RANK-1 MATRIX DIFFERENTIAL EQUATIONS FOR STRUCTURED EIGENVALUE OPTIMIZATION

NICOLA GUGLIELMI*, CHRISTIAN LUBICH‡, AND STEFANO SICILIA*

Abstract. A new approach to solving eigenvalue optimization problems for large structured matrices is proposed and studied. The class of optimization problems considered is related to computing structured pseudospectra and their extremal points, and to structured matrix nearness problems such as computing the distance to instability or to singularity under structured perturbations. The perturbation structure can be a general linear structure. We focus on the practically important cases of large matrices with a given sparsity pattern and on perturbation matrices with given range and co-range. It is known that analogous eigenvalue optimization for *unstructured* complex matrices favorably works with rank-1 matrices. The novelty of the present paper is that *structured* eigenvalue optimization can still be performed with rank-1 matrices, which yields a significant reduction of storage and in some cases of the computational cost. Optimizers are shown to be rank-1 matrices orthogonally projected onto the given structure. This is used in the numerical algorithms.

The numerical approach comes in two different versions. If the rank-1 matrices projected onto the structure form a manifold and if the orthogonal projection onto the corresponding tangent spaces is known and computationally inexpensive, as is the case for matrices with given range and co-range, then the method relies on a gradient system on this manifold. Otherwise, in particular for matrices with a given sparsity pattern, the method relies on the orthogonal projection of the free gradient onto the tangent space of the manifold of complex rank-1 matrices. The solution of this rank-1 system is then projected onto the structure. It is shown that there is a bijective correspondence between the stationary points of the gradient system on the structure space and of the rank-1 system. Near a local minimizer the rank-1 tangent projection is very close to the identity map, and so the computationally favorable rank-1 projected system behaves locally like the gradient system. Numerical experiments illustrate the two approaches for large matrices with a given sparsity pattern and for perturbation matrices with given range and co-range.

Key words. Structured matrix nearness problems, structured pseudospectrum, pseudospectral abscissa, pseudospectral radius, rank-1 perturbations, low-rank dynamics, gradient system.

AMS subject classifications. 15A18, 65F15

1. Introduction.

We describe an approach to structured eigenvalue optimization problems that uses constrained gradient flows and the underlying rank-1 property of the optimizers. We illustrate basic techniques on a class of problems that arise in computing structured pseudospectra or their extremal points and appear as the essential algorithmic building block in structured matrix nearness problems. For example, we determine the largest possible spectral abscissa or radius of a given matrix under perturbations of a prescribed norm that preserve its structure, or - in other words - the structured pseudospectral abscissa or radius. This is an important subtask in the computation of structured stability radii (or structured distance to instability in another terminology). In the literature these quantities are extensively studied with the purpose of analyzing stability properties and robustness of linear dynamical systems (see, e.g., [16]). Similarly, if one is interested in the distance of a matrix to singularity, the unstructured distance is the smallest singular value. However, if the matrix is structured, having a small singular value does not imply the existence of a small structured perturbation that makes it singular, and the structured distance to singularity is not readily obtained.

*Division of Mathematics, Gran Sasso Science Institute, Via Crispi 7, I-67100 L' Aquila, Italy. Email: nicola.guglielmi@gssi.it, stefano.sicilia@gssi.it

‡Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D–72076 Tübingen, Germany. Email: lubich@na.uni-tuebingen.de

The structures considered here are general complex- or real-linear structures, that is, the perturbation matrices are restricted to lie in a structure space $\mathcal{S}$, which can be an arbitrary linear subspace of $\mathbb{C}^{n,n}$ or $\mathbb{R}^{n,n}$. We will put the focus on two very different classes of major interest in applications,

(i) perturbation matrices with a given sparsity pattern and

(ii) perturbation matrices with given range and co-range.

Instead of a direct discrete approach to solve the optimization problems, we present a continuous approach using structure- and norm-constrained gradient flows, which reveals the underlying rank-1 property of optimizers, on which we build our discrete optimization method. The rank-1 property is well-known for unstructured problems (see e.g. [27]) and has been exploited for developing suitable algorithms (see e.g. [14, 22, 10]). The rank-1 differential equation is finally fully discretized, using an appropriate time discretization (here chosen beyond mere gradient descent) and an adaptive, line search-type stepsize selection. We mention that there are several situations previously addressed in the literature where considering a time-continuous approach provides new insight, e.g. [4, 6, 17, 27, 1] and references therein. This list is far from exhaustive.

In previous works, structured eigenvalue optimization problems were addressed for some specific structures. For example when the matrices are required to be real (the unstructured problem would consider them as complex), it has been proved that the optimizers have a rank-2 structure [25] and indeed are obtained as real parts of an underlying rank-1 matrix [11]. Similarly, Hamiltonian eigenvalue optimization has been studied in detail in [24] and [2], in the setting of robust passivity analysis of linear control systems, where eigenvalues of Hamiltonian matrices have to be bounded away from the imaginary axis. In that case it is possible to show that for a real Hamiltonian matrix, extremal perturbations have rank 4 [9]. However, when considering for example a sparse matrix, the low-rank property of optimizers seems to be irremediably lost. *It is a basic goal of this article to uncover the underlying rank-*1 *property and to show how it can be used in algorithms for structured eigenvalue optimization.*

The paper is organized as follows. In Section 2 we set up the framework and present our approach, which is based on a structure- and norm-constrained gradient system. We show that optimizers are orthogonal projections of rank-1 matrices onto the given structure. We discuss the possibilities and difficulties of using a gradient system for structure-projected rank-1 matrices. This works well for the case (ii) of prescribed range and co-range, but is not feasible for the case (i) of a prescribed sparsity pattern. In Section 3 we introduce instead a differential equation on the manifold of rank-1 matrices of unit Frobenius norm, for which the stationary points are shown to be in a bijective correspondence with the stationary points of the structure- and norm-constrained gradient system. In Section 4 we prove local convergence to strong minima under an assumption that appears to be generically satisfied in case (i) of sparse matrices, but is not satisifed in case (ii) of perturbation matrices with prescribed range and co-range. A basic observation, valid for all cases, is that near a local minimizer, the rank-1 tangent projection is very close to the identity map, and so the computationally favorable rank-1 projected system behaves locally like the gradient system. In Section 5 we discretize the rank-1 differential equation by a splitting method. This leads us to a fully discrete algorithm that updates rank-1 matrices in every step. Then, in Section 6 we describe a two-level approach to compute the structured stability radius (or structured distance to instability), used to characterize robustness of spectral stability properties. This is an important use

of the considered class of eigenvalue optimization problems for solving structured matrix nearness problems. The structured distance to singularity is computed in an analogous way. In Section 7 we present some numerical examples showing that the rank-1 system is well-suited for the efficient computation of optimizers.

Finally, in Section 8 we show how the alternative approach of using the gradient system for structure-projected rank-1 matrices can be used for the case (ii) of prescribed range and co-range.

2. Structured eigenvalue optimization and constrained gradient flows. In this section we formulate and discuss a class of eigenvalue optimization problems that are related to structured pseudospectra. We derive and study structure- and norm-constrained gradient systems and their stationary points, which turn out to be structure-projected rank-1 matrices.

2.1. Problem formulation and motivation. For a matrix $A \in \mathbb{C}^{n,n}$, let $\lambda(A) \in \mathbb{C}$ be a target eigenvalue of A, for example:

(i) an eigenvalue of minimal or maximal real part;
(ii) an eigenvalue of minimal or maximal modulus;
(iii) a closest eigenvalue to a given set in the complex plane.

We note that here the target eigenvalue need not depend continuously on the matrix A when several eigenvalues are simultaneously extremal, but it depends continuously on A when the extremal eigenvalue is unique.

Let $\mathcal{S}$ be a subspace of the vector space of complex or real $n \times n$ matrices, e.g., a space of matrices with a prescribed sparsity pattern, or matrices with given range and co-range. We let

$$f : \mathbb{C}^2 \to \mathbb{C} \quad \text{with} \quad f(z, \overline{z}) = f(\overline{z}, z) \in \mathbb{R} \quad \text{for all } z \in \mathbb{C} \tag{2.1}$$

be a given smooth function that will be minimized over target eigenvalues $\lambda(A + \Delta)$ for structured perturbations $\Delta \in \mathcal{S}$ to a given matrix A. While our theory applies to general functions f with (2.1), in our examples we consider specific cases where

f or $-f$ evaluated at $(z, \overline{z})$ equals

$$\operatorname{Re} z = \frac{z + \overline{z}}{2} \quad \text{or} \quad |z|^2 = z\overline{z}.$$

As will be dicussed in more detail in Section 6, the real part function is used in studying the distance to instability (or stability radius) of a Hurwitz matrix, that is with all eigenvalues in the left complex half-plane. The interest is in computing the nearest matrix $A + \Delta$ to A for which the rightmost eigenvalue is on the imaginary axis. Here, the perturbation Δ will be constrained to be in the structure space $\mathcal{S}$, and "nearest" will refer to the Frobenius norm $\|\Delta\|_F$. Similarly, the squared modulus function is used when A is a Schur matrix, that is with all eigenvalues in the unit disk, to compute the nearest matrix $A + \Delta$ to A for which the rightmost eigenvalue is on the unit circle. The squared modulus function is also used to compute the structured distance to singularity of an invertible matrix.

We consider the following *structured eigenvalue optimization problem*: For a given perturbation size $\varepsilon > 0$, find

$$\arg \min_{\Delta \in \mathcal{S},\, \|\Delta\|_F = \varepsilon} f\left(\lambda(A + \Delta), \overline{\lambda}(A + \Delta)\right), \tag{2.2}$$

where $\|\Delta\|_F$ is the Frobenius norm of the structured matrix $\Delta \in \mathcal{S}$, i.e. the Euclidean norm of the vector of its entries and $\lambda(A+\Delta)$ is the considered target eigenvalue of the

perturbed matrix $A+\Delta$. The arg max case is treated analogously, replacing f by $-f$. This problem arises in computing extremal points of the *structured ε-pseudospectrum*

$$\Lambda_\varepsilon^{\mathcal{S}}(A) = \{\lambda \in \mathbb{C} : \lambda \text{ is an eigenvalue of } A+\Delta \text{ for some } \Delta \in \mathcal{S} \text{ with } \|\Delta\|_F \le \varepsilon\}.$$

For $-f(z,\overline{z}) = \operatorname{Re} z$, (2.2) yields the structured pseudospectral abscissa $\alpha_\varepsilon^{\mathcal{S}}(A) = \max\{\operatorname{Re} z : z \in \Lambda_\varepsilon^{\mathcal{S}}(A)\}$, and for $-f(z,\overline{z}) = |z|^2$ it yields the structured pseudospectral radius $\rho_\varepsilon^{\mathcal{S}}(A) = \max\{|z| : z \in \Lambda_\varepsilon^{\mathcal{S}}(A)\}$. The structured distance to instability is then obtained by finding the smallest $\varepsilon > 0$ such that $\alpha_\varepsilon^{\mathcal{S}}(A) = 0$ (for a Hurwitz matrix A) or $\rho_\varepsilon^{\mathcal{S}}(A) = 0$ (for a Schur matrix A).

In the following it is convenient to write

$$\Delta = \varepsilon E \quad \text{with } \|E\|_F = 1 \qquad \text{and} \qquad F_\varepsilon(E) = f\left(\lambda\left(A+\varepsilon E\right), \overline{\lambda}\left(A+\varepsilon E\right)\right)$$

so that Problem (2.2) is equivalent to the problem of finding

$$\arg \min_{E \in \mathcal{S}, \|E\|_F = 1} F_\varepsilon(E). \tag{2.3}$$

Problem (2.2) or (2.3) is a nonconvex, nonsmooth optimization problem.

In a variant to the above problem, the inequality constraints $\|\Delta\|_F \le \varepsilon$ and $\|E\|_F \le 1$ can also be considered in (2.2) and (2.3), respectively.

An extension to functions of several target eigenvalues is direct, but is not considered in this paper for brevity.

2.2. Projection onto the structure. In order to treat the above problem, we shall make use of a projection onto the structure space $\mathcal{S}$.

Given two complex $n \times n$ matrices, we denote by ($\operatorname{tr}(\cdot)$ denotes the trace)

$$\langle X, Y\rangle = \sum_{i,j} \overline{x}_{ij} y_{ij} = \operatorname{tr}(X^* Y)$$

the inner product in $\mathbb{C}^{n,n}$ that induces the Frobenius norm $\|X\|_F = \langle X, X\rangle^{1/2}$.

Let $\Pi^{\mathcal{S}}$ be the orthogonal projection (w.r.t. the Frobenius inner product) onto $\mathcal{S}$: for every $Z \in \mathbb{C}^{n,n}$,

$$\Pi^{\mathcal{S}} Z \in \mathcal{S} \quad \text{and} \quad \operatorname{Re}\langle \Pi^{\mathcal{S}} Z, W\rangle = \operatorname{Re}\langle Z, W\rangle \quad \forall\, W \in \mathcal{S}. \tag{2.4}$$

For a complex-linear subspace $\mathcal{S}$, taking the real part of the complex inner product can be omitted (because with $W \in \mathcal{S}$, then also $\mathrm{i}W \in \mathcal{S}$), but taking the real part is needed for real-linear subspaces. Note that for $\mathcal{S} = \mathbb{R}^{n,n}$, we then have $\Pi^{\mathcal{S}} Z = \operatorname{Re} Z$ for all $Z \in \mathbb{C}^{n,n}$. Our guiding examples in this paper are the following structures, where the stated action of $\Pi^{\mathcal{S}}$ is readily verified.

