

LINEARIZING A NONLINEAR EIGENVALUE PROBLEM WITH QUADRATIC RATIONAL EIGENVECTOR NONLINEARITIES*

VICTOR JANSSENS[†], KARL MEERBERGEN[†], AND WIM MICHIELS[†]

Abstract. Nonlinear eigenvalue problems with eigenvector nonlinearities (NEPv) are algebraic eigenvalue problems whose matrix depends on the eigenvector. Applications range from computational quantum mechanics to machine learning. Due to its nonlinear behavior, existing methods almost exclusively rely on fixed-point iterations, the global convergence properties of which are only understood in specific cases. Recently, a certain class of NEPv with linear rational eigenvector nonlinearities has been linearized, i.e., the spectrum of the linear eigenvalue problem contains the eigenvalues of the NEPv. This linear problem is solved using structure exploiting algorithms to improve both convergence and reliability. We propose a linearization for a different class of NEPv with quadratic rational nonlinearities, inspired by the discretized Gross-Pitaevskii equation. The eigenvalues of this NEPv form a subset of the spectrum of a linear multiparameter eigenvalue problem which is equivalent to a system of generalized eigenvalue problems expressed in terms of operator determinants. A structure exploiting Arnoldi algorithm is used to filter a large portion of spurious solutions and to accelerate convergence.

Key words. nonlinear eigenvalue problems, multiparameter eigenvalue problems, linearization

MSC codes. 65F15, 65H17

1. Introduction. We revisit a specific class of the eigenvector dependent nonlinear eigenvalue problem (NEPv), which involves solving

$$(1.1) \quad Av = \lambda Bv + \sum_{i=1}^m f_i(v) C_i v,$$

where $A, B, C_i \in \mathbb{C}^{n \times n}$ are complex-valued matrices of size n and where the functions $f_i : \mathbb{C}^n \rightarrow \mathbb{C}$ are scaling invariant for $i \in \{1, 2, 3, \dots, m\}$, i.e., $f(\alpha v) = f(v)$ for all nonzero $v \in \mathbb{C}^n \setminus \{0\}$ and all nonzero $\alpha \in \mathbb{C} \setminus \{0\}$ [7]. The eigenpair (λ, v) solves this eigenvalue problem if (1.1) holds with $\lambda \in \mathbb{C}$ a complex number and $v \in \mathbb{C} \setminus \{0\}$ a nonzero vector.

More general forms of the NEPv appear in quantum mechanical applications such as electronic structure calculations [20, 27] and Bose-Einstein condensates [4, 13]. Other applications in machine learning are p-spectral clustering [5] and the trace-ratio optimization problem [24].

Most existing methods rely on fixed point iterations of which the convergence properties are only understood in specific cases. A popular example is the Self-Consistent Field (SCF) iteration, which was originally developed to solve the Hartree and Fock equations in electronic structure calculations [12, 27]. The idea is to guess an initial eigenvector and substitute it in the nonlinear part of the NEPv. This gives a generalized eigenvalue problem (GEP) which can be solved to obtain the eigenvector for the next iteration. Repeating this process until the eigenvector does not change any longer after an iteration, gives the solution of this fixed point problem. Convergence of the SCF method is studied, see [6, 3], e.g., but it often does not converge in its original form, which is why it is usually combined with acceleration techniques

*This work was supported by the project G027624N of the Research Foundation-Flanders (FWO - Vlaanderen).

[†]Department of Computer Science, KU Leuven, Leuven, Belgium (victor.janssens@kuleuven.be, karl.meerbergen@kuleuven.be, wim.michiels@cs.kuleuven.be).

such as Direct Inversion of the Iterative Subspace (DIIS) [25]. Other methods are generalizations of the inverse iteration method [18] and contour integration methods [8].

A recent study [7] shows that if the functions in the NEPv (1.1) are rational with a linear numerator and denominator in the eigenvector, i.e.

$$f_i(v) = \frac{r_i^T v}{s_i^T v},$$

with vectors $r_i, s_i \in \mathbb{C}^n$, the problem can be linearized into a GEP, that is, all eigenvalues of the NEPv are also eigenvalues of the GEP. This is useful as the NEPv can then be solved using efficient and reliable methods developed for the GEP. In [7], a slight modification to inverse iteration is proposed that exploits the structure of the linearized problem to speed up convergence.

In this paper, we propose a linearization of (1.1) with one quadratic rational nonlinearity, i.e.,

$$(1.2) \quad f_i(v) = \frac{v^H P_i v}{v^H Q_i v},$$

where $m = 1$, $A, C_i, P_i \in H_n = \{X \in \mathbb{C}^{n \times n} | X = X^H\}$ are Hermitian and $B, Q_i \in H_n^+ = \{X \in H_n | X \succ 0\}$ positive definite. This type of eigenvector nonlinearity is closely related to a discretization of the Gross-Pitaevskii equation (GPE), also referred to as the nonlinear Schrödinger equation, which needs to be solved in order to obtain the ground state and the wave function of a so called Bose-Einstein condensate (BEC), see for instance [4, 13] and references therein. In case C_i is a rank one matrix and $C_i = P_i$, the eigenvalues of the NEPv can be obtained by solving an eigenvalue problem with eigenvalue nonlinearities (NEP λ) instead [19] for which efficient methods are available. This approach works very well for this low rank case, whereas our linearization is better suited when the rank of C_i is greater than or equal to $n - 1$ and $m = 1$.

We will show how the NEPv with scalar nonlinearities (1.2) for $m = 1$ is related to a multiparameter eigenvalue problem (MEP), which can be constructed directly from the coefficient matrices A, B, C_1, P_1 and Q_1 and an additional matrix R that can be chosen freely as long as its rank is full. Using operator determinants, the MEP may then be converted into a GEP. The paper proposes a modification to the Arnoldi algorithm [1] and exploits the structure of the resulting GEP to obtain the eigenvalues of the NEPv closest to a chosen shift by filtering out some of the spurious solutions.

The paper is organized as follows. An upper bound on the number of eigenvalues of (1.1) with eigenvector nonlinearity (1.2) and the link with a system of polynomial equations are discussed in Section 2. As a first main contribution, Section 3 proposes the linearization which contains all the eigenvalues of the NEPv. The resulting problem is solved using structure-exploiting Arnoldi algorithms discussed in Section 4, which is the second contribution. The theoretical results are supported by numerical examples in Section 5, and lastly Section 6 summarizes the results.

2. Problem properties. Given the Hermitian matrices $A, C, P \in H_n$ and positive definite matrices $B, Q \in H_n^+$, find $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n \setminus \{0\}$ such that

$$(2.1) \quad Av = \lambda Bv + \frac{v^H P v}{v^H Q v} Cv.$$

If such values exist, we say that (λ, v) solves the NEPv (2.1). The following lemma states that this NEPv is equivalent to the system of nonlinear equations

$$(2.2) \quad \begin{cases} M(\lambda, \mu)v = 0, \\ v^H S(\mu)v = 0, \end{cases}$$

where

$$(2.3) \quad \begin{aligned} M(\lambda, \mu) &= A - \lambda B - \mu C, \\ S(\mu) &= P - \mu Q, \end{aligned}$$

which we will write in short as M and S if clear from context.

LEMMA 2.1. *The eigenpair (λ, v) solves (2.1) if and only if there exists a scalar $\mu \in \mathbb{C}$ such that (λ, μ, v) solves (2.2). Moreover, μ equals $\frac{v^H P v}{v^H Q v}$ and both λ and μ are real-valued if they solve the NEPv.*

Proof. If (λ, v) solves (2.1) and we define $\mu = \frac{v^H P v}{v^H Q v}$, then we have $Av = \lambda Bv + \mu Cv$ and $v^H P v = \mu v^H Q v$ which implies that (λ, μ, v) solves (2.2).

Conversely, assume (λ, μ, v) solves (2.2), then $v^H P v = \mu v^H Q v$, which can also be written as $\mu = \frac{v^H P v}{v^H Q v}$ since $v^H Q v$ cannot equal zero¹. Eliminating μ from the first equation in (2.2) gives (2.1).

