + \|\,(J+\varepsilon\Gamma)\,x\|_2^2\Big)$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$$= \frac{1}{2}\Big(\lambda^2 + \nu^2 + x^\top \left(R + \varepsilon\Theta\right)^2 x - x^\top \left(J + \varepsilon\Gamma\right)^2 x\Big),$$

with $\| \left(\Delta, \Gamma, \Theta\right) \|_F = 1$.

We observe that

$$\begin{aligned}
\frac{1}{2\varepsilon}\frac{\mathrm{d}}{\mathrm{d}t}\left(x^\top \left(R + \varepsilon\Theta\right)^2 x\right) &= -x^\top \left(R + \varepsilon\Theta\right)^2 L\dot{\Delta}x + x^\top \left(R + \varepsilon\Theta\right)\dot{\Theta}x \\
&= -\,\langle L\left(R + \varepsilon\Theta\right)^2 xx^\top, \dot{\Delta}\rangle + \langle \left(R + \varepsilon\Theta\right) xx^\top, \dot{\Theta}\rangle,
\end{aligned}$$

and similarly,

$$\begin{aligned}
\frac{1}{2\varepsilon}\frac{\mathrm{d}}{\mathrm{d}t}\left(x^\top \left(J + \varepsilon\Gamma\right)^2 x\right) &= -\,x^\top \left(J + \varepsilon\Gamma\right)^2 L\dot{\Delta}x + x^\top \left(J + \varepsilon\Gamma\right)\dot{\Gamma}x \\
&= -\,\langle L\left(J + \varepsilon\Gamma\right)^2 xx^\top, \dot{\Delta}\rangle + \langle \left(J + \varepsilon\Gamma\right) xx^\top, \dot{\Gamma}\rangle.
\end{aligned}$$

This leads to the system of ODEs

$$\begin{aligned}
\dot{\Delta} &= -\mathrm{Sym}\left(\left(\lambda x + s\right) x^\top\right) + \varrho\Delta, \\
\dot{\Theta} &= -\mathrm{Sym}\left(\nu u\, u^\top + r\, x^\top\right) + \varrho\Theta, \\
\dot{\Gamma} &= \mathrm{Skew}\left(z\, x^\top\right) + \varrho\Gamma,
\end{aligned}$$

where

$$\varrho = \left(\langle \Delta, \mathrm{Sym}\left(\left(\lambda x + s\right) x^\top\right)\rangle + \langle \Theta, \left(\nu u\, u^\top + \mathrm{Sym}\left(t\, x^\top\right)\right)\rangle - \langle \Gamma, \mathrm{Skew}\left(z\, x^\top\right)\rangle\right)$$

and

$$\begin{aligned}
s &= L\Big(\left(J + \varepsilon\Gamma\right)^2 - \left(R + \varepsilon\Theta\right)^2\Big)x, \\
r &= \left(R + \varepsilon\Theta\right)x, \\
z &= \left(J + \varepsilon\Gamma\right)x.
\end{aligned}$$

Although this functional appears simpler to manage and does not require the computation of two pseudoinverses, our experiments seem to indicate that with the previously considered functionals a higher accuracy can be reached.

7. The outer iteration for ε. In this section we present the outer iteration which is aimed to compute the optimal value ε, that is, the norm of the smallest perturbation which determines a common null vector of the considered triplet. An illustrative picture is given in Figure 7.1, which has been computed for the function $f^{\mathrm{ev}}(\varepsilon)$ of the first test example of Section 7.1.

The simplest way to do this is by means of a bisection technique as in Algorithm 2.

7.1. Illustrative examples. To illustrate the performance of the described algorithm, we consider first a scalable linear mass-spring-damper system that has been used as a model reduction test case in [20]. This test case generates matrices $M = M^T \geq 0$, $G = -G^T$, $D = D^T \geq 0$, $K = K^T > 0$ and, in a first-order formulation, leads to a dH pencil

$$\lambda \underbrace{\begin{bmatrix} K & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{E} - \left(\underbrace{\begin{bmatrix} 0 & K & 0 \\ -K & 0 & -G^T \\ 0 & G & 0 \end{bmatrix}}_{J} - \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{R} \right).$$

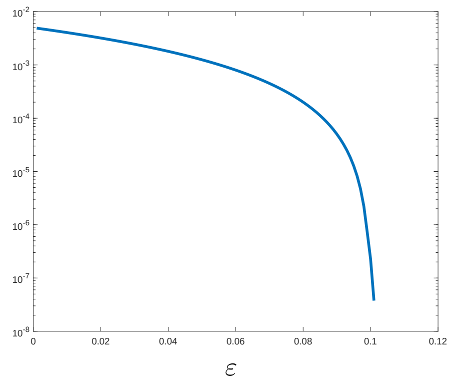

FIG. 7.1. *The function $f^{\mathrm{ev}}(\varepsilon)$ for the mass-spring-damper example.*

Algorithm 2: Bisection method for distance approximation.