Example 2.1 (Sparse matrices). If $\mathcal{S}$ is the space of complex matrices with a prescribed sparsity pattern, then $\Pi^{\mathcal{S}} Z$ leaves the entries of Z on the sparsity pattern unchanged and annihilates those outside the sparsity pattern. If $\mathcal{S}$ is the space of real matrices with a prescribed sparsity pattern, then $\Pi^{\mathcal{S}} Z$ takes instead the real part of the entries of Z on the sparsity pattern.

Example 2.2 (Matrices with prescribed range and co-range). An example of particular interest in control theory is the perturbation space

$$\mathcal{S} = \{B\Delta C : \Delta \in \mathbb{C}^{k,l}\}, \tag{2.5}$$

where $B \in \mathbb{C}^{n,k}$ and $C \in \mathbb{C}^{l,n}$ with $k, l < n$ are given matrices of full rank. Here, $\Pi^{\mathcal{S}} Z = BB^\dagger Z C^\dagger C$, where $B^\dagger$ and $C^\dagger$ are the Moore–Penrose pseudo-inverses of B and C, respectively. In the real case, where B and C are given real matrices and Δ is required to be real, we have $\Pi^{\mathcal{S}} Z = BB^\dagger(\operatorname{Re} Z)C^\dagger C$.

2.3. Free gradient of the functional F_ε. To get the gradient of the functional F_ε, we need the derivative of the target eigenvalue $\lambda(A+\varepsilon E(t))$ along paths of matrices $E(t)$, for t in some interval. In the case of a simple eigenvalue, which is the situation we will consider in the following, this derivative is obtained from the following well-known result; see e.g. [19, Lemma 6.3.10 and Theorem 6.3.12] and [7, Theorem 1].

LEMMA 2.3 (Derivative of simple eigenvalues). *Consider a continuously differentiable path of square complex matrices $M(t)$ for t in an open interval I. Let $\lambda(t)$, $t \in I$, be a continuous path of simple eigenvalues of $M(t)$. Let $x(t)$ and $y(t)$ be left and right eigenvectors, respectively, of $M(t)$ to the eigenvalue $\lambda(t)$. Then, $x(t)^*y(t) \neq 0$ for $t \in I$ and λ is continuously differentiable on I with*

$$\dot\lambda = \frac{x^*\dot M y}{x^*y}, \tag{2.6}$$

where the dot indicates differentiation with respect to t.

Since we have $x(t)^*y(t) \neq 0$, we can apply the normalization

$$\|x(t)\| = 1, \quad \|y(t)\| = 1, \quad x(t)^*y(t) \text{ is real and positive.} \tag{2.7}$$

The norm $\|\cdot\|$ is chosen as the Euclidean norm, and $x^* = \overline{x}^\top$. Clearly, a pair of left and right eigenvectors x and y fulfilling (2.7) may be replaced by μx and μy for any complex μ of modulus 1 without changing the property (2.7).

The following lemma will allow us to compute the steepest descent direction of the functional F_ε in $\mathbb{C}^{n,n}$, which means neglecting any structural constraint. For this reason we refer to it as the *free gradient* of the functional.

LEMMA 2.4 (Free gradient). *Let $E(t) \in \mathbb{C}^{n,n}$, for t near t_0, be a continuously differentiable path of matrices, with the derivative denoted by $\dot E(t)$. Assume that $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$ depending continuously on t, with associated eigenvectors $x(t)$ and $y(t)$ satisfying* (2.7)*, and let the eigenvalue condition number be*

$$\kappa(t) = \frac{1}{x(t)^*y(t)} > 0.$$

Then, $F_\varepsilon(E(t)) = f\big(\lambda(t), \overline{\lambda(t)}\big)$ is continuously differentiable w.r.t. t and we have

$$\frac{1}{\varepsilon\kappa(t)}\,\frac{d}{dt}F_\varepsilon(E(t)) = \operatorname{Re}\big\langle G_\varepsilon(E(t)), \dot E(t)\big\rangle, \tag{2.8}$$

where the (rescaled) gradient of F_ε is the rank-1 matrix

$$G_\varepsilon(E) = 2f_{\overline\lambda}\, xy^* \in \mathbb{C}^{n,n} \qquad \text{with } f_{\overline\lambda} = \frac{\partial f}{\partial\overline\lambda}(\lambda,\overline\lambda) \tag{2.9}$$

for the target eigenvalue $\lambda = \lambda(A+\varepsilon E)$ and the corresponding left and right eigenvectors x and y normalized according to (2.7)*.*

Proof. We first observe that (2.1) implies

$$f_{\overline\lambda} = \overline{f_\lambda} = \overline{\frac{\partial f}{\partial\lambda}(\lambda,\overline\lambda)}.$$

Using Lemma 2.3, we obtain that $F_\varepsilon(E(t))$ is differentiable with

$$\frac{d}{dt}F_\varepsilon\left(E(t)\right) = f_\lambda\,\dot\lambda + f_{\overline\lambda}\,\overline{\dot\lambda} = \frac{\varepsilon}{x^*y}\left(f_\lambda\, x^*\dot E y + f_{\overline\lambda}\,\overline{x^*\dot E y}\right) = \frac{\varepsilon}{x^*y}\,2\operatorname{Re}\left(f_\lambda\, x^*\dot E y\right),$$

where we omit the omnipresent dependence on t on the right-hand side. Noting that

$$\mathrm{Re}\big(f_\lambda\, x^*\dot{E}y\big) = \mathrm{Re}\,\Big\langle \overline{f_\lambda}\, xy^*, \dot{E}\Big\rangle,$$

we obtain (2.8)–(2.9). ∎

EXAMPLE 2.5. For

$$f(\lambda,\overline{\lambda}) = -\tfrac{1}{2}(\lambda+\overline{\lambda}) = -\mathrm{Re}\,\lambda$$

we have $2f_{\overline{\lambda}} = -1$ and hence $G_\varepsilon(E) = -xy^*$, which is nonzero for all λ. For

$$f(\lambda,\overline{\lambda}) = -\tfrac{1}{2}|\lambda|^2 = -\tfrac{1}{2}\lambda\overline{\lambda}$$

we have $2f_{\overline{\lambda}} = -\lambda$. In this case $G_\varepsilon(E) = -\lambda\, xy^*$, which is nonzero whenever $\lambda \neq 0$.

2.4. Structure-projected gradient. The optimization problem (2.3) is set on the manifold $\mathcal{S}_1 = \{E \in \mathbb{C}^{n,n} \,:\, E \in \mathcal{S},\ \|E\|_F = 1\}$.

Preserving the structure. Consider a differentiable path of *structured* matrices $E(t)$ in the linear space $\mathcal{S}$. Since then also $\dot{E}(t) \in \mathcal{S}$, we have by Lemma 2.4

$$\frac{1}{\varepsilon\kappa(t)}\,\frac{d}{dt}F_\varepsilon(E(t)) = \mathrm{Re}\big\langle G_\varepsilon^{\mathcal{S}}(E(t)), \dot{E}(t)\big\rangle \tag{2.10}$$

with the rescaled structured gradient

$$G_\varepsilon^{\mathcal{S}}(E) := \Pi^{\mathcal{S}} G_\varepsilon(E) = \Pi^{\mathcal{S}}(2f_{\overline{\lambda}}\, xy^*) \in \mathcal{S}, \tag{2.11}$$

which is the projection onto $\mathcal{S}$ of a rank-1 matrix.

Preserving the unit norm. To fulfil the constraint $\|E(t)\|_F^2 = 1$, we must have

$$0 = \frac{1}{2}\,\frac{d}{dt}\|E(t)\|_F^2 = \mathrm{Re}\,\big\langle E(t), \dot{E}(t)\big\rangle. \tag{2.12}$$

In view of Lemma 2.4 we are thus led to the following constrained optimization problem for the admissible direction of steepest descent.

LEMMA 2.6 (Direction of steepest admissible descent). *Let $E \in \mathcal{S}$ with $\|E\|_F = 1$ and let $G \in \mathbb{C}^{n,n}$ be such that $G^{\mathcal{S}} := \Pi^{\mathcal{S}} G$ is not a real multiple of E. Then, the unique solution of the optimization problem*

$$\begin{aligned} Z_\star = \arg\min_{Z\in\mathbb{C}^{n,n}} \ & \mathrm{Re}\,\langle G, Z\rangle \\ \text{subject to} \quad & Z \in \mathcal{S} \\ \text{and} \quad & \mathrm{Re}\,\langle E, Z\rangle = 0 \\ \text{and} \quad & \|Z\|_F = 1 \qquad \text{(for uniqueness)} \end{aligned} \tag{2.13}$$

is given by

$$\mu Z_\star = -G^{\mathcal{S}} + \mathrm{Re}\,\langle G^{\mathcal{S}}, E\rangle\, E, \tag{2.14}$$

where μ is the Frobenius norm of the matrix on the right-hand side.

Proof. The result follows by noting that the real part of the complex inner product on $\mathbb{C}^{n,n}$ is a real inner product on $\mathbb{R}^{2n,2n}$, and the real inner product with a given vector (which here is a matrix) is maximized over a subspace by orthogonally projecting the vector onto that subspace. The expression in (2.14) is the orthogonal projection of $-G^{\mathcal{S}}$ to the orthogonal complement of the span of E, which is the tangent space at E of the manifold of matrices of unit Frobenius norm. Since $E, G^{\mathcal{S}} \in \mathcal{S}$ in (2.14), also $Z_\star$ is in $\mathcal{S}$. ∎

2.5. Structure- and norm-constrained gradient flow. Lemmas 2.4 and 2.6 show that the admissible direction of steepest descent of the functional F_ε at a matrix $E \in \mathcal{S}$ of unit Frobenius norm is given by the positive multiples of the matrix $-G_\varepsilon^{\mathcal{S}}(E) + \operatorname{Re}\langle G_\varepsilon^{\mathcal{S}}(E), E\rangle E$. This leads us to consider the (rescaled) *gradient flow on the manifold $\mathcal{S}_1$ of matrices in the structure space $\mathcal{S}$ of unit Frobenius norm*:

$$\dot{E} = -G_\varepsilon^{\mathcal{S}}(E) + \operatorname{Re}\langle G_\varepsilon^{\mathcal{S}}(E), E\rangle E. \tag{2.15}$$

By construction of this ordinary differential equation, we have that $\dot{E} \in \mathcal{S}$ for $E \in \mathcal{S}$ and $\operatorname{Re}\langle E, \dot{E}\rangle = 0$ along its solutions, and so both the structure $\mathcal{S}$ and the Frobenius norm 1 are conserved.

Monotonicity. As we follow the admissible direction of steepest descent of the functional F_ε along solutions $E(t)$ of the ODE (2.15), we obtain the following.

THEOREM 2.7 (Monotonicity). *Assume that $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$ and that $\lambda(\cdot)$ is continuous at t. Let $E(\cdot)$ of unit Frobenius norm satisfy the differential equation* (2.15)*. Then,*

$$\frac{d}{dt} F_\varepsilon(E(t)) \le 0. \tag{2.16}$$

Proof. Although the result follows directly from Lemmas 2.4 and 2.6, we compute the derivative directly.

We write $G = G_\varepsilon(E)$ for short and take the inner product of (2.15) with $\dot{E}$. Using that $\operatorname{Re}\langle E, \dot{E}\rangle = 0$, we find

$$\|\dot{E}\|_F^2 = -\operatorname{Re}\langle G - \operatorname{Re}\langle G, E\rangle E, \dot{E}\rangle = -\operatorname{Re}\langle G, \dot{E}\rangle$$

and hence Lemma 2.4 and (2.15) yield

$$\frac{1}{\varepsilon\kappa}\frac{d}{dt} F_\varepsilon(E(t)) = \operatorname{Re}\langle G, \dot{E}\rangle = -\|\dot{E}\|_F^2 = -\|G - \operatorname{Re}\langle G, E\rangle E\|_F^2 \le 0, \tag{2.17}$$

which gives the precise rate of decay of F_ε along a trajectory $E(t)$ of (2.15). ∎

The following important result states the genericity of the non-vanishing property of the structured gradient.

LEMMA 2.8 (Non-vanishing structured gradient). *Let $A, E \in \mathcal{S}$ and $\varepsilon > 0$, and let λ be a simple target eigenvalue of $A + \varepsilon E$.*

(i) Complex case: $\mathcal{S}$ is a complex-linear subspace of $\mathbb{C}^{n,n}$. Then,

$$G_\varepsilon^{\mathcal{S}}(E) \neq 0 \quad \text{if} \quad \overline{\lambda} f_{\overline{\lambda}} \neq 0.$$

(ii) Real case: $\mathcal{S}$ is a real-linear subspace of $\mathbb{R}^{n,n}$. Then,

$$G_\varepsilon^{\mathcal{S}}(E) \neq 0 \quad \text{if} \quad \operatorname{Re}(\overline{\lambda} f_{\overline{\lambda}}) \neq 0.$$

We emphasize that also A needs to be in $\mathcal{S}$. The result does not hold true when $A \notin \mathcal{S}$.