Because both P and Q are positive definite, the value $\mu = \frac{v^H P v}{v^H Q v}$ equals its complex conjugate $\bar{\mu}$, and thus μ is real-valued. Multiplying (2.1) with v^H on the left and dividing² by $v^H B v$, gives

$$\lambda = \frac{v^H (A - \mu C) v}{v^H B v},$$

and therefore λ is real-valued as well because $A - \mu C$ and B are Hermitian. \square

In order to obtain an upper bound on the number of isolated solutions of the NEPv (2.2), we show how the eigenvalues are related to the solutions of the system of polynomial equations

$$(2.4) \quad \begin{cases} \det(M(\lambda, \mu)) = 0, \\ \text{trace}(S(\mu) \text{adj}(M(\lambda, \mu))) = 0. \end{cases}$$

LEMMA 2.2. *If (λ, μ, v) solves the NEPv (2.2) with v nonzero, then (λ, μ) solves the equations (2.4).*

Proof. Given the solution (λ, μ, v) which satisfies (2.2), we have that $\det(M) = 0$ because $Mv = 0$ for nonzero v . To obtain the second equality in (2.4), we use the fact that the adjugate of the singular matrix M can be written as $\alpha v v^H$ for some complex number $\alpha \in \mathbb{C}$ [17, page 22]. From (2.2), we have

$$\begin{aligned} v^H S v &= 0, \\ \iff \text{trace}(S v v^H) &= 0, \\ \iff \text{trace}(S \text{adj}(M)) &= 0, \end{aligned}$$

and therefore (λ, μ) solves the equations (2.4). \square

¹As the eigenvector v is nonzero and Q is positive definite, the value $v^H Q v$ is strictly positive.

²This division is allowed due to v being nonzero and B positive definite.

The converse is also true under a specific condition and if the eigenvalues are real-valued.

PROPOSITION 2.3. *Let $\lambda, \mu \in \mathbb{R}$ be real-valued numbers such that (2.4) holds, and define $\hat{S} = V^H S(\lambda, \mu) V$ where the columns of $V \in \mathbb{C}^{n \times k}$ form a basis for the null space of $M(\lambda, \mu)$, then there exists a nonzero vector $v \in \mathbb{C}^n \setminus \{0\}$ such that (λ, μ, v) solves (2.2) if and only if \hat{S} is neither positive nor negative definite. Moreover, if V has one column, i.e., $k = 1$, then the condition $\pm \hat{S} \neq 0$ is automatically satisfied.*

Proof. We start by proving the if and only if statement from left to right. Assume (λ, μ, v) solves the NEPv (2.2), then both equations $Mv = 0$ and $v^H S v = 0$ hold. The first equation implies $v \in \text{null}(M)$ and thus there exists a nonzero vector $x \in \mathbb{C}^k \setminus \{0\}$ such that $v = Vx$. The other equation $v^H S v = 0$ is equivalent to $x^H \hat{S} x = 0$ and thus $\pm \hat{S} \neq 0$.

Conversely, assume $\pm \hat{S} \neq 0$ then there exists a nonzero vector $x \in \mathbb{C}^k \setminus \{0\}$ such that $x^H \hat{S} x = 0$, which is equivalent to $x^H V^H S V x = 0$. Note that $MVx = 0$ since V is a basis for the null space of M and therefore (λ, μ, Vx) solves the NEPv (2.2).

For the case where $k = 1$, there must exist a nonzero v such that $Mv = 0$ and $\text{adj}(M) = vv^H$ [17]. From (2.4) the equation $\text{trace}(S \text{adj}(M)) = 0$ holds and thus $\hat{S} = v^H S v = \text{trace}(S vv^H) = \text{trace}(S \text{adj}(M)) = 0$ which is the same as $\pm \hat{S} \neq 0$. \square

THEOREM 2.4. *If the system of polynomial equations (2.4) has a finite number of solutions, the NEPv (2.1) has at most n^2 eigenvalues.*

Proof. From Bézouts theorem, the upper bound on the number of isolated solutions of the system of polynomial equations (2.4) is n^2 since both polynomials have a degree of n . If (λ, v) solves (2.1), then from Lemma 2.1 and Lemma 2.2, the pair $(\lambda, \frac{v^H P v}{v^H Q v})$ solves (2.4). Therefore an upper bound on the number of eigenvalues of the NEPv is n^2 . \square

An interesting fact is that because A, B, C, P and Q are Hermitian, this system of polynomial equations (2.4) in λ and μ has real-valued coefficients, which can be explained as follows. Let $f(\lambda, \mu) = \det(A - \lambda B - \mu C)$, then for any $\lambda, \mu \in \mathbb{C}$

$$\overline{f(\lambda, \mu)} = \det(\overline{A} - \overline{\lambda B} - \overline{\mu C}) = \det(A^\top - \overline{\lambda} B^\top - \overline{\mu} C^\top) = f(\overline{\lambda}, \overline{\mu}).$$

As a result, the coefficients of f have an imaginary part equal to zero. Similarly, let $g(\lambda, \mu) = \text{trace}((P - \mu Q) \text{adj}(A - \lambda B - \mu C))$, then $\overline{g(\lambda, \mu)} = g(\overline{\lambda}, \overline{\mu})$, so the same conclusion can be drawn for the second equation.

Note that we could solve the NEPv by solving the polynomial system (2.4), for instance, using homotopy continuation methods [22]. However, computing its coefficients can be time consuming and numerically unstable because of the evaluation of a determinant and an adjugate in the equations, which further motivates the linearization approach in Section 3.

Example 2.5. Proposition 2.3 indicates that there may be cases where not all real-valued eigenvalues of the system of polynomial equations (2.4) correspond to eigenvalues of the NEPv (2.2). As an example, consider the matrices

$$A = \begin{bmatrix} 8 & -9 - 6i \\ -9 + 6i & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 8 & 3 + 2i \\ 3 - 2i & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 6 + 4i \\ 6 - 4i & 6 \end{bmatrix}$$

$$P = \begin{bmatrix} -4 & 6 + 6i \\ 6 - 6i & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 6 & -3 - 2i \\ -3 + 2i & 3 \end{bmatrix}.$$

In this case $(\lambda, \mu) = (1, -2)$ is a real-valued solution of the equations

$$\begin{cases} \det(M) = 51\lambda^2 - 4\lambda\mu - 52\mu^2 - 110\lambda - 204\mu - 149 = 0, \\ \text{trace}(S \text{ adj}(M)) = 98\lambda\mu + 88\mu^2 + 92\lambda + 222\mu + 196 = 0. \end{cases}$$

Note that $M(1, -2) = 0$, i.e., the dimension of its null space equals 2 and thus the columns of the identity matrix I can be chosen as a basis for the null space of M . This means that $\hat{S} = IS(-2)I = P + 2Q$, as defined in Proposition 2.3. Since \hat{S} has strictly positive eigenvalues, it is positive definite, which means $\lambda = 1$ is not an eigenvalue of (2.2).

Example 2.6. The upper bound of n^2 isolated solutions for the NEPv (2.1) can be reached in certain cases. For instance, when $n = 2$ and the NEPv matrices are

$$A = \begin{bmatrix} 4 & 3+i \\ 3-i & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 16 & 2-2i \\ 2+2i & 9 \end{bmatrix}, \quad C = \begin{bmatrix} -8 & 5-10i \\ 5+10i & -17 \end{bmatrix}$$

$$P = \begin{bmatrix} 6 & -1+18i \\ -1-18i & 4 \end{bmatrix}, \quad Q = \begin{bmatrix} 6 & 2+i \\ 2-i & 4 \end{bmatrix},$$

the problem has 4 different eigenvalues. These are given in Table 1 with the corresponding value of μ and eigenvector $v = (v_1 \ v_2)^\top$.

	λ	μ	v_1	v_2
Solution 1	11.936	4.0164	$0.0438 + 0.4424i$	$0.7073 - 0.5497i$
Solution 2	-0.0684	0.0207	$0.5209 - 0.0291i$	$-0.7875 + 0.3282i$
Solution 3	0.1906	-1.4229	$0.7836 + 0.3013i$	$0.1508 + 0.5220i$
Solution 4	0.2612	-0.3510	$0.3963 - 0.7437i$	$0.4312 - 0.3225i$

TABLE 1
All solution of the NEPv (2.1) in Theorem 2.6.

3. Linearization. This section makes the link between the NEPv and a GEP. While the linearized problem contains all the eigenvalues (at most n^2) of the NEPv, the opposite is not true since the GEP we propose in Subsection 3.2 has size $2n^2 - n$, i.e., it may contain eigenvalues that do not correspond to a solution of the NEPv. We will refer to these eigenvalues as spurious solutions and their origin is explained in Subsection 3.3. Fortunately, these spurious solutions have a special structure in the eigenvectors of the GEP which will be exploited in numerical methods proposed in Section 4.

3.1. Kronecker linearization. The NEPv with quadratic eigenvector nonlinearities can be reformulated as a NEPv similar to those studied in [7]. This allows us to apply the linearization technique described there to our problem as well. However, the resulting linearized problem is singular and exhibits rapid growth in size with respect to the dimension n . While there are workarounds for both issues, the details of which are omitted here, the scaling challenge remains significant. This motivates the more compact linearization approach presented in Subsection 3.2.