Data: Matrices E, R, J, $k_{\max}$ (max number of iterations), δ, and tolerance tol
ε_0, $\varepsilon_{\mathrm{lb}}$, and $\varepsilon_{\mathrm{ub}}$ (starting values for the lower and upper bounds for ε^*)
Result: ε_δ (upper bound for the distance), $\Delta(\varepsilon^*), \Theta(\varepsilon^*), \Gamma(\varepsilon^*)$
begin
- Compute $\Delta(\varepsilon_0), \Theta(\varepsilon_0), \Gamma(\varepsilon_0)$.
- Compute $f(\varepsilon_0)$.
- Set $k = 0$.
- **while** $k \le 1$ *or* $|\varepsilon_{\mathrm{ub}} - \varepsilon_{\mathrm{lb}}| > \mathrm{tol}$ **do**
  - **if** $f(\varepsilon_k) < \mathrm{tol}$ **then**
    - Set $\varepsilon_{\mathrm{ub}} = \min(\varepsilon_{\mathrm{ub}}, \varepsilon_k)$.
  - **else**
    - Set $\varepsilon_{\mathrm{lb}} = \max(\varepsilon_{\mathrm{lb}}, \varepsilon_k)$.
  - Compute $\varepsilon_{k+1} = (\varepsilon_{\mathrm{lb}} + \varepsilon_{\mathrm{ub}})/2$ (bisection step).
  - **if** $k = k_{\max}$ **then**
    - Return interval $[\varepsilon_{\mathrm{lb}}, \varepsilon_{\mathrm{ub}}]$.
    - Halt.
  - **else**
    - Set $k = k + 1$.
  - Compute $\Delta(\varepsilon_k), \Theta(\varepsilon_k), \Gamma(\varepsilon_k)$.
  - Compute $f(\varepsilon_k)$.
- Return $\varepsilon^* = \varepsilon_k$.

This pencil is regular and of index two. If one puts γI in the $(3,3)$-block of the matrix R, then the distance to index 2 and instability is γ. Choosing the dimension $N = 100$ we obtain matrices $E, R, J \in \mathbb{R}^{3N+1}$.

We fix $\gamma = 10^{-1}$. The plot of the function $f(\varepsilon)$ obtained by integrating (5.4) for increasing ε is given in Figure 7.1. As a second example we consider the linear electrical circuit (1.2) from [29] (see Figure 7.2). We choose the values $L = 2$, $C_1 = 0.01$, $C_2 = 0.02$, $R_L = 0.1$, $R_G = 6$, $R_R = 3$, and $E_G = 1$. If we let J unchanged, then we find—as expected—$\Delta E \approx 0$ and

$$\Delta R = \mathrm{diag}(0, -C_1, 0, 0, 0) = \mathrm{diag}(0, -0.01, 0, 0, 0).$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

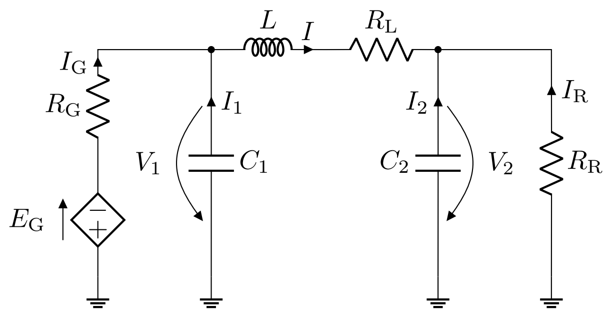

FIG. 7.2. *Simple DC power network example from* [29].

If we vary also J, then we find $\varepsilon^* = 1.46\ldots$ and

$$\Delta E = \begin{bmatrix} -0.0227 & -0.0696 & -0.0715 & -0.0302 & 0.0072 \\ -0.0696 & 0.0019 & -0.0067 & 0.0013 & -0.0004 \\ -0.0715 & -0.0067 & -0.0084 & 0.0007 & 0.0001 \\ -0.0302 & 0.0013 & 0.0007 & 0.0085 & -0.0002 \\ 0.0072 & -0.0004 & 0.0001 & -0.0002 & 0.0059 \end{bmatrix},$$