Proof. We give the proof for the real case. The complex case is analogous but slightly simpler. We take the real inner product of $G_\varepsilon^{\mathcal{S}}(E)$ with $A + \varepsilon E \in \mathcal{S}$ and use the definition (2.11) of $G_\varepsilon^{\mathcal{S}}(E)$:

$$\begin{aligned}
&\langle G_\varepsilon^{\mathcal{S}}(E), A + \varepsilon E\rangle = \operatorname{Re}\langle \Pi^{\mathcal{S}}(2 f_{\overline{\lambda}} x y^*), A + \varepsilon E\rangle = \operatorname{Re}\langle 2 f_{\overline{\lambda}} x y^*, A + \varepsilon E\rangle \\
&= \operatorname{Re}\big(2 f_\lambda\, x^*(A + \varepsilon E) y\big) = \operatorname{Re}\big(2 f_\lambda \lambda\, x^* y\big) = 2\operatorname{Re}\big(f_{\overline{\lambda}}\overline{\lambda}\big)\,(x^* y),
\end{aligned}$$

where $x^*y > 0$. This yields the result. ∎

If the identity matrix I is in $\mathcal{S}$, then the condition for $G_\varepsilon^{\mathcal{S}}(E) \neq 0$ can be weakened:

– In the complex case, it then suffices to have $f_{\overline{\lambda}} \neq 0$. This is seen by taking the inner product with $A + \varepsilon E - \mu I \in \mathcal{S}$ for an arbitrary $\mu \in \mathbb{C}$.

– In the real case, if λ is real, then it suffices to have $\operatorname{Re} f_{\overline{\lambda}} \neq 0$. If λ is non-real, then it even suffices to have $f_{\overline{\lambda}} \neq 0$. In both cases this is seen by taking the inner product with $A + \varepsilon E - \mu I \in \mathcal{S}$ for an arbitrary $\mu \in \mathbb{R}$.

Stationary points. We have the following characterization of stationary points of the norm- and structure-constrained gradient system (2.15) on $\mathcal{S}_1$.

THEOREM 2.9. *Let $E_\star \in \mathcal{S}$ with $\|E_\star\|_F = 1$ be such that*

(i) The target eigenvalue $\lambda(A+\varepsilon E)$ is simple at $E = E_\star$ and depends continuously on E in a neighborhood of $E_\star$.

(ii) The structure-projected gradient $G_\varepsilon^{\mathcal{S}}(E_\star)$ is nonzero.

Let $E(t) \in \mathcal{S}$ be the solution of (2.15) passing through $E_\star$. Then the following are equivalent:

$$\begin{aligned} &1. \quad \frac{d}{dt} F_\varepsilon\left(E(t)\right) = 0 \\ &2. \quad E_\star \text{ is a stationary point of the differential equation (2.15).} \\ &3. \quad E_\star \text{ is a real multiple of } G_\varepsilon^{\mathcal{S}}(E_\star). \end{aligned} \tag{2.18}$$

Proof. Clearly, 3. implies 2., which implies 1. Since the Cauchy–Schwarz inequality in (2.17) is strict unless E is a real multiple of $G_\varepsilon^{\mathcal{S}}(E)$, 1. implies 3. ∎

REMARK 2.10. *In degenerate situations where $G_\varepsilon^{\mathcal{S}}(E_\star) = 0$, we cannot conclude from 2. to 3., i.e., that the stationary point is a projection onto $\mathcal{S}$ of a rank-1 matrix. Such degenerate situations might arise. For example, for the case $f(\lambda, \overline{\lambda}) = -\operatorname{Re}\lambda$ we have seen in Example 2.5 that $G_\varepsilon(E) = -xy^* \neq 0$. However, for a structure space $\mathcal{S}$ we might nevertheless have $G_\varepsilon^{\mathcal{S}}(E_\star) = \Pi^{\mathcal{S}} G_\varepsilon(E_\star) = 0$. An instructive minuscule example is when $n = 2$, $x = y = (1,0)^\top$, and $\mathcal{S}$ is given by a sparsity pattern that does not contain the $(1,1)$ entry. We consider such degenerate cases as exceptional situations and do not consider them further in this paper.*

We call an optimizer $E_\star$ of (2.3) *non-degenerate* if conditions (i) and (ii) of Theorem 2.9 are satisfied. Since non-degenerate optimizers are stationary points of the norm- and structure-constrained gradient system (2.15), Theorem 2.9 immediately yields the following corollary.

THEOREM 2.11. *Non-degenerate optimizers $E_\star$ of (2.3) are projections onto $\mathcal{S}$ of rank-1 matrices.*

This provides the motivation to search for a differential equation that retains the rank-1 property along its solutions. We describe a first, seemingly obvious approach in the next subsection and then turn to a less obvious alternative in Section 3 on which we focus in Sections 3 to 7.

2.6. Constrained gradient flow for structure-projected rank-1 matrices. Let $\mathcal{M}_1 = \mathcal{M}_1(\mathbb{C}^{n\times n})$ be the manifold of complex $n \times n$ rank-1 matrices, and let $\mathcal{M}_1^{\mathcal{S}} = \Pi^{\mathcal{S}}\mathcal{M}_1$ be the set of $\mathcal{S}$-projected rank-1 matrices. We note that $\mathcal{M}_1^{\mathcal{S}}$ need not be a manifold. For example, for $\mathcal{S} = \mathbb{R}^{n,n} = \operatorname{Re}\mathbb{C}^{n,n}$ we have $\mathcal{M}_1^{\mathcal{S}} = \operatorname{Re}\mathcal{M}_1$, which is the union of $\{0\}$ and the two manifolds of real rank-1 and rank-2 matrices. Let us suppose in this short subsection that $\mathcal{M}_1^{\mathcal{S}}$ *is* a manifold, at least locally in a neighborhood of interest. For $E \in \mathcal{M}_1^{\mathcal{S}}$ (in such a neighborhood), we then let $T_E\mathcal{M}_1^{\mathcal{S}}$ be the tangent space at E of $\mathcal{M}_1^{\mathcal{S}}$.

We further suppose that the orthogonal projection $P_E^{\mathcal{S}}$ onto the tangent space $T_E\mathcal{M}_1^{\mathcal{S}}$ is computationally readily available. This is the case when the structure space $\mathcal{S}$ consists of matrices of prescribed range and co-range, as will be discussed in Section 8. However, this is not the case when the structure is given by a sparsity pattern, for which we therefore propose the alternative approach of Section 3.

We consider the projected gradient system on the manifold $\mathcal{M}_1^{\mathcal{S}}$:

$$\dot{E} = -P_E^{\mathcal{S}} G_\varepsilon^{\mathcal{S}}(E) + \operatorname{Re}\langle P_E^{\mathcal{S}} G_\varepsilon^{\mathcal{S}}(E), E\rangle E. \tag{2.19}$$

We note that $P_E^{\mathcal{S}} E = E$ for $E \in \mathcal{M}_1^{\mathcal{S}}$, because the fact that scalar multiples of E are again in $\mathcal{M}_1^{\mathcal{S}}$ implies that $E \in T_E\mathcal{M}_1^{\mathcal{S}}$. Therefore, the right-hand side of (2.19) is in the tangent space $T_E\mathcal{M}_1^{\mathcal{S}}$, and so we have a differential equation on $\mathcal{M}_1^{\mathcal{S}}$.

Since $\operatorname{Re}\langle E, \dot{E}\rangle = 0$, the unit Frobenius norm is preserved. Moreover, we again have the monotonicity property (2.16) by the same argument as before.

Under the non-degeneracy condition $P_E^{\mathcal{S}} G_\varepsilon^{\mathcal{S}}(E_\star) \neq 0$, we have that $E_\star \in \mathcal{M}_1^{\mathcal{S}}$ is a stationary point of (2.19) if and only if $E_\star$ is a real multiple of $P_{E_\star}^{\mathcal{S}} G_\varepsilon^{\mathcal{S}}(E_\star)$. Clearly, every stationary point $E_\star$ of (2.15) is also a stationary point of (2.19). In fact if the right-hand side $R(E)$ of (2.15) vanishes, then also $P_E^{\mathcal{S}} R(E)$ vanishes, which is the right-hand side of (2.19). However, in general we cannot exclude that (2.19) may have additional, spurious stationary points $E_\odot$ that are not a real multiple of $G_\varepsilon^{\mathcal{S}}(E_\odot)$.

In Section 8 we show how the projected gradient system (2.19) can actually be used in computations when the structure space $\mathcal{S}$ consists of complex matrices with prescribed range and co-range. Moreover, we find that in this particular case no spurious stationary points are possible.

3. A rank-1 matrix differential equation. When the structure space $\mathcal{S}$ consists of matrices with a prescribed sparsity pattern, the above tangent space projection $P_E^{\mathcal{S}}$ is not readily available, and the projected gradient system (2.19) of the previous subsection can apparently not be used in a computationally efficient way. As a more accessible alternative, we consider a differential equation on the manifold $\mathcal{M}_1$ of rank-1 matrices, which uses only the known and computationally very simple orthogonal projections $\Pi^{\mathcal{S}}$ onto the structure and P_Y onto the tangent space $T_Y\mathcal{M}_1$ at $Y \in \mathcal{M}_1$ (note that in general $P_E^{\mathcal{S}} \neq \Pi^{\mathcal{S}} P_Y$ for $E = \Pi^{\mathcal{S}} Y$ and $Y \in \mathcal{M}_1$, since $\Pi^{\mathcal{S}}$ and P_Y do not commute). The alternative differential equation is shown to lead to the same stationary points as the structure- and norm-constrained gradient flow (2.15), without any spurious stationary points (under a non-degeneracy condition). However, this differential equation is not a gradient system, and the monotonicity property (2.16) is therefore not guaranteed (though it is usually observed in numerical experiments). This alternative differential equation, reformulated for the factors of the rank-1 matrices, is to be numerically solved into a stationary point.

3.1. Formulation and properties of the rank-1 differential equation. Solutions of (2.15) can be written as $E(t) = \Pi^{\mathcal{S}} Z(t)$, where $Z(t)$ solves

$$\dot{Z} = -G_\varepsilon(\Pi^{\mathcal{S}} Z) + \operatorname{Re}\langle G_\varepsilon(\Pi^{\mathcal{S}} Z), \Pi^{\mathcal{S}} Z\rangle Z, \tag{3.1}$$

as is immediately seen by projecting both sides onto $\mathcal{S}$ with $\Pi^{\mathcal{S}}$ and comparing with (2.15). We note that if $E(t) = \Pi^{\mathcal{S}} Z(t)$ has unit Frobenius norm, then

$$\operatorname{Re}\langle E, \dot{E}\rangle = -\operatorname{Re}\langle E, G_\varepsilon(E)\rangle + \operatorname{Re}\langle G_\varepsilon(E), E\rangle \operatorname{Re}\langle E, E\rangle = 0.$$

Therefore, the unit Frobenius norm of $E(t) = \Pi^{\mathcal{S}} Z(t)$ is conserved for all t.

Since $G_\varepsilon(E)$ is of rank 1 (unless $G_\varepsilon(E) = 0$, which we exclude), every stationary point $Z_\star$ of the differential equation (3.1) is of rank 1. We therefore project the right-hand side onto the tangent space $T_Y\mathcal{M}_1$ at $Y \in \mathcal{M}_1$ and consider instead the projected differential equation with solutions of rank 1:

$$\dot{Y} = -P_Y G_\varepsilon(\Pi^{\mathcal{S}} Y) + \mathrm{Re}\langle P_Y G_\varepsilon(\Pi^{\mathcal{S}} Y), \Pi^{\mathcal{S}} Y\rangle Y. \tag{3.2}$$

Here, $P_Y : \mathbb{C}^{n,n} \to T_Y\mathcal{M}_1$ is the orthogonal projection onto the tangent space $T_Y\mathcal{M}_1$, which for a rank-1 matrix $Y = \sigma uv^*$ with $\|u\| = \|v\| = 1$ is given as (see [20])

$$P_Y(Z) = Z - (I - uu^*)Z(I - vv^*). \tag{3.3}$$

It is useful to note that $P_Y(Y) = Y$. For $E = \Pi^{\mathcal{S}} Y$ of unit Frobenius norm in (3.2), we find

$$\mathrm{Re}\langle E, \dot{E}\rangle = \mathrm{Re}\langle E, \dot{Y}\rangle = -\mathrm{Re}\langle E, P_Y G_\varepsilon(E)\rangle + \mathrm{Re}\langle P_Y G_\varepsilon(E), E\rangle\, \mathrm{Re}\langle E, Y\rangle = 0,$$

where we used that $\mathrm{Re}\langle E, Y\rangle = \mathrm{Re}\langle \Pi^{\mathcal{S}} E, Y\rangle = \mathrm{Re}\langle E, \Pi^{\mathcal{S}} Y\rangle = \mathrm{Re}\langle E, E\rangle = \|E\|_F^2 = 1$. So we have for $E = \Pi^{\mathcal{S}} Y$

$$\|E(t)\|_F = 1 \qquad \text{for all } t.$$

Stationary points. The following theorem states that under a non-degeneracy condition, the differential equations (2.15) and (3.2) yield the same stationary points.