Starting from the original NEPv (2.1), the conversion to the different class of NEPv can be realized as follows: Let $\mu = \frac{v^H P v}{v^H Q v}$ as in Section 2 and define the vector $u \in \mathbb{C}^{n^2}$ as $u = \bar{v} \otimes v$, then $v^H P v$ equals $p^H u$ where $p = \text{vec}(P) \in \mathbb{C}^{n^2}$ is a vector

containing all columns of P stacked on top of each other. Similarly, $v^H Q v = q^H u$ with $q = \text{vec}(Q)$. From Lemma 2.1, both λ and μ are real-valued, and therefore

$$\begin{aligned} (\bar{A} \oplus A)u &= (\bar{A} \otimes I + I \otimes A)(\bar{v} \otimes v) = \bar{A}v \otimes v + \bar{v} \otimes Av, \\ &= (\lambda \bar{B}v + \mu \bar{C}v) \otimes v + \bar{v} \otimes (\lambda Bv + \mu Cv), \\ &= \lambda(\bar{C} \oplus C)u + \mu(\bar{C} \oplus C)u. \end{aligned}$$

Eliminating μ from the equations gives the NEP_v with eigenvector nonlinearities described in [7]:

$$(\bar{A} \oplus A)u = \lambda(\bar{C} \oplus C)u + \frac{p^H u}{q^H u}(\bar{C} \oplus C)u.$$

Linearizing this NEP_v results in a GEP of size n^4 which scales worse than the linearization of size $2n^2 - n$ proposed in the next section. Therefore, the remaining parts of this paper will focus on the latter.

3.2. Compact linearization. First, the NEP_v (2.1) is converted into a MEP. Let $R \in \mathbb{C}^{n \times (n-1)}$ be a matrix of full column rank, then the linearization is defined as the MEP

$$(3.1) \quad \begin{cases} M(\lambda, \mu)v = 0, \\ \begin{bmatrix} 0 & R^H M(\lambda, \mu) \\ M(\lambda, \mu)R & S(\mu) \end{bmatrix} w = 0. \end{cases}$$

This is a two-parameter eigenvalue problem with solution (λ, μ, v, w) , i.e., eigentuple $(\lambda, \mu) \in \mathbb{C} \times \mathbb{C}$ and eigenvectors $v \in \mathbb{C}^n \setminus \{0\}$, $w \in \mathbb{C}^{2n-1} \setminus \{0\}$, which is clearer to see if we rewrite (3.1) using $M(\lambda, \mu) = A - \lambda B - \mu C$ and $S(\mu) = P - \mu Q$ as

$$\begin{cases} Av = \lambda Bv + \mu Cv, \\ \hat{A}w = \lambda \hat{B}w + \mu \hat{C}w, \end{cases}$$

where

$$\hat{A} = \begin{bmatrix} 0 & R^H A \\ AR & P \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0 & R^H B \\ BR & 0 \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} 0 & R^H C \\ CR & Q \end{bmatrix}.$$

In general, the number of isolated solutions equals $2n^2 - n$ [2]. In what follows, we will show how the NEP_v (2.2) is related to this MEP.

THEOREM 3.1. *If (λ, v) solves the NEP_v (2.1), then there exists a scalar $\mu \in \mathbb{R}$ and a vector $w \in \mathbb{C}^{2n-1} \setminus \{0\}$ such that (λ, μ, v, w) solves the MEP (3.1) and such that w equals $\begin{bmatrix} w_1^T & \alpha v^T \end{bmatrix}^T$ for some $w_1 \in \mathbb{C}^{n-1}$ and $\alpha \in \mathbb{C}$.*

Proof. Suppose (λ, v) solves (2.1), then it follows from Lemma 2.1 that λ is real-valued and that there exists a scalar $\mu \in \mathbb{R}$ such that

$$\begin{cases} M(\lambda, \mu)v = 0, \\ v^H S(\mu)v = 0, \end{cases}$$

where M and S are defined in (2.3). Because M is Hermitian, the vector v must lie in the left null space of the square matrix $\begin{bmatrix} MR & Sv \end{bmatrix}$, i.e., $v^H \begin{bmatrix} MR & Sv \end{bmatrix} = 0$ with

$R \in \mathbb{C}^{n \times (n-1)}$ a matrix of full rank. As a result, there exists a vector $w_1 \in \mathbb{C}^{n-1}$ and a scalar $\alpha \in \mathbb{C}$ such that $\begin{bmatrix} w_1^\top & \alpha \end{bmatrix}^\top$ lies in the right null space of this matrix, that is, $MRw_1 + \alpha Sv = 0$. Also since $Mv = 0$, the equality $\alpha R^H Mv = 0$ holds and therefore

$$\begin{cases} Mv = 0, \\ \begin{bmatrix} 0 & R^H M \\ MR & S \end{bmatrix} \begin{bmatrix} w_1 \\ \alpha v \end{bmatrix} = 0, \end{cases}$$

meaning $(\lambda, \mu, v, \begin{bmatrix} w_1^\top & \alpha v^\top \end{bmatrix}^\top)$ solves the MEP (3.1). \square

We thus have a linear problem of which the solutions contain all eigenvalues of the original nonlinear problem. Of course, since our MEP has more isolated solutions than the NEPv, the converse cannot be true unless some assumptions are specified.

PROPOSITION 3.2. *Assume λ and μ are real-valued, that $\text{rank}(M(\lambda, \mu))$ equals $\text{rank}(M(\lambda, \mu)R)$ and that $w = \begin{bmatrix} w_1^\top & w_2^\top \end{bmatrix}^\top \in \mathbb{C}^{2n-1}$ is some vector such that $w_2 \in \mathbb{C}^n$ is nonzero. If (λ, μ, v, w) solves the MEP (3.1), then (λ, w_2) solves (2.1).*

Proof. Since (λ, μ, v, w) solves (3.1) the equations

$$(3.2a) \quad Mv = 0,$$

$$(3.2b) \quad R^H Mw_2 = 0,$$

$$(3.2c) \quad MRw_1 + Sw_2 = 0,$$

must hold. From the assumptions we have that M is Hermitian and that the ranks of M , MR and $R^H M$ are equal. By consequence, M and $R^H M$ must share the same null space, hence $Mw_2 = 0$ and $w_2^H M = 0$ because of (3.2b). The vector w_2 cannot be the zero vector and therefore (λ, μ, w_2) solves

$$\begin{cases} Mw_2 = 0, \\ w_2^H Sw_2 = 0. \end{cases}$$

Lemma 2.1 then states that (λ, w_2) is a solution of (2.1). \square

Remark 3.3. Note that the assumption that $w_2 \neq 0$ in the last proposition is automatically satisfied if $\text{rank}(MR) = n - 1$. Indeed, if we assume w_2 is zero anyway, we have that $MRw_1 = 0$ from (3.2c) which is only possible if w_1 is zero since MR is of full rank. But now the vector w is zero which contradicts with the fact that eigenvectors must be nonzero and we must deduce that $w_2 \neq 0$.

Remark 3.4. If (λ, μ, v, w) solves the MEP (3.1), then the matrix $M(\lambda, \mu)$ is singular. Hence, the assumption that the rank of M must equal the rank of MR in Proposition 3.2 is satisfied with probability one for a randomly chosen matrix $R \in \mathbb{C}^{n \times (n-1)}$.

The linearization (3.1) discussed up until now can thus be solved using MEP solving techniques [14, 28, 10, 26], but in this article we will convert it to a GEP that is solved in Section 4 to filter out a portion of the spurious solutions. This conversion can be obtained after defining the operator determinants

$$\begin{cases} \Delta_0 = B \otimes \hat{C} - C \otimes \hat{B}, \\ \Delta_1 = A \otimes \hat{C} - C \otimes \hat{A}, \\ \Delta_2 = B \otimes \hat{A} - A \otimes \hat{B}. \end{cases}$$

THEOREM 3.5. [2, Chapter 6] If Δ_0 is nonsingular, the MEP (3.1) is equivalent to the system of generalized eigenvalue problems

$$(3.3) \quad \begin{cases} \Delta_1 z = \lambda \Delta_0 z, \\ \Delta_2 z = \mu \Delta_0 z, \end{cases}$$

with $z = v \otimes w$.