$$\Delta R = \begin{bmatrix} -0.0011 & -0.0034 & -0.0033 & -0.0088 & 0.0068 \\ -0.0034 & 0.0036 & -0.0005 & -0.0449 & 0.0551 \\ -0.0033 & -0.0005 & 0.0036 & -0.0438 & 0.0581 \\ -0.0088 & -0.0449 & -0.0438 & -0.0035 & 0.0026 \\ 0.0068 & 0.0551 & 0.0581 & 0.0026 & -0.0042 \end{bmatrix},$$

$$\Delta J = \begin{bmatrix} 0 & -0.0142 & -0.0090 & 0.1104 & 0.1130 \\ 0.0142 & 0 & -0.0460 & 0.4787 & 0.4866 \\ 0.0090 & 0.0460 & 0 & 0.4925 & 0.5148 \\ -0.1104 & -0.4787 & -0.4925 & 0 & 0.2078 \\ -0.1130 & -0.4866 & -0.5148 & -0.2078 & 0 \end{bmatrix}.$$

Conclusions and further work. We have investigated a structured distance problem related to the study of dissipative Hamiltonian systems that is determining the closest triplet of matrices to a given one, sharing a common null-space. The approach is based on the numerical integration of suitable gradient systems derived for an associated functional. A remarkable low-rank property of extremizers leads to low-rank ODEs which are appealing from a computational and storage point of view. Extensions to related problems can be derived with a similar approach. A Matlab package with the algorithms to estimate the distance to singularity based on the procedures presented in this paper is available at the webpage of the authors.

Acknowledgments. We would like to thank both referees for their careful reading and for their many helpful comments and remarks.

The first author acknowledges that his research was supported by funds from the Italian MUR (Ministero dell'Università e della Ricerca) within the PRIN 2017 Project "Discontinuous dynamical systems: theory, numerics and applications" and by the INdAM Research group GNCS (Gruppo Nazionale di Calcolo Scientifico). The Matlab routines implementing full ODE integration in the odd and even cases are available at the web-

ETNA
Kent State University and
Johann Radon Institute (RICAM)

page https://www.gssi.it/people/professors/lectures-maths/item/545-guglielmi-nicola. In order to use them it is necessary to download the package psapsr (see [19]) from the webpage https://cs.nyu.edu/~overton/software/psapsr/.

The second author thanks Deutsche Forschungsgemeinschaft (DFG) for support within the project B03 in CRC TRR 154.

REFERENCES

[1] F. ACHLEITNER, A. ARNOLD, AND V. MEHRMANN, *Hypocoercivity and controllability in linear semi-dissipative ODEs and DAEs*, ZAMM Z. Angew. Math. Mech., in press, 2022. https://doi.org/10.1002/zamm.202100171

[2] N. ALIYEV, P. BENNER, E. MENGI, P. SCHWERDTNER, AND M. VOIGT, *Large-scale computation of L_∞-norms by a greedy subspace method*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1496–1516.

[3] N. ALIYEV, V. MEHRMANN, AND E. MENGI, *Approximation of stability radii for large-scale dissipative Hamiltonian systems*, Adv. Comput. Math., 46 (2020), Art. 6, 28 pages.

[4] R. ALTMANN, V. MEHRMANN, AND B. UNGER, *Port-Hamiltonian formulations of poroelastic network models*, Preprint on arXiv, 2020. http://arxiv.org/abs/2012.01949.

[5] C. BEATTIE, V. MEHRMANN, H. XU, AND H. ZWART, *Linear port-Hamiltonian descriptor systems*, Math. Control Signals Systems, 30 (2018), Art. 17, 27 pages.

[6] C. BEATTIE, V. MEHRMANN, AND P. VAN DOOREN, *Robust port-Hamiltonian representations of passive systems*, Automatica J. IFAC, 100 (2019), pp. 182–186.

[7] T. BERGER, H. GERNANDT, C. TRUNK, H. WINKLER, M. WOJTYLAK, *A new bound for the distance to singularity of a regular matrix pencil*, PAMM. Proc. Appl. Math. Mech., 17 (2017), pp. 863–864.

[8] ———, *The gap distance to the set of singular matrix pencils*, Linear Algebra Appl., 564 (2019), pp. 28–57.

[9] R. BYERS, C. HE, AND V. MEHRMANN, *Where is the nearest non-regular pencil?*, Linear Algebra Appl., 285 (1998), pp. 81–105.

[10] G. CERUTI AND C. LUBICH, *Time integration of symmetric and anti-symmetric low-rank matrices and Tucker tensors*, BIT, 60 (2020), pp. 591–614.