THEOREM 3.1 (Relating stationary points).

(a) *Let $E \in \mathcal{S}$ of unit Frobenius norm be a stationary point of the gradient system* (2.15) *that satisfies $\Pi^{\mathcal{S}} G_\varepsilon(E) \neq 0$. Then, $E = \Pi^{\mathcal{S}} Y$ for some matrix $Y \in \mathcal{M}_1$ that is a stationary point of the differential equation* (3.2).

(b) *Conversely, let $Y \in \mathcal{M}_1$ be a stationary point of the differential equation* (3.2) *such that $E = \Pi^{\mathcal{S}} Y$ has unit Frobenius norm and $P_Y G_\varepsilon(E) \neq 0$. Then, $P_Y G_\varepsilon(E) = G_\varepsilon(E)$, Y is a nonzero real multiple of $G_\varepsilon(E)$, and E is a stationary point of the gradient system* (2.15).

Proof. Let $G = G_\varepsilon(E)$ in this proof for short.

(a) By (2.18), $E = \mu^{-1}\Pi^{\mathcal{S}} G$ for some nonzero real μ. Then, $Y := \mu^{-1} G$ is of rank 1 and we have $E = \Pi^{\mathcal{S}} Y$. We further note that $P_Y G = \mu P_Y Y = \mu Y = G$. We thus have

$$-P_Y G + \mathrm{Re}\langle P_Y G, E\rangle Y = -G + \mathrm{Re}\langle G, E\rangle Y.$$

Here we find that

$$\mathrm{Re}\langle G, E\rangle = \mathrm{Re}\langle \Pi^{\mathcal{S}} G, E\rangle = \mathrm{Re}\langle \mu E, E\rangle = \mu \|E\|_F^2 = \mu.$$

So we have

$$-G + \mathrm{Re}\langle G, E\rangle Y = -G + \mu Y = 0$$

by the definition of Y. This shows that Y is a stationary point of (3.2).

(b) We show that Y is a nonzero real multiple of G. By Theorem 2.9, E is then a stationary point of the differential equation (2.15).

For a stationary point Y of (3.2), we have that $P_Y(G)$ is a nonzero real multiple of Y. Hence, in view of $P_Y(Y) = Y$, we can write G as

$$G = \mu Y + W, \qquad \text{where } \mu \neq 0 \text{ is real and } P_Y(W) = 0.$$

Writing the rank-1 matrix $Y = \rho uv^*$ with $\rho \neq 0$ and $\|u\| = \|v\| = 1$, we then have by (3.3) that

$$W = W - P_Y(W) = (I - uu^*)W(I - vv^*).$$

On the other hand, $G = 2\overline{f}_\lambda xy^*$ is also of rank 1. So we have

$$2\overline{f}_\lambda xy^* = \mu uv^* + (I - uu^*)W(I - vv^*).$$

Multiplying from the right with v yields that x is a complex multiple of u, and multiplying from the left by u^* yields that y is a complex multiple of v. Hence, G is a complex multiple of Y. Since we already know that $P_Y(G)$ is a nonzero real multiple of $P_Y(Y) = Y$, it follows that G is the same real multiple of Y.

Thus stationary points $Y \in \mathcal{M}_1$ of the differential equation (3.2) are characterized as real multiples of G. Hence, $E = \Pi^{\mathcal{S}} Y$ is a real multiple of $\Pi^{\mathcal{S}} G$, and by Theorem 2.9, $E = \Pi^{\mathcal{S}} Y$ is a stationary point of (2.15). □

Possible loss of monotonicity. Since the projections $\Pi^{\mathcal{S}}$ and P_Y do not commute, along solutions of (3.2) we cannot guarantee the monotonicity property (2.16) that we have for the constrained gradient system (2.15).

However, in all our numerical experiments we observed that starting with an initial datum given by the negative free gradient of the considered functional (2.9), i.e. $Y(0) = -G_\varepsilon(0)$, we always obtained a monotone convergence behavior to a (local) optimum. Only in very few cases, by starting from a randomly chosen initial datum, we were able to observe a nonmonotonic convergence. However the loss of monotonicity occurred only once, after the first step, and monotonicity was recovered from the following step onwards. In the following section we will explain this behavior locally near a stationary point, but we have no theoretical explanation for the favorable numerically observed monotonic behavior far from stationary points.

3.2. Differential equations for the factors of rank-1 matrices. Equation (3.2) is an abstract differential equation on the rank-1 manifold $\mathcal{M}_1$. We write a rank-1 matrix $Y \in \mathcal{M}_1$ in a non-unique way as

$$Y = \rho uv^*,$$

where $\rho \in \mathbb{R}$, $\rho > 0$ and $u, v \in \mathbb{C}^n$ have unit norm.

The following lemma shows how we can rewrite the rank-1 differential equation (3.2) in terms of differential equations for the factors u, v and an explicit formula for ρ.

LEMMA 3.2 (Differential equations for the factors). *Every solution $Y(t) \in \mathcal{M}_1$ of the rank-1 differential equation (3.2) with $\|\Pi^{\mathcal{S}} Y(t)\|_F = 1$ can be written as $Y(t) = \rho(t)u(t)v(t)^*$ from the following differential equations for the factors u and v of unit norm,*

$$\begin{aligned} \rho \dot{u} &= -(I - uu^*)Gv - \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^*Gv)u, \\ \rho \dot{v} &= -(I - vv^*)G^*u + \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^*Gv)v, \end{aligned}$$

where $G = G_\varepsilon(E)$ for $E = \Pi^{\mathcal{S}} Y = \rho\, \Pi^{\mathcal{S}}(uv^*)$ and $\rho = 1/\|\Pi^{\mathcal{S}}(uv^*)\|_F$.

The positive factor ρ on the left-hand sides of the differential equations for u and v only determines the speed with which the trajectory is traversed, but has no influence on the trajectory itself.

Proof. The equation for ρ is obvious because $1 = \|E\|_F = \rho\|\Pi^{\mathcal{S}}(uv^*)\|_F$. We write the right-hand side of (3.2) and use (3.3) to obtain for $Y = \rho uv^*$

$$\begin{aligned}
\dot{Y} &= -P_Y G + \mathrm{Re}\langle P_Y G, E\rangle Y \\
&= -(I-uu^*)Gvv^* - uu^*G(I-vv^*) - uu^*Gvv^* + \mathrm{Re}\Big\langle P_Y G, E\Big\rangle Y \\
&= -\Big((I-uu^*)Gvv^* + \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^*Gv)u\Big)v^* - u\Big(u^*G(I-vv^*) + \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^*Gv)v^*\Big) \\
&\quad - \Big(\mathrm{Re}(u^*Gv) + \mathrm{Re}\langle P_Y G, E\rangle\rho\Big)uv^*.
\end{aligned}$$

Since this is equal to $\dot{Y} = (\rho\dot{u})v^* + u(\rho\dot{v}^*) + \dot{\rho}uv^*$, we can read off $\rho\dot{u}$, $\rho\dot{v}^*$ and $\dot{\rho}$ as the three terms in big brackets. This yields the stated differential equations for u and v (and another one for ρ, which will not be needed). Note that $(d/dt)\|u\|^2 = 2\,\mathrm{Re}(u^*\dot{u}) = 0$ and analogously for v, so that the unit norm of u and v is conserved. ∎

We note that for $G = G_\varepsilon(E) = 2f_{\overline{\lambda}}xy^*$ (see Lemma 2.4) and with $\alpha = u^*x$, $\beta = v^*y$ and $\gamma = 2f_{\overline{\lambda}}$, we obtain the differential equations

$$\begin{aligned}
\rho\dot{u} &= \alpha\overline{\beta}\gamma\, u - \overline{\beta}\gamma\, x - \tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\alpha\overline{\beta}\gamma)u \\
\rho\dot{v} &= \overline{\alpha}\beta\overline{\gamma}\, v - \overline{\alpha\gamma}\, y - \tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\overline{\alpha}\beta\overline{\gamma})v.
\end{aligned} \tag{3.4}$$

3.3. Cases of interest. The real dimension of the manifold of complex $n \times n$ rank-1 matrices of unit norm is $4n-2$. Integrating (3.2) instead of (2.15) is very appealing in those cases where $\dim \mathcal{S}$ is significantly larger than $4n-2$. An important example is given by sparse matrices with a sparsity pattern with a number of nonzero elements of order cn with $c > 4$ (and ideally much larger than 4).

In the case of a real target eigenvalue the dimension of the manifold of real $n \times n$ rank-1 matrices of unit norm is $2n-1$ so that for structured matrices it is meaningful to make use of (3.2) if $c > 2$.

Similarly, when considering matrices with prescribed range and co-range,

$$\mathcal{S} = \{B\Delta C \,:\, \Delta \in \mathbb{R}^{k,l}\},$$

where $B \in \mathbb{R}^{n,k}$ and $C \in \mathbb{R}^{l,n}$ with $k, l < n$, replacing the unknown matrix Δ, which is a full $k \times l$ real matrix, by a rank-1 matrix, significantly reduces the memory requirements when k and l are large. As for the computational cost, we may argue that the reduced number of variables may lead to faster convergence of the method.

4. Local convergence to local minima of solutions to the rank-1 projected differential equation. In this section we show that solutions of the rank-1 projected differential equation (3.2) converge locally to strong local minima of the functional F_ε, that is the solution converges to a local minimum E for which the Hessian matrix $H_\varepsilon(E)$ of F_ε at E yields a positive definite quadratic form when restricted to the tangent space $T_E\mathcal{S}_1$ of the manifold $\mathcal{S}_1$ at E. Here, $\mathcal{S}_1$ is the manifold of matrices in $\mathcal{S}$ of unit Frobenius norm.

We first state the result, then formulate and prove a key lemma, and finally give the proof of the local convergence result.

4.1. Statement of the local convergence result. For the formulation of our local convergence result we need the following assumptions. There, $\mathcal{M}_1$ is the manifold of rank-1 matrices in $\mathbb{C}^{n,n}$, and $\mathcal{M}_1^{\mathcal{S}} = \Pi^{\mathcal{S}}\mathcal{M}_1$ consists of structure-projected rank-1 matrices. The first assumption is made on the structure space $\mathcal{S}$. It excludes, in particular, spaces $\mathcal{S}$ that are too low-dimensional: it requires $\dim \mathcal{S} \geq \dim \mathcal{M}_1 = 4n-2$ (as before dim indicates the *real* dimension).

ASSUMPTION 4.1. *The restricted projection* $\Pi^{\mathcal{S}}\big|_{\mathcal{M}_1} : \mathcal{M}_1 \to \mathcal{M}_1^{\mathcal{S}} \subset \mathcal{S}$ *is a local diffeomorphism, or equivalently:*

(i) If $E = \Pi^{\mathcal{S}}Y \in \mathcal{M}_1^{\mathcal{S}}$ *for some* $Y \in \mathcal{M}_1$*, then* Y *is locally unique.*

(ii) The local inverse map $E \mapsto Y$ *is continuously differentiable.*

REMARK 4.2. We comment on Assumption 4.1 to (a) make it plausible in the case of perturbation matrices E with prescribed sparsity pattern and (b) show that it is not satisfied in the case of perturbation matrices with prescribed range and co-range.

(a) Consider the structure $\mathcal{S}$ of real $n \times n$ matrices with a prescribed sparsity pattern. Let $E \in \mathcal{M}_1^{\mathcal{S}} \subset \mathcal{S}$ be given. So $E = \Pi^{\mathcal{S}}\widehat{Y}$ for some $\widehat{Y} \in \mathcal{M}_1$. In principle, in order to determine all solutions of the equation

$$\Pi^{\mathcal{S}}Y = E \tag{4.1}$$

we should form $Y = uv^*$ with $u, v \in \mathbb{C}^n$ with $\|u\| = 1$ and $v \neq 0$, and write a system of quadratic equations in the variables $\{u_i\}_{i=1}^n$ and $\{v_j\}_{j=1}^n$ that reads

$$\operatorname{Re}\left(u_i v_j^*\right) = E_{ij} \qquad \text{for all } (i,j) \in \mathcal{P}$$

where $\mathcal{P}$ is the considered sparsity pattern, together with the norm constraint $\|u\|^2 = 1$, and moreover the first nonzero entry of u can be chosen to be real and positive to guarantee uniqueness of the representation $Y = uv^*$.

This gives a system of $s+1$ quadratic equations where $s = \#\mathcal{P} = \dim \mathcal{S}$ is the number of entries of E which are not prescribed to be zero. In terms of the real variables $\operatorname{Re}(u_i), \operatorname{Im}(u_i), \operatorname{Re}(v_i), \operatorname{Im}(v_i)$ (excluding $\operatorname{Im}(u_1) = 0$), the system has $s+1$ quadratic equations in $4n-1$ variables: $\Phi(u,v) \equiv \Pi^{\mathcal{S}}Y = E$. We have local uniqueness of $\widehat{Y} = \widehat{u}\widehat{v}^*$ if the derivative matrix $D\Phi(\widehat{u},\widehat{v}) \in \mathbb{C}^{s+1,4n-1}$ has only the trivial kernel 0. This can be expected to hold true generically if $s \geq 4n-2$ (and the more so as s gets larger).