Since we are interested in the eigenvalue λ , it suffices to solve the first GEP in (3.3)

$$(3.4) \quad \Delta_1 z = \lambda \Delta_0 z,$$

which will be referred to as the linearization of the class of NEPv (2.1). The methods developed in Section 4 solve this GEP using the Arnoldi iteration to obtain the NEPv eigenvalues. In order for this to work, the pencil must be nonsingular which is true if Δ_0 is nonsingular as stated in Theorem 3.5. The following lemma and theorem shed light on the invertibility of Δ_0 , but before providing both statements, recall that

$$\hat{B} = \begin{bmatrix} 0 & R^H B \\ BR & 0 \end{bmatrix} \quad \text{and} \quad \hat{C} = \begin{bmatrix} 0 & R^H C \\ CR & Q \end{bmatrix}.$$

LEMMA 3.6. Let $\lambda \in \mathbb{C}$ be some complex number. There exists a nonzero vector $z \in \mathbb{C}^{2n-1} \setminus \{0\}$ such that (λ, z) solves $\hat{C}z = \lambda \hat{B}z$ with $\lambda \in \mathbb{R}$ if and only if (λ, x) solves $CRx = \lambda BRx$ for some $x \in \mathbb{C}^{n-1} \setminus \{0\}$.

Proof. We start by proving the statement from right to left: Assume (λ, x) solves $CRx = \lambda BRx$, then (λ, Rx) is an eigenpair of the pencil (C, B) as R is a full rank matrix. Since both B and C are Hermitian and B is positive definite, λ must be real-valued. Define $z = [x^T \ 0^T]^T \in \mathbb{C}^{2n-1}$, then $\hat{C}z = \lambda \hat{B}z$ and thus λ is a real-valued eigenvalue of the pencil (\hat{C}, \hat{B}) .

The converse is also true: Assume (λ, z) solves $\hat{C}z = \lambda \hat{B}z$ with λ real-valued and $z = [x^T \ y^T]^T$ nonzero, then the following two equations must hold:

$$\begin{cases} R^H C y = \lambda R^H B y, \\ CRx + Qy = \lambda BRx. \end{cases}$$

Because $Q \succ 0$, it has a Cholesky factorization $Q = LL^H$. Left multiplying the first equation with x^H and defining $b = L^{-1}BRx$ and $c = L^{-1}CRx$ gives

$$(3.5) \quad \begin{cases} c^H L^H y = \lambda b^H L^H y, \\ L^H y = \lambda b - c. \end{cases}$$

Eliminating $L^H y$ and rearranging terms yields

$$(3.6) \quad \|b\|^2 \lambda^2 - 2\Re(b^H c) \lambda + \|c\|^2 = 0,$$

where $\Re(b^H c)$ is the real part of $b^H c$. This last equation is always quadratic in λ as b cannot be the zero vector due to the following reasoning: Assume $b = 0$ then from the definition of b , $x = R^\dagger B^{-1} L b = 0$ because $R \in \mathbb{C}^{n \times (n-1)}$ is of full rank, and $c = 0$ from equation (3.6). But now $y = 0$ because of equation (3.5) and thus $z = 0$ which

is not possible because this is a (nonzero) eigenvector of $\hat{C} - \lambda\hat{B}$, thus $b \neq 0$. Solving equation (3.6) for λ gives

$$\lambda = \frac{\Re(b^H c) \pm \sqrt{D}}{\|b\|^2},$$

where $D = \Re(b^H c)^2 - \|b\|^2 \|c\|^2$ is the discriminant. Using the Cauchy-Schwarz inequality, we obtain

$$(3.7) \quad \Re(b^H c)^2 \leq |b^H c|^2 \leq \|b\|^2 \|c\|^2,$$

and therefore $D \leq 0$ meaning the only way to get a real-valued eigenvalue of this problem is by setting the discriminant equal to zero. This is true if and only if $c = \alpha b$ with $\alpha \in \mathbb{R}$ [17, Theorem 5.1.4], in which case $\lambda = \alpha$. Using our definitions of b and c , we can replace $c = \lambda b$ with

$$L^{-1}CRx = \lambda L^{-1}BRx,$$

and thus (λ, x) solves $CRx = \lambda BRx$. \square

A direct consequence of this lemma is that if the rectangular pencil (CR, BR) has no solutions, all eigenvalues of (\hat{C}, \hat{B}) have a nonzero imaginary part. This fact is indirectly used in the following theorem.

THEOREM 3.7. *The operator determinant Δ_0 is singular if and only if there exists a value $\lambda \in \mathbb{C}$ and a nonzero vector $x \in \mathbb{C}^{n-1} \setminus \{0\}$ such that $CRx = \lambda BRx$.*

Proof. We start the proof from right to left: Assume $CRx = \lambda BRx$ is solved for some $\lambda \in \mathbb{C}$ and nonzero $x \in \mathbb{C}^{n-1} \setminus \{0\}$, then Lemma 3.6 tells us that λ must be real-valued and that $\hat{C}y = \lambda \hat{B}y$ for some nonzero vector $y \in \mathbb{C}^{2n-1} \setminus \{0\}$. Define the vector $z = Rx \otimes y$, then

$$\begin{aligned} \Delta_0 z &= BRx \otimes \hat{C}y - CRx \otimes \hat{B}y, \\ &= \lambda BRx \otimes \hat{B}y - \lambda BRx \otimes \hat{B}y = 0, \end{aligned}$$

and since z is nonzero, Δ_0 must be singular.

Conversely, if Δ_0 is singular, there exists a nonzero vector $z \in \mathbb{C}^{2n^2-n} \setminus \{0\}$ such that $\Delta_0 z = 0$. Define $Z \in \mathbb{C}^{(2n-1) \times n}$ such that $z = \text{vec}(Z)$, then $\Delta_0 z = 0$ can be written as a Sylvester equation in Z :

$$\hat{C}ZB^T - \hat{B}ZC^T = 0$$

Such an equation has a nonzero solution if and only if the pencils (C, B) and (\hat{C}, \hat{B}) have a common eigenvalue³ [11], that is, there exists a λ such that both $C - \lambda B$ and $\hat{C} - \lambda \hat{B}$ are rank deficient. This eigenvalue must be real-valued due to C being Hermitian and B positive definite, and consequently, there exists a nonzero vector $x \in \mathbb{C}^{n-1} \setminus \{0\}$ such that $CRx = \lambda BRx$ because of Lemma 3.6. \square

A problem occurs when $\text{rank}(C) < n - 1$, in which case CR is rank deficient meaning there must be some nonzero vector $x \in \mathbb{C}^{n-1}$ such that $CRx = 0$. More

³The correct statement says there is a unique solution if and only if the two pencils have disjoint spectra and they are both regular. However, in our case $C - \lambda B$ is regular since B is invertible, and $\hat{C} - \lambda \hat{B}$ is regular because otherwise Lemma 3.6 states that $C - \lambda B$ is singular which is impossible.

precisely, $(\lambda = 0, x)$ solves the equation $CRx = \lambda BRx$ and Δ_0 must be singular because of Theorem 3.7, so the assumption of Theorem 3.5 does not hold. Even worse, the vector $Rx \otimes [x^\top \ 0]^\top \in \mathbb{C}^{n(2n-1)}$ lives in the null space of both Δ_0 and Δ_1 , so the entire pencil (Δ_1, Δ_0) is singular. This case is further investigated in Subsection 4.2 as we are still able to find the NEPv eigenvalues using this linearization (3.4).

3.3. Spurious solutions. The MEP (3.1) and its corresponding GEP (3.3) have at most $2n^2 - n$ isolated solutions, which exceeds the upper bound of n^2 isolated solutions of the NEPv (2.1), so after solving the linearized problem we should expect some false eigenvalues to appear. Part of the solutions change if a different choice for R is taken in the linearization, and because of their dependence on R , they must be spurious. More specifically, consider the problem of finding the numbers $\lambda, \mu \in \mathbb{C}$ and the nonzero vector $x \in \mathbb{C}^{n-1} \setminus \{0\}$ such that

$$(3.8) \quad M(\lambda, \mu)Rx = 0,$$

i.e., we are looking for a scalar μ such that there is an eigenvalue of $(A - \mu C, B)$ with an eigenvector in the column space of R . If we use the fact that $M = A - \lambda B - \mu C$, it becomes clear that we are essentially trying to solve the rectangular multiparameter eigenvalue problem (rMEP)

$$ARx = \lambda BRx + \mu CRx.$$

In the generic case, this problem has exactly $\ell = \frac{1}{2}n(n-1)$ eigenvalues counting multiplicities [15, 29, Lemma 1]. Note that because (λ, μ, x) solves the rMEP, the equation $MRx = 0$ holds, which also means that

$$\begin{cases} M(Rx) = 0, \\ \begin{bmatrix} 0 & R^H M \\ MR & S \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 0. \end{cases}$$

The vector Rx cannot be zero since R is of full rank and thus $(\lambda, \mu, Rx, [x^\top \ 0]^\top)$ solves our linearization (3.1).