[11] N. H. DU, V. H. LINH, AND V. MEHRMANN, *Robust stability of differential-algebraic equations*, in Surveys in Differential-Algebraic Equations. I, A. Ilchmann and T. Reis, eds., Differ.-Algebr. Equ. Forum, Springer, Heidelberg, 2013, pp. 63–95.

[12] H. EGGER, T. KUGLER, B. LILJEGREN-SAILER, N. MARHEINEKE, AND V. MEHRMANN, *On structure-preserving model reduction for damped wave propagation in transport networks*, SIAM J. Sci. Comput., 40 (2018), pp. A331–A365.

[13] A. FAZZI, N. GUGLIELMI, AND C. LUBICH, *Finding the nearest passive or nonpassive system via Hamiltonian eigenvalue optimization*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 1553–1580.

[14] F. R. GANTMACHER, *Theory of Matrices. Vol. 1*, Chelsea, New York, 1959.

[15] N. GILLIS, V. MEHRMANN, AND P. SHARMA, *Computing the nearest stable matrix pairs*, Numer. Linear Algebra Appl., 25 (2018), Art. e2153, 16 pages.

[16] N. GILLIS AND P. SHARMA, *On computing the distance to stability for matrices using linear dissipative Hamiltonian systems*, Automatica J. IFAC, 85 (2017), pp. 113–121.

[17] ———, *Finding the nearest positive-real system*, SIAM J. Numer. Anal., 56 (2018), pp. 1022–1047.

[18] N. GUGLIELMI, C. LUBICH, AND V. MEHRMANN, *On the nearest singular matrix pencil*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 776–806.

[19] N. GUGLIELMI AND M. L. OVERTON, *Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1166–1192.

[20] S.-A. HAUSCHILD, N. MARHEINEKE, AND V. MEHRMANN, *Model reduction techniques for linear constant coefficient port-Hamiltonian differential-algebraic systems*, Control Cybernet., 48 (2019), pp. 125–152.

[21] B. JACOB AND H. J. ZWART, *Linear Port-Hamiltonian Systems on Infinite-Dimensional Spaces*, Birkhäuser, Basel, 2012.

[22] T. KATO, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1995.

[23] O. KOCH AND C. LUBICH, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454.

[24] C. MEHL, V. MEHRMANN, AND P. SHARMA, *Stability radii for linear Hamiltonian systems with dissipation under structure-preserving perturbations*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 1625–1654.

[25] ———, *Stability radii for real linear Hamiltonian systems with perturbed dissipation*, BIT, 57 (2017), pp. 811–843.

[26] C. MEHL, V. MEHRMANN, AND M. WOJTYLAK, *On the distance to singularity via low rank perturbations*, Oper. Matrices, 9 (2015), pp. 733–772.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[27] ———, *Linear algebra properties of dissipative Hamiltonian descriptor systems*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1489–1519.

[28] ———, *Distance problems for dissipative Hamiltonian systems and related matrix polynomials*, Linear Algebra Appl., 623 (2021), pp. 335–366.

[29] V. MEHRMANN AND R. MORANDIN, *Structure-preserving discretization for port-Hamiltonian descriptor systems*, in 58th IEEE Conf. Decision and Control (CDC), IEEE Conference Proceedings, Los Alamitos, 2019, pp. 6863–6868.

[30] V. MEHRMANN AND P. VAN DOOREN, *Optimal robustness of port-Hamiltonian systems*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 134–151.

[31] C. MEYER AND G. W. STEWART, *Derivatives and perturbations of eigenvectors*, SIAM J. Numer. Anal., 25 (1988), pp. 679–691.

[32] M. L. OVERTON, *HANSO: Hybrid Algorithm for Non-Smooth Optimization*, Software, 2020. http://www.cs.nyu.edu/overton/software/hanso

[33] A. J. VAN DER SCHAFT, *Port-Hamiltonian Differential-Algebraic Systems*, in Surveys in Differential-Algebraic Equations. I, A. Ilchmann and T. Reis, eds., Differ.-Algebr. Equ. Forum, Springer, Heidelberg, 2013, pp. 173–226.

[34] A. J. VAN DER SCHAFT AND B. MASCHKE, *Hamiltonian formulation of distributed-parameter systems with boundary energy flow*, J. Geom. Phys., 42 (2002), pp. 166–194.

[35] ———, *Generalized port-Hamiltonian DAE systems*, Systems Control Lett., 121 (2018), pp. 31–37.

[36] A. J. VAN DER SCHAFT AND D. JELTSEMA, *Port-Hamiltonian systems theory: an introductory overview*, Found. and Trends in Systems and Control, 1 (2014), pp. 173–378.