On the other hand, if $s < 4n-2$, then integrating the gradient system (2.15), translated into a system of differential equations in terms of the s nonzero entries of E, would be favorable over integrating the rank-1 matrix differential equation (2.15). This further indicates that Assumption 4.1 is reasonable in the case where the structure is given by a sparsity pattern.

For the structure space $\mathcal{S}$ of matrices with a prescribed sparsity pattern, Assumption 4.1 is reminiscent of the problem of matrix completion, where the aim is to minimize the rank r such that there exists a unique matrix M of rank r with $\Pi^{\mathcal{S}}M = E$ for a given matrix $E \in \mathcal{S}$; see e.g. [5]. Note, however, that in Assumption 4.1 the condition is not about existence but about local uniqueness, and the rank is fixed to 1.

(b) Assumption 4.1 is not satisfied in the case where the structure $\mathcal{S}$ is given by matrices with prescribed range and co-range. Since in this case the orthogonal projection onto the structure is given by $\Pi^{\mathcal{S}}Y = BB^{\dagger}YC^{\dagger}C$ (see Example 2.2), we have that for a rank-1 matrix $Y = uv^*$, the projected matrix $E = \Pi^{\mathcal{S}}Y$ can also be written as $E = \Pi^{\mathcal{S}}\widetilde{Y}$ with $Y = (u+\widetilde{u})(v+\widetilde{v})^*$ for arbitrary $\widetilde{u} \in \operatorname{Ker} B^{\top}$ and

$\widetilde{v} \in \operatorname{Ker} C^\top$, and so condition (i) in Assumption 4.1 is violated. This could be remedied by requiring that $Y = uv^*$ be such that $u \in (\operatorname{Ker} B^\top)^\perp = \operatorname{Ran} B$ and $v \in (\operatorname{Ker} C^\top)^\perp = \operatorname{Ran} C$ and incorporating these constraints in the differential equation. We will not carry this out in detail for two reasons: On the one hand it did not seem necessary in our numerical experiments, and on the other hand we can here work instead with the projected gradient system (2.19) on $\mathcal{M}_1^{\mathcal{S}}$, as is described in Section 8.

The next assumption is made on the Hessian of the functional F_ε at a stationary point of the differential equation (2.15).

ASSUMPTION 4.3. *Let $E_0 \in \mathcal{S}_1$ be a stationary point of the constrained gradient system* (2.15)*. We assume that E_0 is a strong minimum of the functional F_ε on $\mathcal{S}_1$, that is, the Hessian matrix $H_\varepsilon(E_0)$ of F_ε at E_0 yields a positive definite quadratic form when restricted to the tangent space $T_{E_0}\mathcal{S}_1$ of the manifold $\mathcal{S}_1$ at E_0: there exists $\alpha > 0$ such that*

$$\langle Z, H_\varepsilon(E_0)Z\rangle \geq \alpha \|Z\|_F^2 \qquad \forall\, Z \in T_{E_0}\mathcal{S}_1. \tag{4.2}$$

Under these assumptions we have the following result.

THEOREM 4.4 (Local convergence to a strong local minimum). *Under Assumption* 4.1*, let the rank-1 matrix $Y_0 \in \mathcal{M}_1$ be a stationary point of the projected differential equation* (3.2) *such that $E_0 = \Pi^{\mathcal{S}} Y_0 \in \mathcal{S}_1$ is of unit Frobenius norm and $P_{Y_0} G_\varepsilon(E_0) \neq 0$. We assume that E_0 satisfies Assumption* 4.3*.*

Then, for an initial datum $Y(0)$ sufficiently close to Y_0, the solution $Y(t)$ of (3.2) *converges to Y_0 exponentially as $t \to \infty$. Moreover, $F_\varepsilon\left(\Pi^{\mathcal{S}} Y(t)\right)$ decreases monotonically with t and converges exponentially to the local minimum value $F_\varepsilon(E_0)$ as $t \to \infty$.*

Note that $E_0 = \Pi^{\mathcal{S}} Y_0 \in \mathcal{S}_1$ is a stationary point of (2.15) by Theorem 3.1 (b). So the assumption on E_0 reduces to the condition (4.2) on the Hessian $H_\varepsilon(E_0)$.

The proof of Theorem 4.4 will be given in Section 4.3.

4.2. A basic lemma. The following remarkable lemma provides the key to the proof of Theorem 4.4.

LEMMA 4.5. *Let $Y_\star \in \mathcal{M}_1$ with $E_\star = \Pi^{\mathcal{S}} Y_\star \in \mathcal{S}$ of unit Frobenius norm. Let $Y_\star$ be a stationary point of the rank-1 projected differential equation* (3.2)*, with an associated target eigenvalue λ of $A + \varepsilon E_\star$ that is simple. Then, there exists $\bar{\delta} > 0$ such that for all positive $\delta \leq \bar{\delta}$ and all $Y \in \mathcal{M}_1$ with $\|Y - Y_\star\|_F \leq \delta$ and $\Pi^{\mathcal{S}} Y$ of unit norm, we have*

$$\|P_Y G_\varepsilon(\Pi^{\mathcal{S}} Y) - G_\varepsilon(\Pi^{\mathcal{S}} Y)\|_F \leq C\delta^2 \tag{4.3}$$

with C independent of δ.

Proof. Let us consider a smooth regular path $Y(\tau) = u(\tau)v(\tau)^* \in \mathcal{M}_1$ (with nonzero $u(\tau), v(\tau) \in \mathbb{C}^n$) such that $E(\tau) = \Pi^{\mathcal{S}} Y(\tau)$ is of unit Frobenius norm and

$$Y(0) = Y_\star = \alpha G_\star \quad \text{for some real } \alpha, \text{ where } G_\star = G_\varepsilon\left(E(0)\right) = 2\overline{f}_\lambda x y^*,$$

where (λ, x, y) is the eigentriplet of $A + \varepsilon E(0)$ associated with the target eigenvalue λ. Similarly, for $\tau \in [0, \delta]$ with δ such that $\lambda(\tau)$ remains simple, we have

$$G(\tau) = G_\varepsilon(E(\tau)) = 2\overline{f}_\lambda(\tau) x(\tau) y(\tau)^*.$$

We may assume that the path is parametrized such that $\|\dot{Y}(\tau)\|_F = 1$ and hence we have $\|Y(\tau) - Y_\star\|_F \sim \tau$ for small τ. By the given assumptions all quantities

are smooth w.r.t. τ. In particular, for a simple eigenvalue, under a smooth matrix perturbation, the derivatives $\dot{x}(\tau)$ and $\dot{y}(\tau)$ of the associated eigenvectors - under the assumed normalization (2.7) - are given by (see e.g. [23, 13])

$$\begin{aligned}\frac{1}{\varepsilon}\,\dot{x}(\tau)^* &= -x(\tau)^*\dot{E}(\tau)N(\tau) + \mathrm{Re}\left(x(\tau)^*\dot{E}(\tau)N(\tau)x(\tau)\right)x(\tau)^*,\\ \frac{1}{\varepsilon}\,\dot{y}(\tau) &= -N(\tau)\dot{E}(\tau)y(\tau) + \mathrm{Re}\left(y(\tau)^*N(\tau)\dot{E}(\tau)y(\tau)\right)y(\tau),\end{aligned}$$

where $N(\tau)$ is the group inverse of $A+\varepsilon E(\tau)-\lambda(\tau)I$ and the last terms on the right-hand side of both differential equations account for the unit norm preservation for both eigenvectors and for the positivity of their inner product. Note that by the simplicity of $\lambda(\tau)$, the group inverse $N(\tau)$ and thus also $\dot{x}(\tau)$ and $\dot{y}(\tau)$ as well as their derivatives are bounded.

With the formula (3.3) for the projection P_Y, we thus have the following first order expansion near $\tau=0$. Here we indicate by u, v, x, y and $f_{\overline{\lambda}}$ (and further $\dot{u}, \dot{v}, \dot{x}, \dot{y}$ and $\dot{f}_{\overline{\lambda}}$) the associated functions of τ at $\tau=0$, i.e. corresponding to the stationary point. We have

$$\begin{aligned}&P_{Y(\tau)}G(\tau) = P_{Y(\tau)}\Big(f_{\overline{\lambda}}(\tau)x(\tau)y(\tau)^*\Big) = \Big(f_{\overline{\lambda}}+\tau\dot{f}_{\overline{\lambda}}\Big)\cdot\\ &\Bigg(\Big(xx^*+\tau\left(\dot{u}x^*+x\dot{u}^*\right)\Big)\Big(xy^*+\tau\left(\dot{x}y^*+x\dot{y}^*\right)\Big)+\\ &\Big(xy^*+\tau\left(\dot{x}y^*+x\dot{y}^*\right)\Big)\Big(yy^*+\tau\left(\dot{v}y^*+y\dot{v}^*\right)\Big)-\\ &\Big(xx^*+\tau\left(\dot{u}x^*+x\dot{u}^*\right)\Big)\Big(xy^*+\tau\left(\dot{x}y^*+x\dot{y}^*\right)\Big)\Big(yy^*+\tau\left(\dot{v}y^*+y\dot{v}^*\right)\Big)\Bigg)+\mathcal{O}(\tau^2) =\\ &f_{\overline{\lambda}}xy^*+\tau\Big(\dot{f}_{\overline{\lambda}}xy^*+f_{\overline{\lambda}}x\dot{y}^*+f_{\overline{\lambda}}\dot{x}y^*\Big)+\mathcal{O}(\tau^2).\end{aligned}\tag{4.4}$$

Consequently, (4.4) has the same first order expansion as

$$G(\tau) = f_{\overline{\lambda}}(\tau)x(\tau)y(\tau)^* = f_{\overline{\lambda}}xy^*+\tau\Big(\dot{f}_{\overline{\lambda}}xy^*+f_{\overline{\lambda}}x\dot{y}^*+f_{\lambda}\dot{x}y^*\Big)+\mathcal{O}(\tau^2),$$

which yields the result. □

4.3. Proof of Theorem 4.4. With $E(t)=\Pi^{\mathcal{S}}Y(t)\in\mathcal{S}_1$, the differential equation (3.2) for $Y(t)\in\mathcal{M}_1$ is equivalent to

$$\dot{E} = -\Pi^{\mathcal{S}}P_Y G_\varepsilon(E) + \mathrm{Re}\langle\Pi^{\mathcal{S}}P_Y G_\varepsilon(E), E\rangle E.$$

By Lemma 4.5, this can be rewritten as a perturbation to the constrained gradient system (2.15) (recall that $G_\varepsilon^{\mathcal{S}}=\Pi^{\mathcal{S}}G_\varepsilon$):

$$\dot{E} = -G_\varepsilon^{\mathcal{S}}(E) + \mathrm{Re}\langle G_\varepsilon^{\mathcal{S}}(E), E\rangle E + D \qquad \text{with} \qquad \|D(t)\| = \mathcal{O}(\|Y(t)-Y_\star\|^2).$$

By Assumption 4.1 (ii), this bound further implies

$$\|D(t)\| = \mathcal{O}(\|E(t)-E_\star\|^2).$$

The orthogonal projection $\widehat{\Pi}_E$ of $Z \in \mathbb{C}^{n,n}$ onto the tangent space $T_E\mathcal{S}_1$ at $E \in \mathcal{S}_1$ is given by

$$\widehat{\Pi}_E Z = \Pi^{\mathcal{S}} Z - \mathrm{Re}\langle \Pi^{\mathcal{S}} Z, E\rangle E.$$

We write

$$\widehat{G}(E) = \widehat{\Pi}_E G_\varepsilon(E) = G_\varepsilon^{\mathcal{S}}(E) - \mathrm{Re}\langle G_\varepsilon^{\mathcal{S}}(E), E\rangle E$$

for short. We have

$$\frac{1}{2}\frac{d}{dt}\,\|E(t) - E_\star\|^2 = \mathrm{Re}\langle E - E_\star, \dot{E}\rangle = \mathrm{Re}\langle E - E_\star, -\widehat{G}(E) + D\rangle.$$