A similar reasoning can be said about the rMEP,

$$y^H R^H M(\lambda, \mu) = 0$$

which also has exactly ℓ solutions (λ, μ, y) counting multiplicities in the generic case. Indeed,

$$\begin{cases} (Ry)^H M = 0 \\ \begin{bmatrix} y^H & 0 \end{bmatrix} \begin{bmatrix} 0 & R^H M \\ MR & S \end{bmatrix} = 0 \end{cases}$$

and therefore (λ, μ) is also an eigenvalue of (3.1).

The two rMEPs discussed above explain the appearance of spurious eigenvalues that depend on R when solving $\Delta_1 z = \lambda \Delta_0 z$ and in Subsection 4.1 we use a property about their right and left eigenvectors in order to filter them out. However, these eigenvalues may not account for all the spurious solutions. Recall that the upper bound n^2 on the number of isolated solutions of (2.1) is based on the number of solutions of the polynomial equations (2.4). It turns out that all solutions of these equations that do not correspond to eigenvalues of the NEPv (2.1), such as the complex conjugate pairs, are also spurious solutions, which is a consequence of Proposition 3.8.

PROPOSITION 3.8. *If (λ, μ) solves the polynomial equations (2.4), then there exist nonzero vectors $v \in \mathbb{C}^n$ and $w \in \mathbb{C}^{2n-1}$ such that (λ, μ, v, w) solves the MEP (3.1).*

Proof. Given (λ, μ) such that (2.4) holds, then $k = \dim(\text{null}(M)) \geq 1$ since $\det(M) = 0$. We distinguish two cases:

- If $k = 1$, there exist nonzero vectors $u, v \in \mathbb{C}^n \setminus \{0\}$ such that $Mv = 0$, $u^H M = 0$ and $\text{adj}(M) = vu^H$ [17, page 22]. Using (2.4) we then have $0 = \text{trace}(S \text{adj}(M)) = \text{trace}(Svu^H) = u^H Sv$.
- In case $k > 1$, let the matrices $X \in \mathbb{C}^{n \times k}$ and $Y \in \mathbb{C}^{n \times k}$ be such that their columns form a basis for the left and right null space of M , respectively, and define $\hat{S} = X^H SY \in \mathbb{C}^{k \times k}$. Choose any nonzero $x \in \mathbb{C}^k \setminus \{0\}$ and $y \in \mathbb{C}^k \setminus \{0\}$ such that $x^H \hat{S} y = 0$ which is always possible since $k > 1$, then $Mv = 0$, $u^H M = 0$ and $u^H Sv = 0$ with $u = Xx$ and $v = Yy$.

From the equalities $u^H M = 0$ and $u^H Sv = 0$ we have that u lies in the left null space of the square matrix $\begin{bmatrix} MR & Sv \end{bmatrix}$. Let $\begin{bmatrix} w_1^T & \alpha \end{bmatrix}^T \in \mathbb{C}^n$ be a nonzero vector in its right null space, then it is easy to check that $(\lambda, \mu, v, \begin{bmatrix} w_1^T & \alpha v^T \end{bmatrix}^T)$ solves (3.1). \square

4. Numerical methods. Before proposing numerical methods to solve the resulting linearized GEP, we must know whether our pencil (Δ_1, Δ_0) is singular. In case it is nonsingular, we may rely on well developed algorithms in the literature such as the Arnoldi algorithm which will be adapted in two ways such that it filters a large portion of the spurious solutions in Subsection 4.1. In case the pencil is singular, the problem becomes more difficult to solve. Subsection 4.2 studies the case where $\text{rank}(C) < n - 1$ as it induces such a singular pencil. For these problems, the relation between the MEP and the system of GEPs is less understood [21], but the eigenvalues can still be retrieved using projection, augmentation or rank-completing perturbations [16, 23] and for this specific problem we can show how it may be solved using the methods from Subsection 4.1 as well.

4.1. Nonsingular pencil. To obtain the eigenvalues of the NEPv (2.1), it suffices to select the real-valued solutions of the generalized eigenvalue problem

$$\Delta_1 z = \lambda \Delta_0 z.$$

Finding all $2n^2 - n$ eigenvalues is possible, for instance by using the QZ algorithm, but this is applicable to small-scale problems only, and usually we are interested in one or a couple of eigenvalues close to a shift for which faster algorithms are available. This section focuses on Krylov methods, specifically the Arnoldi algorithm which will also search for the undesirable spurious solutions making it expensive for large problems. To elevate this problem, we choose a starting vector for the algorithm such that all iterates are restricted to an invariant subspace which excludes a portion of the spurious eigenvalues.

4.1.1. Filtering Arnoldi method. Recall that the Arnoldi algorithm [1] builds and orthogonalizes the Krylov space

$$\mathcal{K}_k = \{z_0, (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 z_0, \dots, ((\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0)^{k-1} z_0\},$$

where z_0 is the initial starting vector and σ a shift. The orthonormal basis is given by $Z_k \in \mathbb{C}^{(2n^2-n) \times k}$ such that $\mathcal{K}_k = \text{span}(Z_k)$. The Ritz values are the eigenvalues of the projected problem in Hessenberg form

$$(4.1) \quad H_k = Z_k^H (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 Z_k.$$

This method can be improved by observing the following fact: Theorem 3.1 and Theorem 3.5 show that an eigenvalue of the NEP_v (2.2) corresponds to an eigenvector $z = v \otimes [w_1^\top \quad \alpha v^\top]^\top$ of the linear GEP (3.3). This vector has a special structure and is an element of the vector space

$$(4.2) \quad \mathcal{Z} = \left\{ z \in \mathbb{C}^{(2n^2-n)} \mid z = \text{vec} \left(\begin{bmatrix} W \\ V \end{bmatrix} \right), V = V^\top \in \mathbb{C}^{n \times n}, W \in \mathbb{C}^{(n-1) \times n} \right\}$$

of dimension $n^2 + \ell$ where $\ell = \frac{1}{2}n(n-1)$. If the Arnoldi algorithm is applied by choosing the starting vector $z_0 \in \mathcal{Z}$, then ℓ of the spurious eigenvalues can be filtered out. We prove this in Theorem 4.2 using Lemma 4.1 and the two technical lemmas in Section A.

LEMMA 4.1. *Assume the matrices Δ_0 , $\Delta_1 - \sigma\Delta_0$ and $\mathbf{A} - \sigma\mathbf{B} \in \mathbb{C}^{(\frac{n^2-n}{2}) \times (\frac{n^2-n}{2})}$ are all nonsingular with*

$$\begin{cases} \mathbf{A} = L^\top (R^H \otimes R^H) (A \otimes C - C \otimes A) T, \\ \mathbf{B} = L^\top (R^H \otimes R^H) (B \otimes C - C \otimes B) T, \end{cases}$$

where $L \in \mathbb{C}^{(n-1)^2 \times (\frac{n^2-n}{2})}$ and $T \in \mathbb{C}^{n^2 \times (\frac{n^2-n}{2})}$ are full rank matrices defined as

$$\begin{cases} \text{col}(L) = \{e_j^{(n-1)} \otimes e_i^{(n-1)} \mid 1 \leq i \leq j \leq n-1\}, \\ \text{col}(T) = \{e_j^{(n)} \otimes e_i^{(n)} - e_i^{(n)} \otimes e_j^{(n)} \mid 1 \leq i < j \leq n\}, \end{cases}$$

and $e_i^{(m)} \in \mathbb{R}^m$ is the i -th standard unit vector. If $z_k \in \mathcal{Z}$ and $z_{k+1} = (\Delta_1 - \sigma\Delta_0)^{-1} \Delta_0 z_k$, then $z_{k+1} \in \mathcal{Z}$.