Since $\widehat{G}(E_\star) = 0$ and

$$E - E_\star = \widehat{\Pi}_{E_\star}(E - E_\star) + \mathcal{O}(\|E - E_\star\|^2),$$

which is due to the fact that both E and $E_\star$ lie on $\mathcal{S}_1$ so that letting $\delta := \|E - E_\star\|$, the orthogonal projection $\widehat{\Pi}_{E_\star}(E - E_\star)$ onto the tangent plane at E_* is δ^2-close to $E - E_*$, we find

$$\widehat{G}(E) = \widehat{G}(E) - \widehat{G}(E_\star) = \widehat{\Pi}_{E_\star} H_\varepsilon(E_\star)\widehat{\Pi}_{E_\star}(E - E_\star) + \mathcal{O}(\|E - E_\star\|^2),$$

where $H_\varepsilon(E_\star)$ is the Hessian matrix of the functional F_ε at $E_\star$. By Assumption 4.3, $H_\varepsilon(E_\star)$ is positive definite on $T_{E_\star}\mathcal{S}_1$. So we obtain

$$\begin{aligned}
&\mathrm{Re}\langle E - E_\star, -\widehat{G}(E) + D\rangle \\
&= \mathrm{Re}\langle \widehat{\Pi}_{E_\star}(E - E_\star) + \mathcal{O}(\|E - E_\star\|^2), -\widehat{\Pi}_{E_\star} H(E_\star)\widehat{\Pi}_{E_\star}(E - E_\star) + \mathcal{O}(\|E - E_\star\|^2)\rangle \\
&= -\langle \widehat{\Pi}_{E_\star}(E - E_\star), -H(E_\star)\widehat{\Pi}_{E_\star}(E - E_\star)\rangle + \mathcal{O}(\|E - E_\star\|^3) \\
&\le -\alpha\|\widehat{\Pi}_{E_\star}(E - E_\star)\|^2 + \mathcal{O}(\|E - E_\star\|^3) \\
&\le -\tfrac{1}{2}\alpha\|E - E_\star\|^2,
\end{aligned}$$

provided that E is sufficiently close to $E_\star$. This yields that $\|E(t) - E_\star\|$ decreases monotonically with growing t and converges exponentially fast to 0 as $t \to \infty$. Similarly we obtain, with the projected Hessian $\widehat{H}(E_\star) = \widehat{\Pi}_{E_\star} H_\varepsilon(E_\star)\widehat{\Pi}_{E_\star}$ for short,

$$\begin{aligned}
&\frac{1}{\kappa\varepsilon}\frac{d}{dt}F_\varepsilon(E(t)) = \mathrm{Re}\langle G_\varepsilon(E), \dot{E}\rangle = \mathrm{Re}\langle \widehat{G}(E), \dot{E}\rangle = \mathrm{Re}\langle \widehat{G}(E), -\widehat{G}(E) + D\rangle \\
&= -\|\widehat{H}(E_\star)\widehat{\Pi}_{E_\star}(E - E_\star)\|^2 + \mathcal{O}(\|E - E_\star\|^3) \\
&\le -\alpha^2\|\widehat{\Pi}_{E_\star}(E - E_\star)\|^2 + \mathcal{O}(\|E - E_\star\|^3) \\
&\le -\tfrac{1}{2}\alpha^2\|E - E_\star\|^2,
\end{aligned}$$

provided that E is sufficiently close to $E_\star$. We conclude that $F_\varepsilon(E(t))$ decreases monotonically and exponentially to $F_\varepsilon(E_\star)$. □

5. Numerical integration by a splitting method. We need to integrate numerically the differential equations (3.4). The objective here is not to follow a particular trajectory accurately, but to arrive quickly at a stationary point. The simplest method is the normalized Euler method, or normalized gradient descent method, where the result after an Euler step (i.e., a steepest descent step) is normalized to unit norm for both the u- and v-component. This can be combined with a standard line search strategy to determine the step size adaptively. We found, however, that a more efficient method is obtained with a splitting method instead of the Euler method.

5.1. Splitting. The splitting method consists of a first step applied to the differential equations (with $\alpha = u^*x$, $\beta = v^*y$ and $\gamma = 2f_{\overline{\lambda}}$ as in (3.4))

$$\begin{aligned} \rho\dot{u} &= \alpha\overline{\beta}\gamma\, u - \overline{\beta}\gamma\, x \\ \rho\dot{v} &= \overline{\alpha}\beta\overline{\gamma}\, v - \overline{\alpha\gamma}\, y \end{aligned} \tag{5.1}$$

followed by a step for the differential equations

$$\begin{aligned} \rho\dot{u} &= -\tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\alpha\overline{\beta}\gamma)u \\ \rho\dot{v} &= +\tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\alpha\overline{\beta}\gamma)v. \end{aligned} \tag{5.2}$$

Note that the second differential equation is a mere rotation of u and v. In the case of a real eigenvalue of a real matrix, the system (5.2) has a vanishing right-hand side and can therefore be ignored. As is unusual for splitting methods, this method preserves stationary points.

LEMMA 5.1 (Stationary points). *(u, v) is a stationary point of the differential equations* (3.4) *if and only if (u, v) is a stationary point of the differential equations* (5.1) *and* (5.2).

Proof. If (u, v) is a stationary point of (3.4), then u is proportional to x and v is proportional to y. Hence, $x = \alpha u$ and $y = \beta v$. This implies that (u, v) is a stationary point of (5.1), and hence also of (5.2). The converse direction is evident. ∎

5.2. Fully discrete splitting algorithm. Starting from vectors u_k, v_k of unit norm and

$$\rho_k = \frac{1}{\|\Pi^{\mathcal{S}}(u_k v_k^*)\|_F}, \tag{5.3}$$

we denote by x_k and y_k the left and right eigenvectors to the target eigenvalue λ_k of $A + \varepsilon\rho_k\Pi^{\mathcal{S}}(u_k v_k^*)$, and set

$$\alpha_k = u_k^* x_k, \qquad \beta_k = v_k^* y_k, \qquad \gamma_k = 2f_{\overline{\lambda}_k}. \tag{5.4}$$

We apply the Euler method with step size h to (5.1) to obtain

$$\begin{aligned} \widehat{u}(h) &= u_k + (h/\rho_k)\left(\alpha_k\overline{\beta}_k\gamma_k\, u_k - \overline{\beta}_k\gamma_k\, x_k\right) \\ \widehat{v}(h) &= v_k + (h/\rho_k)\left(\overline{\alpha}_k\beta_k\overline{\gamma}_k\, v_k - \overline{\alpha_k\gamma_k}\, y_k\right), \end{aligned} \tag{5.5}$$

followed by a normalization to unit norm

$$\widetilde{u}(h) = \frac{\widehat{u}(h)}{\|\widehat{u}(h)\|}, \quad \widetilde{v}(h) = \frac{\widehat{v}(h)}{\|\widehat{v}(h)\|}. \tag{5.6}$$

Then, as a second step, we integrate the rotating differential equations (5.2) by setting, with $\vartheta = -\frac{1}{2\rho_k}\,\mathrm{Im}(\alpha_k\overline{\beta_k}\gamma_k)$,

$$u(h) = \mathrm{e}^{\mathrm{i}\vartheta h}\,\widetilde{u}(h), \qquad v(h) = \mathrm{e}^{-\mathrm{i}\vartheta h}\,\widetilde{v}(h), \tag{5.7}$$

set $\rho(h) = 1/\|\Pi^{\mathcal{S}}(u(h)v(h)^*)\|_F$, and compute the target eigenvalue $\lambda(h)$ of the perturbed matrix $A + \varepsilon\rho(h)\Pi^{\mathcal{S}}\big(u(h)v(h)^*\big)$.

We note that this fully discrete algorithm still preserves stationary points.

Algorithm 1: Integration step for the rank-1 differential equation

Data: $A, \varepsilon, \theta > 1, u_k \approx u(t_k), v_k \approx v(t_k)$, h_k (proposed step size), λ_k (target eigenvalue of $A + \Delta_k$ with $\Delta_k = \varepsilon \Pi^{\mathcal{S}}(u_k v_k^*)/\|\Pi^{\mathcal{S}}(u_k v_k^*)\|_F$)
Result: u_{k+1}, v_{k+1}, h_{k+1}, λ_{k+1}
begin
Initialize the step size by the proposed step size, $h = h_k$
Compute left/right eigenvectors x_k, y_k of $A + \Delta_k$ to λ_k such that $\|x_k\| = \|y_k\| = 1, x_k^* y_k > 0$
Compute $\alpha_k, \beta_k, \gamma_k$ by (5.4) and g_k by (5.8)
Initialize $f(h) = f_k$
while $f(h) \geq \max(f_k, f_k - h\theta g_k)$ **do**
  Compute $u(h), v(h)$ according to (5.5)-(5.7)
  Compute $\Delta(h) = \varepsilon \rho(h) \Pi^{\mathcal{S}}(u(h)v(h)^*)$ with $\rho(h) = 1/\|\Pi^{\mathcal{S}}(u(h)v(h)^*)\|_F$
  Compute $\lambda(h)$ target eigenvalue of $A + \Delta(h)$
  Compute the value $f(h) = f\big(\lambda(h), \overline{\lambda(h)}\big)$
  if $f(h) \geq \max(f_k, f_k - h\theta g_k)$ **then**
    Reduce the step size, $h := h/\theta$
if $\big(g_k \geq 0$ and $f(h) \geq f_k - (h/\theta) g_k\big)$ or $\big(g_k < 0$ and $f(h) \geq f_k - h\theta g_k\big)$ **then**
  Reduce the step size for the next step, $h_{\text{next}} := h/\theta$
else if $h = h_k$ **then**
  Set $h_{\text{next}} := \theta h_k$ (augment the stepsize if no rejection has occurred)
else
  Set $h_{\text{next}} := h_k$
Set $h_{k+1} := h_{\text{next}}$, $\lambda_{k+1} := \lambda(h)$, and the starting values for the next step as $u_{k+1} := u(h)$, $v_{k+1} := v(h)$
return

One motivation for choosing this method is that near a *non real* stationary point, the motion is almost rotational since $x \approx \alpha u$ and $y \approx \beta v$. The dominant term determining the motion is then the rotational term on the right-hand side of (3.4), which is integrated by a rotation in the above scheme (the integration would be exact if α, β, γ were constant). This algorithm requires in each step one computation of target eigenvalues and associated eigenvectors of structure-projected rank-1 perturbations to the matrix A, which can be computed at moderate computational cost for large sparse matrices A by a Krylov Schur algorithm [26], as implemented in the MATLAB function *eigs*.

5.3. Step-size selection. We use an Armijo-type line search strategy, adapted to the possibility that the functional $f(\lambda, \overline{\lambda})$ is not everywhere reduced along the flow of the differential equation (3.2) (even though this was never observed in our numerical experiments when we chose the initial value $Y(0)$ as a positive multiple of the negative free gradient $-2 f_{\overline{\lambda}} x y^*$ where (λ, x, y) is the target eigentriplet of the matrix A). By Lemma 2.4, the change of the functional along solutions of (3.2) equals

(with $G = G_\varepsilon(E)$) and omitting the argument t on the right-hand side)

$$\frac{d}{dt}F_\varepsilon(E(t)) = \varepsilon\kappa\,\mathrm{Re}\langle G_\varepsilon(E), \dot{E}\rangle$$
$$= -\varepsilon\kappa\Big(\mathrm{Re}\langle \Pi^{\mathcal{S}} G, \Pi^{\mathcal{S}} P_Y G\rangle - \mathrm{Re}\langle \Pi^{\mathcal{S}} P_Y G, E\rangle\,\mathrm{Re}\langle \Pi^{\mathcal{S}} G, E\rangle\Big) =: -g \quad (5.8)$$

We write $g_k = g$ for the choice $E = E_k = u_k v_k^*$, $G = G_\varepsilon(E_k) = 2f_{\overline{\lambda}}(\lambda_k, \overline{\lambda_k})x_k y_k^*$, and $\kappa = \kappa_k = 1/(x_k^* y_k)$. Let

$$f_k = f(\lambda_k, \overline{\lambda_k}), \qquad f(h) = f(\lambda(h), \overline{\lambda(h)}).$$

We accept the result of the step with step size h if, for a given parameter $\theta > 1$,

$$f(h) < \max(f_k, f_k - h\theta g_k).$$

If $g_k \geq 0$ and $f(h) \geq f_k - (h/\theta)g_k$, or if $g_k < 0$ and $f(h) \geq f_k - h\theta g_k$, then we reduce the step size for the next step to h/θ. If the step size has not been reduced in the previous step, we try for a larger step size. Algorithm 1 describes the step from t_k to $t_{k+1} = t_k + h_k$.

6. Application to structured matrix nearness problems. We consider matrix nearness problems that are closely related to the eigenvalue optimization problems considered in this article. We pose the problem in the structure space $\mathcal{S}$. Let again $A \in \mathbb{C}^{n,n}$ be a given matrix and let $\lambda(A) \in \mathbb{C}$ be a target eigenvalue of A. We again consider the smooth function $f(\lambda, \overline{\lambda})$ satisfying (2.1) that is to be minimized. For a prescribed real number r in the range of f we assume that

$$f(\lambda(A), \overline{\lambda}(A)) > r,$$

so that for sufficiently small $\varepsilon > 0$ we have $\phi(\varepsilon) > r$, where

$$\phi(\varepsilon) := \min_{\Delta\in\mathcal{S},\, \|\Delta\|_F=\varepsilon} f\left(\lambda\left(A+\Delta\right), \overline{\lambda}\left(A+\Delta\right)\right).$$

The objective now is to find the smallest $\varepsilon > 0$ such that $\phi(\varepsilon) = r$:

$$\varepsilon_\star = \min\big\{\varepsilon > 0 \,:\, \phi(\varepsilon) \leq r\big\}. \quad (6.1)$$

Determining $\varepsilon_\star$ is a one-dimensional root-finding problem for the function ϕ that is defined by the considered eigenvalue optimization problem.

6.1. Structured distances to singularity and to instability. Let us consider two examples, with a peculiar difference. In the first case the problem reduces to the search of the simple (unique) zero of a smooth function, while in the second case the function is not smooth at its smallest zero, and (generically) vanishes identically right to it.