Proof. First, we rewrite the equation $(\Delta_1 - \sigma\Delta_0)z_{k+1} = \Delta_0 z_k$ as

$$(4.3) \quad \hat{C}Z_{k+1}\Gamma^\top - \hat{\Gamma}Z_{k+1}C^\top = \hat{C}Z_kB^\top - \hat{B}Z_kC^\top,$$

where $\Gamma = A - \sigma B$, $\hat{\Gamma} = \hat{A} - \sigma\hat{B}$, $z_k = \text{vec}(Z_k)$ and $z_{k+1} = \text{vec}(Z_{k+1})$. Let

$$Z_k = \begin{bmatrix} W_k \\ V_k \end{bmatrix} \quad \text{and} \quad Z_{k+1} = \begin{bmatrix} W_{k+1} \\ V_{k+1} \end{bmatrix},$$

then the first $n-1$ rows of equation (4.3) read

$$R^H(CV_{k+1}\Gamma^\top - \Gamma V_{k+1}C^\top) = R^H(CV_kB^\top - BV_kC^\top).$$

Right multiplying with \bar{R} and defining $F = R^H(CV_kB^\top - BV_kC^\top)\bar{R}$ yields

$$R^H(CV_{k+1}\Gamma^\top - \Gamma V_{k+1}C^\top)\bar{R} = F.$$

Since $z_k \in \mathcal{Z}$, V_k is a symmetric matrix and as a result F is skew symmetric, i.e., the symmetric matrix $F + F^\top$ equals zero and therefore

$$R^H(C(V_{k+1} - V_{k+1}^\top)\Gamma^\top - \Gamma(V_{k+1} - V_{k+1}^\top)C^\top)\bar{R} = F + F^\top = 0.$$

This is a homogeneous system of linear equations in $X = V_{k+1} - V_{k+1}^\top$ that may equivalently be written in terms of Kronecker products as

$$(R^H \otimes R^H)(\Gamma \otimes C - C \otimes \Gamma) \text{vec}(X) = 0.$$

Due to the symmetry, the equation has redundant rows that can be eliminated by left multiplying with L^\top . Moreover, X is skew symmetric and can be written as $\text{vec}(X) = Tx$ where $x \in \mathbb{C}^{n(n-1)/2}$ contains the strictly upper triangular elements of X . Combining these reductions gives

$$L^\top(R^H \otimes R^H)(\Gamma \otimes C - C \otimes \Gamma)Tx = 0,$$

and therefore

$$(\mathbf{A} - \sigma \mathbf{B})x = 0.$$

We made the assumption that this matrix is invertible, that is, the unique solution is the zero vector $x = 0$ which implies $X = V_{k+1} - V_{k+1}^\top = 0$. By consequence, V_{k+1} is symmetric and $z_{k+1} \in \mathcal{Z}$. \square

THEOREM 4.2. *Assume Δ_0 is nonsingular and that $\sigma \in \mathbb{C}$ is not an eigenvalue of the pencil $\Delta_1 - \lambda \Delta_0$. Let $z_k \in \mathcal{Z}$ and $z_{k+1} = (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 z_k$. If either*

- (1) *the square pencil (\mathbf{A}, \mathbf{B}) of size $\frac{1}{2}n(n-1)$ is nonsingular with \mathbf{A} and \mathbf{B} defined in Lemma 4.1, or*
- (2) *the system of GEPs (3.3) has $2n^2 - n$ distinct eigentuples (λ, μ) ,*

then $z_{k+1} \in \mathcal{Z}$.

Proof. We adopt two complementary approaches. Our first approach is based on directly analyzing the matrix vector product for z_{k+1} . Assume that condition (1) holds and choose a σ_0 , not necessarily equal to the shift σ , such that both $\Delta_1 - \sigma_0 \Delta_0$ and $\mathbf{A} - \sigma_0 \mathbf{B}$ are nonsingular, then from Lemma 4.1 the subspace \mathcal{Z} is invariant under multiplication with the matrix $\Delta = (\Delta_1 - \sigma_0 \Delta_0)^{-1} \Delta_0$. Note that

$$(\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 = (I - (\sigma - \sigma_0) \Delta)^{-1} \Delta = p(\Delta)$$

for some polynomial p of degree less than or equal to $2n^2 - n$ because of the Cayley-Hamilton theorem [17, Theorem 2.4.3.2, Corollary 2.4.3.4], and as a result $z_{k+1} = (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 z_k = p(\Delta) z_k \in \mathcal{Z}$.

The second approach is inspired by [7, Theorem 4.7] and relies on the property that the invariant vector space \mathcal{Z} is spanned by a subset of the eigenvectors of the pencil (Δ_1, Δ_0) . More precisely, as shown in Lemma A.2, it is stated that if condition (2) holds, then $z_k \in \mathcal{Z}$ can be decomposed as

$$z_k = \sum_{i=1}^{n^2+\ell} \alpha_i \hat{z}_i,$$

where \hat{z}_i is an eigenvector of (3.3) that lies in \mathcal{Z} and $\ell = \frac{1}{2}n(n-1)$. Now z_{k+1} equals,

$$z_{k+1} = (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 z_k = \sum_{i=1}^{n^2+\ell} \frac{\alpha_i}{\lambda_i - \sigma} \hat{z}_i,$$

and thus $z_{k+1} \in \mathcal{Z}$. \square

The consequence of this last statement is that when we choose the initial vector z_0 in this invariant subspace \mathcal{Z} , then the next iteration will lie in this space as well. As a result, the ℓ spurious solutions corresponding to the problem $y^H R^H M(\lambda, \mu) = 0$ will be filtered. This leads to Algorithm 4.1: Line 5 solves a large linear system

Algorithm 4.1 Arnoldi method**Require:** $z_0 \in \mathcal{Z}$, σ , k_{max}

```

1:  $z_0 \leftarrow z_0 / \|z_0\|_2$ 
2:  $Z_1 = [z_0]$ 
3:  $H_0 = []$ 
4: for  $k = 1, 2, \dots, k_{max}$  do
5:    $\hat{z}_k = (\Delta_1 - \sigma \Delta_0)^{-1} \Delta_0 z_{k-1}$            (Solve using Sylvester equation)
6:    $h_k = Z_{k-1}^H \hat{z}_k$ 
7:    $\hat{z}_k \leftarrow \hat{z}_k - Z_{k-1} h_k$ 
8:    $\beta_k = \|\hat{z}_k\|_2$ 
9:    $\hat{z}_k \leftarrow \hat{z}_k / \beta_k$ 
10:   $\begin{bmatrix} W_k \\ \hat{V}_k \end{bmatrix} = \text{vec}^{-1}(\hat{z}_k); \quad V_k = \frac{1}{2}(\hat{V}_k + \hat{V}_k^T) \in \mathbb{C}^{n \times n}; \quad z_k = \text{vec} \left( \begin{bmatrix} W_k \\ V_k \end{bmatrix} \right)$ 
11:   $Z_{k+1} = [Z_{k-1} \quad z_k]$ 
12:   $H_k = \begin{bmatrix} H_{k-1} & h_k \\ 0 & \beta_k \end{bmatrix}$ 
13: end for
14: return  $H_{k_{max}}, Z_{k_{max}}$ 

```

of dimension $2n^2 - n$ which is the most expensive step. Because of the structure in the matrices Δ_i , this cost can be significantly reduced by solving the following corresponding generalized Sylvester equation instead:

$$\hat{C}X_k(A - \sigma B)^T - (\hat{A} - \sigma \hat{B})X_k C^T = E$$

with $\text{vec}(E) = \Delta_0 z_{k-1}$ and where the solution is $z_k = \text{vec}(X_k)$. If this linear matrix equation is solved using a Bartels-Stewart like method [11, 30], for instance, then the time complexity can be reduced from $\mathcal{O}(n^6)$ to $\mathcal{O}(n^3)$. Lines 6 to 9 in Algorithm 4.1 correspond to the orthogonalization step which can be implemented using modified Gram-Schmidt with reorthogonalization to minimize rounding errors. In line 10, the new vector z_k is projected onto \mathcal{Z} to correct for potential deviations caused by floating-point precision errors.

4.1.2. Two-sided projection. Instead of projecting the problem to a Hessenberg matrix (4.1), we also consider the alternative projection

$$\underbrace{Z_k^H \Delta_1 Z_k}_{H_1} y = \lambda \underbrace{Z_k^H \Delta_0 Z_k}_{H_0} y,$$

after building the same Krylov space using Algorithm 4.1, which leads to Algorithm 4.2. This new method does not differ much except now the Ritz values are obtained by solving the pencil $H_1 - \lambda H_0$. Both H_1 and H_0 are Hermitian like the original pencil, so the projection preserves this structure. Moreover, for real eigenvalues λ of $\Delta_1 - \lambda \Delta_0$ we expect faster convergence as the column space of Z_k lies closer to both its right and left eigenvector after each iteration.

Also, an improved filtering of spurious eigenvalues is observed in the numerical experiments, in the sense that twice as many false eigenvalues will be ignored in comparison with the previous filtering. We can explain this using the set of vectors

$$(4.4) \quad \mathcal{W} = \left\{ z \in \mathbb{C}^{2n^2-n} \mid z = \text{vec} \left(\begin{bmatrix} W \\ 0 \end{bmatrix} \right), RW = (RW)^T \right\} \subset \mathcal{Z}$$

and the following theorem.

THEOREM 4.3. *If $z \in \mathcal{W}$ (4.4) and $\text{col}(Z) \subseteq \mathcal{Z}$ (4.2), then $Z^H(\Delta_1 - \lambda\Delta_0)z = 0$ for any $\lambda \in \mathbb{C}$. Consequence: If such a vector $z \in \text{col}(Z)$, then (H_1, H_0) is a singular pencil where $H_j = Z^H\Delta_j Z$ for $j = 1, 2$.*

Proof. Define $\Gamma := A - \lambda B$, then

$$\Delta(\lambda) := \Delta_1 - \lambda\Delta_0 = \Gamma \otimes \begin{bmatrix} 0 & R^H C \\ CR & Q \end{bmatrix} - C \otimes \begin{bmatrix} 0 & R^H \Gamma \\ \Gamma R & P \end{bmatrix}.$$

The matrix vector product $\Delta(\lambda)z$ with $z \in \mathcal{W}$ can be expressed using the vec operator:

$$\Delta(\lambda)z = \text{vec} \left(\begin{bmatrix} 0 \\ \hat{\Gamma} \end{bmatrix} \right),$$

where

$$\hat{\Gamma} = CRW\Gamma^T - \Gamma RWC^T.$$

Since RW is symmetric, $\hat{\Gamma}$ is skew symmetric, i.e., $\hat{\Gamma}^T = -\hat{\Gamma}$. Let z_i be the i -th column of Z , then $z_i \in \mathcal{Z}$ because $\text{col}(Z) \subseteq \mathcal{Z}$. Therefore there exists a matrix $W_i \in \mathbb{C}^{(n-1) \times n}$ and a symmetric matrix $V_i = V_i^T \in \mathbb{C}^{n \times n}$ such that

$$z_i = \text{vec} \left(\begin{bmatrix} W_i \\ V_i \end{bmatrix} \right).$$

The inner product of two vectorized matrices can be expressed using the trace operator

$$\alpha_i := z_i^H \Delta(\lambda)z = \text{trace}(V_i^H \hat{\Gamma}).$$

Note that since V_i is symmetric, and $\hat{\Gamma}$ is skew symmetric, we have

$$\begin{aligned} \alpha_i &= -\text{trace}(V_i^H \hat{\Gamma}^T), & (\hat{\Gamma} &= -\hat{\Gamma}^T) \\ &= -\text{trace}(\hat{\Gamma} V_i^H), & (V_i &= V_i^T) \\ &= -\text{trace}(V_i^H \hat{\Gamma}) = -\alpha_i, & (\text{trace}(XY) &= \text{trace}(YX)) \end{aligned}$$

and thus $\alpha_i = 0$. This is true for any column of Z and therefore

$$Z^H \Delta(\lambda)z = Z^H(\Delta_1 - \lambda\Delta_0)z = 0$$

for any $\lambda \in \mathbb{C}$. □

A consequence of this theorem is that the eigenvalues of the pencil (Δ_1, Δ_0) that have a corresponding eigenvector z in the set $\mathcal{W} \subset \mathcal{Z}$ are no eigenvalues of the projected pencil (H_1, H_0) , because in that case $(H_1 - \sigma H_0)y = 0$ for any scalar σ , with $z = Zy$. Coming back to the characterization of spurious eigenvalues in Subsection 3.3, note that the solutions (λ, μ, x) of the rMEP $M(\lambda, \mu)Rx$ correspond to a solution (λ, z) of the pencil (Δ_1, Δ_0) with $z \in \mathcal{W}$, so these $\ell = \frac{1}{2}n(n-1)$ spurious solutions will be filtered by Algorithm 4.2. By symmetry, a similar reasoning reveals that the ℓ solutions (λ, μ, x) of the pencil $x^H R^H M(\lambda, \mu)$ are filtered as well and thus we expect 2ℓ spurious solutions to be ignored.

The downside of Algorithm 4.2 is the extra computation required for the projection and the fact that (H_1, H_0) gets closer to a singular pencil after every additional

iteration. The issue is that as Z_k grows in size, its column space gets closer to vectors in the set \mathcal{W} that make the pencil singular. In fact, from a certain iteration on, the column space of Z_k in Algorithm 4.2 will contain a vector $z \in \mathcal{W}$ such that the projected pencil is singular.

PROPOSITION 4.4. *Let $Z_k \in \mathbb{C}^{(2n^2-n) \times k}$ be generated by k iterations of Algorithm 4.1, $H_i = Z_k^H \Delta_j Z_k$ for $j = 0, 1$ and $\text{col}(Z_k) \subseteq \mathcal{Z}$. If $k > n^2$, then the pencil $H_1 - \lambda H_0$ is singular.*

Proof. Let z_i be the i -th column of Z_k , then

$$z_i = \text{vec} \left(\begin{bmatrix} W_i \\ V_i \end{bmatrix} \right),$$

with $V_i = W_i^T$ since $\text{col}(Z_k) \subseteq \mathcal{Z}$. From Theorem 4.3, the pencil is singular if there exists a vector $y = [y_1 \ \dots \ y_k]^T \in \mathbb{C}^k$ such that

$$Z_k y = \text{vec} \left(\begin{bmatrix} W \\ 0 \end{bmatrix} \right),$$

where RW is a symmetric matrix. This is possible if

$$\begin{cases} \sum_{i=1}^k \Lambda_i y_i = 0, \\ \sum_{i=1}^k V_i y_i = 0, \end{cases}$$

with $\Lambda_i := RW_i - W_i^T R^T$ a skew symmetric matrix. The first equation accounts to $\frac{1}{2}n(n-1)$ equations due to the skew symmetry of Λ_i , while the second one has $\frac{1}{2}n(n+1)$ equations because of the symmetric V_i . Hence, there are a total of n^2 linear equations in k unknowns which is guaranteed to have a nonzero solution if $k > n^2$. As a result, the pencil (H_1, H_0) has to be singular for these values of k because of Theorem 4.3. \square

Of course, k does not need to be greater than n^2 for the pencil (H_1, H_0) to become singular, as the system we have to solve in the proof of the last corollary could be rank deficient, for example if $Z \subseteq \mathcal{Z}$ contains a vector $z = \text{vec} \left(\begin{bmatrix} W^T & 0 \end{bmatrix}^T \right)$ with RW symmetric. It is unlikely for this to happen, but if our Krylov space Z gets closer to one of these vectors z , the projected eigenvalue problem (H_1, H_0) becomes close to singular rather quickly. We will illustrate this aspect in Section 5.

Algorithm 4.2 Two-sided projection method

Require: $z_0 \in \mathcal{Z}$, σ , k_{max}

- 1: Obtain $Z_{k_{max}}$ from Algorithm 4.1
 - 2: Compute $H_j = Z_{k_{max}}^H \Delta_j Z_{k_{max}}$ for $j = 0, 1$
 - 3: Solve for eigenpairs (θ_i, x_i) : $H_1 x_i = \theta_i H_0 x_i$
 - 4: **return** θ_i
-

4.2. Singular pencil. The methods developed in Subsection 4.1 assume that (Δ_1, Δ_0) is a regular pencil which holds under the conditions of Theorem 3.7. In case $\text{rank}(C) < n - 1$ these conditions are not satisfied and, as discussed in the last paragraph of Subsection 3.2, it can be shown that the whole pencil (Δ_1, Δ_0) becomes

singular. For such an eigenvalue problem, an eigenvalue λ is defined as the value for which the rank of $\Delta_1 - \lambda\Delta_0$ is smaller than the normal rank

$$\text{nrnk}(\Delta_1 - \nu\Delta_0) = \max_{\nu \in \mathbb{C}} \text{rank}(\Delta_1 - \nu\Delta_0).$$

Fortunately, the eigenvalues of this GEP corresponding to solutions of the NEPv (2.1) hold to this definition under the assumptions of Lemma 4.5 and Theorem 4.6.