Example 6.1 (Structured distance to instability). *Let A be a Hurwitz matrix, i.e. with negative spectral abscissa $\alpha(A) = \max\{\mathrm{Re}\,\lambda \,:\, \lambda \text{ is an eigenvalue of } A\} < 0$. With the function $f(\lambda, \overline{\lambda}) = -\frac{1}{2}(\lambda + \overline{\lambda}) = -\mathrm{Re}\,\lambda$ and the target eigenvalue λ given by the eigenvalue of largest real part, and $r = 0$, we arrive at the problem of computing the structured distance to instability of A :*

$$\varepsilon_\star = \min\{\varepsilon > 0 \,:\, \alpha_\varepsilon^{\mathcal{S}}(A) = 0\}, \quad \text{where} \quad \alpha_\varepsilon^{\mathcal{S}}(A) = \max_{E\in\mathcal{S},\|E\|_F=1} \alpha(A + \varepsilon E)$$

is the ε-pseudospectral abscissa with respect to the structure space $\mathcal{S}$.

EXAMPLE 6.2 (Structured distance to singularity). *Let A be a nonsingular matrix. With $f(\lambda, \overline{\lambda}) = \lambda\overline{\lambda} = |\lambda|^2$ and the target eigenvalue λ the one of smallest modulus, we arrive at the problem of computing the distance to singularity of A :*

$$\varepsilon_\star = \min\{\varepsilon > 0 \,:\, \varrho_\varepsilon^{\mathcal{S}}(A) = 0\}, \quad \text{where} \quad \varrho_\varepsilon^{\mathcal{S}}(A) = \min_{E\in\mathcal{S}, \|E\|_F=1} \varrho(A + \varepsilon E)$$

where $\varrho(M)$ is the smallest modulus of eigenvalues of a matrix M.[1]

6.2. Two-level iterative method. As in previous work (see e.g. [9, 8]), we use a two-level method:

(i) **Inner iteration:** Given $\varepsilon > 0$, we aim to compute a matrix $E(\varepsilon) \in \mathcal{S}$ of unit Frobenius norm, such that $F_\varepsilon(E) = f\left(\lambda\left(A + \varepsilon E\right), \overline{\lambda}\left(A + \varepsilon E\right)\right)$ is minimized:

$$E(\varepsilon) = \arg \min_{E\in\mathcal{S}, \|E\|_F=1} F_\varepsilon(E). \tag{6.2}$$

(ii) **Outer iteration:** We compute the smallest positive value $\varepsilon_\star$ with

$$\phi(\varepsilon_\star) = r, \tag{6.3}$$

where $\phi(\varepsilon) = F_\varepsilon\left(E(\varepsilon)\right) = f\left(\lambda\left(A + \varepsilon E(\varepsilon)\right), \overline{\lambda}\left(A + \varepsilon E(\varepsilon)\right)\right)$.

6.3. Inner iteration. The eigenvalue optimization problem (6.2) is precisely of the type studied in the previous sections. To compute $E(\varepsilon)$ for a given $\varepsilon > 0$, we integrate numerically either the ODE system (2.15) or (3.2); see Section 5.

The computational cost can be significantly reduced if we are able to compute efficiently $\Pi^{\mathcal{S}}(Y)$ and the matrix vector multiplication $\Pi^{\mathcal{S}}(Y)v$ (with $v \in \mathbb{C}^n$) which is typically used by an iterative eigensolver applied to $A + \varepsilon\Pi^{\mathcal{S}}(Y)$. This is true for example when $\mathcal{S}$ is the set of matrices with a prescribed sparsity pattern. Note that often also linear system solves are required to find the desired eigenvalue and a convenient solution of the structured linear systems is desirable.

6.4. Outer iteration. The outer iteration determines the smallest positive solution of the one-dimensional root-finding problem (6.3). We make use of a locally quadratically convergent Newton-type method, which can be justified under appropriate regularity assumptions. It turns out that the derivative of ϕ is then simply

$$\phi'(\varepsilon) = -\|\Pi^{\mathcal{S}} G_\varepsilon(E(\varepsilon))\|_F / (x(\varepsilon)^* y(\varepsilon)), \tag{6.4}$$

where $x(\varepsilon)$ and $y(\varepsilon)$ with $x(\varepsilon)^* y(\varepsilon) > 0$ are the eigenvectors to the (simple) target eigenvalue $\lambda(\varepsilon)$ of $A + \varepsilon E(\varepsilon)$ at the extremizer $E(\varepsilon)$; cf. [8, 12] for related derivative formulas. If the assumptions justifying this formula are not met, we can always resort to bijection. The algorithm we use is indeed a combined Newton / bisection approach, similar to [9, 8, 12].

7. Numerical experiments. In this section we show the behavior of Algorithm 1, which is based on the rank-1 differential equation (3.2), on two sparse matrices and an example with prescribed range and corange.

We start by considering two well-known sparse matrices.

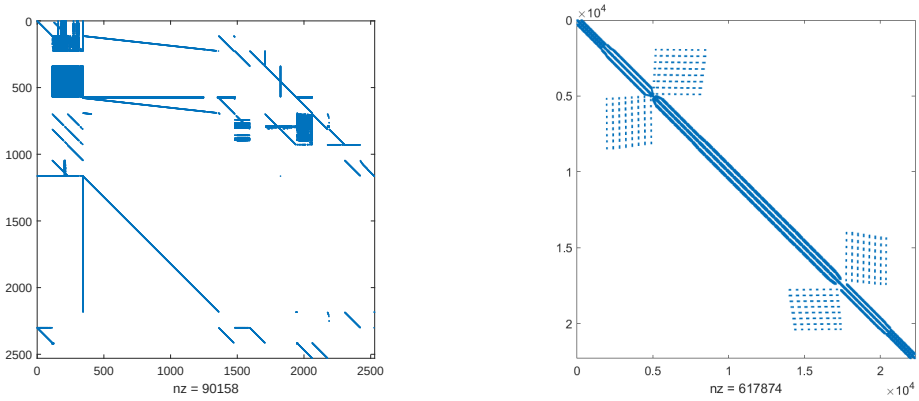

FIG. 7.1. *Sparsity patterns of the matrices ORANI678 (left) and FIDAPM11 (right).*

7.1. The matrix ORANI678 from the Harwell Boeing collection. The matrix A is a sparse real unsymmetric square matrix taken from the set ECONAUS. It has dimension $n = 2529$ and a number of nonzero entries $nz = 90158 \approx 40n$. Its sparsity pattern is plotted in Figure 7.1.

(i) We have set $\varepsilon = 1$ and applied our algorithms to the minimization problem (2.3) with $f(\lambda, \overline{\lambda}) = -\frac{1}{2}(\lambda + \overline{\lambda}) = -\mathrm{Re}(\lambda)$ and $\mathcal{S}$ the space of real matrices with the sparsity pattern of A. The target eigenvalue is the one with largest real part. We thus aim to compute the *structured ε-pseudospectral abscissa* of A with an accuracy of 14 digits. We denote by n_{eig} the total number of eigenvalue computations (that is

k	$\mathrm{Re}(\lambda_k)$
0	−1.232670912085709
1	−1.745212357950066
2	−1.917229680782718
3	−2.076407232182272
4	−**2**.249359154133923
5	−**2.3**43018078428841
6	−**2.3**43036033336665
7	−**2.3**49611649664635
8	−**2.350**556073486847
9	−**2.350**620017092603
...	...
25	−2.350634775262768

TABLE 7.1
Computed values using Algorithm 1 *for the ORANI678 matrix.*

the number of calls to the Matlab routine eigs).

We integrated (3.2) by Algorithm 1 and obtained the results in Table 7.1. The main cost is the number of eigentriplets evaluations by the Matlab routine eigs and is given by $n_{\mathrm{eig}} = 38$. The CPU time is around 1.5 seconds. For comparison we also integrated the full-rank ODE (2.15) by the Euler method (gradient descent) with variable stepsize and obtained a similar behavior. The number of eigentriplets

[1] Instead of eigenvalues of smallest modulus, we could take the smallest singular value.

evaluations is $n_{\text{eig}} = 35$ and the final approximation to the ε-pseudospectral abscissa is 2.350634775261177, which coincides with the value computed by the rank-1 method up to the 11-th digit. The CPU time is 1.6 seconds.

Since u and v turn out to be real, the gain in terms of memory requirements for the rank-1 algorithm is $90158/5058 \approx 17.82$, which is a significant reduction in the storage of the iterates.

(ii) Setting next $f(\lambda, \overline{\lambda}) = \lambda\overline{\lambda} = |\lambda|^2$ and the target eigenvalue the one - say $\lambda_{\min}$ - with smallest modulus, we approximated the *structured distance to singularity* of A. Given the convergence to a local optimizer of Algorithm 1 we obtain this way an upper bound to this distance. An immediate lower bound is the unstructured distance $\sigma_{\min}(A)$, i.e. the smallest singular value, which is equal to 0.0033388. As we see in Table 7.2, the effective structured distance to singularity is one order of magnitude larger. Applying a Newton-bisection method we obtained the results shown in Table 7.2. Since the function ϕ and its derivative (see (6.4)) are computed inexactly (by Algorithm 1), we do not observe quadratic convergence.

k	ε_k	$\phi(\varepsilon_k)$	# eigs
1	0.0104015	$1.1019564 \cdot 10^{-2}$	13
2	0.0176409	$9.5284061 \cdot 10^{-4}$	13
3	0.0219541	$2.5263758 \cdot 10^{-4}$	14
4	0.0243116	$6.5050153 \cdot 10^{-5}$	13
5	0.0255439	$1.6503282 \cdot 10^{-6}$	13
6	0.0261739	$4.1561289 \cdot 10^{-6}$	13
7	0.0264923	$1.0428313 \cdot 10^{-6}$	13
8	0.0266524	$2.6118300 \cdot 10^{-7}$	13
9	0.0267327	$6.5355110 \cdot 10^{-8}$	13
10	0.0267728	$9.6346192 \cdot 10^{-9}$	13
11	0.0267930	$1.7293467 \cdot 10^{-10}$	4

TABLE 7.2

Distance to singularity for the ORANI678 matrix: computed values ε_k, $\phi(\varepsilon_k) = |\lambda_{\min}(A + \varepsilon_k E_k)|^2$ and number of eigenvalue computations of the inner rank-1 algorithm.

The average CPU time of an outer iteration is around 528.6 seconds, which is due to the augmented computational cost required by the routine eigs for linear systems solves. The average number of eigentriplets evaluation is $n_{\text{eig}} = 12$.

7.2. The matrix FIDAPM11 from the SPARSKIT collection. The matrix A is now a sparse real unsymmetric square matrix taken from the set ECONAUS. It has dimension $n = 22294$ and a number of nonzero entries $nz = 623554 \approx 30n$. Its sparsity pattern is plotted in Figure 7.1.

We have set $\varepsilon = 0.5$ and applied our algorithms to the minimization problem (2.3) with $f(\lambda, \overline{\lambda}) = -\lambda\overline{\lambda} = -|\lambda|^2$ and $\mathcal{S}$ the space of real matrices with the sparsity pattern of A, and the target eigenvalue is the one with largest real part. We are thus aiming to compute the *structured ε-pseudospectral radius* of A.

Integrating both ODEs (2.15) and (3.2), we obtain the same optimizer $\lambda = 1.9716893$. The number of computed eigen-triplets is $n_{\text{eig}} = 107$ and $n_{\text{eig}} = 99$, with a slight advantage of the rank-1 method. The CPU time is close to 32.95 and 31.32 seconds respectively.

Also in this case u and v turn out to be real so that the gain in terms of memory requirements is significant, $623554/44588 \approx 13.98$.

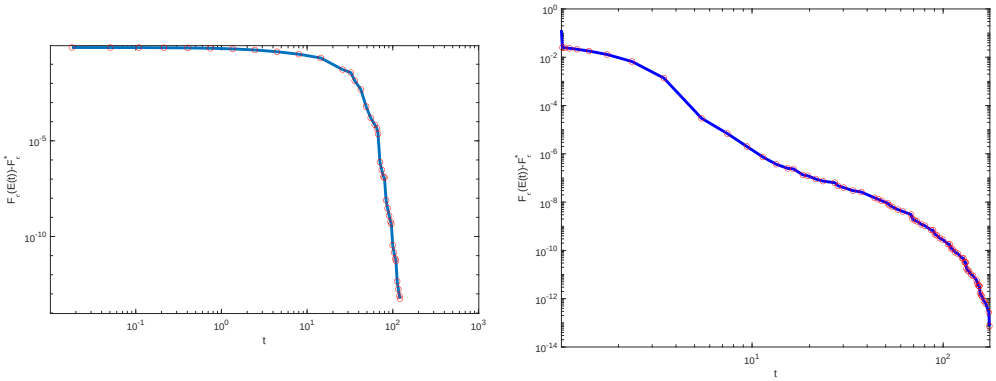

FIG. 7.2. *Behavior $F_\varepsilon(E(t)) - F_\varepsilon^*$ (where F_ε^* is the computed value of F_ε at the stationary point) in the numerical integration by Algorithm 1 for the matrix ORANI678 with $f(\lambda, \overline{\lambda}) = -\mathrm{Re}(\lambda)$ (left picture) and for the matrix FIDAPM11 with $f(\lambda, \overline{\lambda}) = -|\lambda|^2$ (right picture). In both cases $F_\varepsilon(E(t_k))$ decays monotonically with k.*

7.3. A comparison with Manopt. We made experiments on the sparse matrices considered above using Manopt, a well-known toolbox for optimization on manifolds and matrices [3].

Applying Manopt to the same problem considered for the ORANI678 matrix, providing the Riemannian gradient on the manifold of sparse matrices with unit Frobenius norm, the method yields a result very close to the one computed with our method (the difference is around 10^{-13}. The CPU time for our algorithm is approximately 1.4 seconds Concerning the algorithm implemented in Manopt, with the conjugate-gradient method. we have found that the method converges in 20 iterations using a CPU time which is approximately 14 seconds; with the BFGS solver it converges in 15 iterations using a CPU time of approximately 16 seconds and finally with the Barzilai-Borwein method it converges in 117 iterations in about 57 seconds. The trust region method (which is the default choice) instead turns out to converge very slowly.

Applying next Manopt, with the conjugate-gradient solver, to the considered example with the FIDAPM11 matrix, we have obtained that the result coincides with the one we obtain to 5 digits, but the CPU time exceeds 5 hours when the default accuracy is used, and it drops to around 8 minutes when a tolerance of 10^{-2} is required, which still gives 3 correct digits. With the option of adaptive line search and the same tolerance, a result with the same accuracy is obtained in a slightly larger CPU time.

7.4. An example of control of the Stokes problem. We consider an example from [15], which arises in the discretization of the 2-dimensional Stokes problem on a uniform quadratic grid. Setting 25 grid points on both sides of the square, we get a sparse matrix A ($J - R$ in the notation of [15]) which has dimension $n = 1824$, while we choose the control matrices B and $C = B^\top$ to have size $n \times k$ and $l \times n$, respectively with $k = l = 40$, randomly i.i.d. entries and unit Frobenius norm.

The matrix A has the rightmost eigenvalue $\lambda = -6.4343098 \cdot 10^{-4}$, which suggests a non-robust Hurwitz stability.

Although Assumption 4.1 is not fulfilled we succesfully execute the rank-1 algorithm. Running it on this example, we find the structured stability radius to be 0.0384039, which is 60 times larger than $|\lambda|$.

Since the matrix is sparse we can exploit favorably the matrix vector products of

the form (with $p \in \mathbb{R}^n$ the vector, $Z = uv^* \in \mathcal{M}_1$, and ρ the normalization factor)

$$\big(A + \varepsilon\rho B\, B^\dagger Z C^\dagger C\big)\; p,$$

whose cost is linear in n.

In Table 7.3 we show the Newton iteration where the number of eigentriplets evaluation is again indicated by $n_{\rm eig}$. The quadratically convergent behavior is evident.

k	ε_k	$\phi(\varepsilon_k)$	# eigs
0	0	$-6.4343098 \cdot 10^{-4}$	1
1	0.02	$-3.1062242 \cdot 10^{-4}$	23
2	0.0385299	$2.1414201 \cdot 10^{-6}$	26
3	0.0384039	$9.7779625 \cdot 10^{-11}$	18

TABLE 7.3

Iterates for computing the structured stability radius for the Stokes problem matrix with range- and corange-constrained perturbations; the optimal perturbation size ε^* is where $\phi(\varepsilon^*) = 0$.

8. Use of the projected gradient system (2.19) for eigenvalue optimization with perturbation matrices of prescribed range and co-range. We consider the (complex) structure space $\mathcal{S}$ of (2.5), which only allows for perturbations of given range and co-range. We recall that the orthogonal projection onto $\mathcal{S}$ is given by $\Pi^{\mathcal{S}} Z = BB^\dagger Z C^\dagger C$. In this case, the set $\mathcal{M}_1^{\mathcal{S}} = \Pi^{\mathcal{S}}\mathcal{M}_1$ of structure-projected rank-1 matrices equals the submanifold of rank-1 matrices that have the prescribed range and co-range:

$$\mathcal{M}_1^{\mathcal{S}} = \{E \in \mathbb{C}^{n,n} \,:\, E = \rho uv^* \;\text{ with }\; \rho > 0,\; u \in \mathrm{Ran}(B),\; v \in \mathrm{Ran}(C)\} \subset \mathcal{M}_1.$$

For such an $E = \rho uv^*$ with u and v of unit norm and in the range of B and C, respectively, the orthogonal projection $P_E^{\mathcal{S}}$ onto the tangent space $T_E\mathcal{M}_1^{\mathcal{S}}$ turns out to be given by the same expression as in (3.3):

$$P_E^{\mathcal{S}}(Z) = Z - (I - uu^*)Z(I - vv^*).$$

This has important consequences:

– On the theoretical side, it allows us to use the same argument as in the proof of part (b) of Theorem 3.1 to show, under the non-degeneracy condition $P_E^{\mathcal{S}} G^{\mathcal{S}}(E) \neq 0$, that every stationary point of the gradient system (2.19) on $\mathcal{M}_1^{\mathcal{S}}$ is also a stationary point of the gradient system (2.15) on $\mathcal{S}$; hence, there are no (non-degenerate) spurious stationary points.

– On the computational side, for a solution $E(t) = u(t)v(t)^* \in \mathcal{M}_1^{\mathcal{S}}$ of unit Frobenius norm of the differential equation (2.19) we therefore obtain differential equations for the factors u and v of unit norm that are formally the same as in Lemma 3.2: with the projected gradient $G^{\mathcal{S}} = \Pi^{\mathcal{S}} G(E)$ for short,

$$\begin{aligned}
\dot u &= -(I - uu^*)G^{\mathcal{S}} v - \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^* G^{\mathcal{S}} v)u,\\
\dot v &= -(I - vv^*)(G^{\mathcal{S}})^* u + \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(u^* G^{\mathcal{S}} v)v.
\end{aligned}$$

Note that here $\dot u$ and $\dot v$ are in the range of B and C, respectively, so that u and v stay in these ranges. In order to obtain a further compression we set $u = Bp$ and

$v = C^*q$ with $p \in \mathbb{C}^k$ and $q \in \mathbb{C}^l$. In this way - with $G = G_\varepsilon(E)$ - we obtain for $E = uv^* = Bpq^*C$ the differential equations

$$\begin{aligned}\dot{p} &= -B^\dagger GC^*q \;+\; pp^*B^*GC^*q \;-\; \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(p^*B^*GC^*q)p\\ \dot{q} &= -(C^*)^\dagger G^*Bp + qq^*CG^*Bp \;+\; \tfrac{\mathrm{i}}{2}\mathrm{Im}\,(p^*B^*GC^*q)q.\end{aligned}$$

With the rank-1 matrix $G = G_\varepsilon(E) = 2f_{\overline{\lambda}}\,xy^*$ (see Lemma 2.4) and with $\alpha = p^*B^*x$, $\beta = q^*Cy$ and $\gamma = 2f_{\overline{\lambda}}$, we thus obtain the differential equations (cf. (3.4))

$$\begin{aligned}\dot{p} &= \alpha\overline{\beta}\gamma\, p - \overline{\beta}\gamma\, B^\dagger x - \tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\alpha\overline{\beta}\gamma)p\\ \dot{q} &= \overline{\alpha}\beta\overline{\gamma}\, q - \overline{\alpha\gamma}\,(C^*)^\dagger y - \tfrac{\mathrm{i}}{2}\,\mathrm{Im}\,(\overline{\alpha}\beta\overline{\gamma})q.\end{aligned} \tag{8.1}$$

This system of differential equations is treated numerically in the same way as described in Section 5, using a splitting between the first two terms on the right-hand side and the third term.

We present numerical results for the Stokes example of Section 7.4, now treated with the above implementation of the gradient system (2.19) for comparison.

k	ε_k	$\phi(\varepsilon_k)$	# eigs
1	0.02	$-3.1061082 \cdot 10^{-4}$	26
2	0.0386113	$4.3234455 \cdot 10^{-5}$	33
3	0.0384036	$8.2212343 \cdot 10^{-10}$	16

TABLE 8.1

Iterates for computing the structured stability radius for the Stokes problem matrix with range- and corange-constrained perturbations with inner iteration realized integrating (8.1).

Acknowledgments. We are grateful to two anonymous referees for their constructive comments on a previous version of this paper.

Nicola Guglielmi acknowledges that his research was supported by funds from the Italian MUR (Ministero dell'Università e della Ricerca) within the PRIN-17 Project "Discontinuous dynamical systems: theory, numerics and applications". He also acknowledges affiliation to INdAM Research group GNCS (Gruppo Nazionale di Calcolo Scientifico).

REFERENCES

[1] P.-A. Absil, R, Mahony and R. Sepulchre. Optimization algorithms on matrix manifolds. *Princeton University Press, Princeton, NJ*, 2008.

[2] R. Alam, S. Bora, M. Karow, V. Mehrmann, and J. Moro. Perturbation theory for Hamiltonian matrices and the distance to bounded-realness. SIAM J. Matrix Anal. Appl., 32(2): 484–514, 2011.

[3] N. Boumal, B. Mishra, P.-A. Absil and R. Sepulchre. Manopt, a Matlab Toolbox for Optimization on Manifolds. Journal of Machine Learning Research, 42(15): 1455–1459, 2014.

[4] R.W. Brockett, Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. Linear Algebra Appl. 146 (1991), 79–91.

[5] E. J. Candès and T. Tao. The power of convex relaxation: near-optimal matrix completion IEEE Trans. Inform. Theory, 56(5): 2053–2080, 2010.

[6] M.T. Chu, Linear algebra algorithms as dynamical systems. Acta Numer. 17 (2008), 1–86.

[7] A. Greenbaum, R.-C. Li, and M. L. Overton. First-order perturbation theory for eigenvalues and eigenvectors. SIAM Rev., 62(2): 463–482, 2020.

[8] N. Guglielmi. On the method by Rostami for computing the real stability radius of large and sparse matrices. SIAM J. Sci. Comput., 38(3): A1662–A1681, 2016.

[9] N. Guglielmi, D. Kressner, and C. Lubich. Low rank differential equations for Hamiltonian matrix nearness problems. Numer. Math., 129(2): 279–319, 2015.
[10] N. Guglielmi and C. Lubich. Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. SIAM J. Numer. Anal., 49: 1194–1209, 2011.
[11] N. Guglielmi and C. Lubich. Low-rank dynamics for computing extremal points of real pseudospectra. SIAM J. Matrix Anal. Appl., 34(1): 40–66, 2013.
[12] N. Guglielmi and C. Lubich. Matrix stabilization using differential equations. SIAM J. Numer. Anal., 55: 3097–3119, 2017.
[13] N. Guglielmi, C. Lubich, Matrix nearness problems and eigenvalue optimization. Book in preparation, 2022.
[14] 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(4): 1166–1192, 2011.
[15] S. A. Hauschild, N. Marheineke and V. Mehrmann, Model reduction techniques for linear constant coefficient port-Hamiltonian differential-algebraic systems. Control and Cybernetics, 32(1): 125+, 2019.
[16] D. Hinrichsen and A. J. Pritchard. Mathematical systems theory I: modelling, state space analysis, stability and robustness. Springer, Berlin, 2005.
[17] U. Helmke, J.B. Moore, Optimization and dynamical systems. *Communications and Control Engineering Series*. Springer-Verlag, London, 1994.
[18] D. Hinrichsen and A.J. Pritchard. Mathematical systems theory. I, Modelling, state space analysis, stability and robustness. Volume 48 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2005.
[19] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.
[20] O. Koch and C. Lubich, Dynamical low-rank approximation. SIAM J. Matrix Anal. Appl., 29(2): 434–454, 2007.
[21] H.-O. Kreiss, Über die Stabilitätsdefinition für Differenzengleichungen, die partielle Differentialgleichungen approximieren. BIT 2 (1962), 153–181.
[22] D. Kressner and B. Vandereycken. Subspace methods for computing the pseudospectral abscissa and the stability radius. SIAM J. Matrix Anal. Appl., 35(1): 292–313, 2014.
[23] C. D. Meyer and G. W. Stewart, Derivatives and perturbations of eigenvectors. SIAM J. Numer. Anal., 25(3): 679–691, 1988.
[24] V. Mehrmann and H. Xu. Perturbation of purely imaginary eigenvalues of Hamiltonian matrices under structured perturbations. Electron. J. Lin. Alg., 17: 234–257, 2008.
[25] L. Qiu, B. Bernhardsson, A. Rantzer, E. J. Davison, P. M. Young, J. C. Doyle, A formula for computation of the real stability radius. Automatica J. IFAC, 31(6): 879–890, 1995.
[26] G. W. Stewart, A Krylov-Schur algorithm for large eigenproblems, SIAM J. Matrix Anal. Appl., 23(3), 601–614, 2001/02.
[27] L.N. Trefethen, M. Embree, Spectra and pseudospectra. The behavior of nonnormal matrices and operators. Princeton University Press, Princeton, NJ, 2005.
[28] T.G. Wright. Eigtool: a graphical tool for nonsymmetric eigenproblems. *Oxford University Computing Laboratory, http://www.comlab.ox.ac.uk/pseudospectra/eigtool/*, 2002.