LEMMA 4.5. *Let $\text{rank}(C) < n - 1$ and assume that the eigenvectors of the pencil (C, B) corresponding to nonzero eigenvalues do not lie in the column space of R . Let z be a nonzero vector such that $\Delta_0 z = 0$, then $(\Delta_1 - \sigma\Delta_0)z = 0$ for any $\sigma \in \mathbb{C}$.*

Proof. Let $z = \text{vec}(Z)$ be a nonzero vector such that $\Delta_0 z = 0$ or equivalently

$$(4.5) \quad \hat{C}ZB^\top - \hat{B}ZC^\top = 0,$$

and denote with $C^\top X = B^\top X \Theta$ an eigenvalue decomposition of (C^\top, B^\top) with $X \in \mathbb{C}^{n \times n}$ and $\Theta = \text{diag}(\theta_1, \dots, \theta_n) \in \mathbb{R}^{n \times n}$, then

$$\hat{C}\hat{Z} - \hat{B}\hat{Z}\Theta = 0,$$

where $\hat{Z} = ZB^\top X$. The nonzero columns \hat{z}_i of \hat{Z} are eigenvectors of the pencil (\hat{C}, \hat{B}) corresponding to real-valued eigenvalues θ_i , meaning there must exist nonzero $x_i \in \mathbb{C}^{n-1}$ such that $CRx_i = \theta_i BRx_i$ because of Lemma 3.6. From the assumption that no eigenvectors corresponding to nonzero eigenvalues of (C, B) lie in the column space of R , we must deduce that $\theta_i = 0$ or $\hat{z}_i = 0$. As a result, $\hat{C}\hat{Z} = 0$ and since B is nonsingular we have $\hat{C}Z = 0$. This implies the second term in (4.5) must be zero as well, that is, $\hat{B}ZC^\top = 0$. Suppose now that $ZC^\top \neq 0$, then $(\hat{C} - \theta\hat{B})ZC^\top = 0$ for any value $\theta \in \mathbb{C}$, which is not possible according to Lemma 3.6 as (C, B) has a finite number of eigenvalues and R is of full rank. Therefore

$$\hat{C}Z = 0 \quad \text{and} \quad ZC^\top = 0.$$

It is now easy to see that $\Delta_1 z$ must be zero as well and thus $(\Delta_1 - \sigma\Delta_0)z = 0$ for any shift σ . \square

A direct consequence of this Lemma is that the rank of $(\Delta_1 - \sigma\Delta_0)$ is smaller than or equal to the rank of $\text{rank}(\Delta_0)$, which implies that $\text{nrnk}(\Delta_1 - \nu\Delta_0) = \text{rank}(\Delta_0)$. Moreover, both matrices have a common null space which is an important fact that will make the Sylvester equations in Algorithm 4.1 solvable.

Also note that if the nonzero eigenvalues of the pencil (C, B) are simple, the assumption of Lemma 4.5 holds for almost all choices of R .

THEOREM 4.6. *Suppose $\text{rank}(C) < n - 1$ and that the assumption of Lemma 4.5 holds. If (λ, v) solves the NEPv (2.1) with $v \in \mathbb{C}^n$ such that Cv is nonzero, then $\text{rank}(\Delta_1 - \lambda\Delta_0) < \text{nrnk}(\Delta_1 - \sigma\Delta_0)$.*

Proof. Let (λ, v) be a solution of the NEPv (2.1) then from Lemma 2.1 and Theorem 3.1 there exists a scalar $\mu \in \mathbb{C}$ and a nonzero vector $w \in \mathbb{C}^{2n-1} \setminus \{0\}$ such that (λ, μ, v, w) solves the MEP (3.1). Note that even though Δ_0 is singular, the eigenpair (λ, z) where $z = v \otimes w$ still solves $\Delta_1 z = \lambda\Delta_0 z$:

$$\begin{aligned} (\Delta_1 - \lambda\Delta_0)z &= (A - \lambda B)v \otimes \hat{C}w - Cv \otimes (\hat{A} - \lambda\hat{B})w, \\ &= (\mu C)v \otimes \hat{C}w - Cv \otimes (\mu\hat{C})w = 0. \end{aligned}$$

Suppose this eigenvector also lies in the right null space of Δ_0 , then

$$\Delta_0 z = 0,$$

or written in terms of Kronecker products

$$Bv \otimes \hat{C}w - Cv \otimes \hat{B}w = 0.$$

Therefore, there must exist a value $\theta \in \mathbb{C}$ such that

$$\begin{cases} Cv = \theta Bv, \\ \hat{C}w = \theta \hat{B}w. \end{cases}$$

The first equation implies that θ is real-valued and thus Lemma 3.6 can be used once more with the second equation, i.e., there exists some $x \in \mathbb{C}^{n-1}$ such that $CRx = \theta BRx$. From the assumption of Lemma 4.5, θ can only be equal to zero which implies $Cv = 0$. This contradicts with the assumption that Cv is nonzero, hence

$$\Delta_0 z \neq 0.$$

Indeed, we have that z lies in the null space of $\Delta_1 - \lambda\Delta_0$ but not in the null space of Δ_0 whereas every vector \hat{z} in the null space of Δ_0 also satisfies $(\Delta_1 - \lambda\Delta_0)\hat{z} = 0$ according to Lemma 4.5, and therefore

$$\text{rank}(\Delta_1 - \lambda\Delta_0) < \text{rank}(\Delta_0) = \text{nrank}(\Delta_1 - \sigma\Delta_0). \quad \square$$

This theorem indicates that an eigenvalue of the NEPv (2.1) is also an eigenvalue of the singular pencil (Δ_1, Δ_0) under the two assumptions stated in *Theorem 4.6*, allowing us to use dedicated algorithms for solving singular matrix pencils [16, 23]. An observation is that the standard Arnoldi algorithm (Algorithm 4.1 without the projection onto \mathcal{Z} in line 10) can be used as well: The system

$$(4.6) \quad (\Delta_1 - \sigma\Delta_0)\hat{z}_k = \Delta_0 z_{k-1}$$

that needs to be solved in each iteration is singular, but if the shift σ is not an eigenvalue then both $\Delta_1 - \sigma\Delta_0$ and Δ_0 have the same column space as a result of Lemma 4.5 and the fact that both Δ_1 and Δ_0 are Hermitian⁴. In other words, the Sylvester equation equivalent to (4.6) has an infinite number of solutions of which we can pick one to continue the Arnoldi iterations. Unfortunately, the filtering proposed in Subsection 4.1 is not possible due to the fact that the singular pencil (Δ_1, Δ_0) will not have enough eigenvectors to form a basis for the invariant subspace \mathcal{Z} (4.2). Moreover, the upper bound on the number of solutions of the NEPv (2.1) drops from n^2 to $n(2r + 1) - r(r + 1)$ with $r = \text{rank}(C)$ (Theorem 4.7), which can be seen as a disadvantage as the difference between the number of NEPv solutions and the dimension of Δ_0 may become larger. A proof for Theorem 4.7 is given in Section B.

⁴If σ is not an eigenvalue then neither is the complex conjugate of σ because both Δ_0 and Δ_1 are Hermitian. The null spaces of $\Delta_1 - \bar{\sigma}\Delta_0$ and Δ_0 must be equal as a consequence of Lemma 4.5, and thus their row spaces must be equal as well. As a result, $\Delta_1 - \sigma\Delta_0$ and Δ_0 share the same column space.

THEOREM 4.7. *Let $r = \text{rank}(C) < n - 1$, then the number of isolated solutions $(\lambda, \mu) \in \mathbb{C} \setminus \{0\} \times \mathbb{C} \setminus \{0\}$ of the system of polynomial equations (2.4)*

$$\begin{cases} f(\lambda, \mu) = \det(A - \lambda B - \mu C) = 0, \\ g(\lambda, \mu) = \text{trace}((P - \mu Q) \text{adj}(A - \lambda B - \mu C)) = 0, \end{cases}$$

is bounded by $n(2r + 1) - r(r + 1)$.

5. Numerical examples. This section is concerned with solving the NEPv (2.1) for different matrices using the two algorithms from Section 4 to solve the exact linearization in Section 3. Recall that the linearization is,

$$\Delta_1 z = \lambda \Delta_0 z,$$

with,

$$\begin{aligned} \Delta_1 &= A \otimes \hat{C} - C \otimes \hat{A}, & \Delta_0 &= B \otimes \hat{C} - C \otimes \hat{B}, \\ \hat{A} &= \begin{bmatrix} 0 & R^H A \\ AR & P \end{bmatrix}, & \hat{B} &= \begin{bmatrix} 0 & R^H B \\ BR & 0 \end{bmatrix}, & \hat{C} &= \begin{bmatrix} 0 & R^H C \\ CR & Q \end{bmatrix} \end{aligned}$$

where the matrix $R \in \mathbb{C}^{n \times (n-1)}$ can be chosen at random, but of full rank. The NEPv matrices A, B, C, P and Q are symmetric matrices of size n and additionally B and Q are positive definite. The example code used to generate the figures is available at <https://gitlab.kuleuven.be/numa/public/nepv>.

5.1. Example 1. Consider the NEPv of dimension $n = 100$ where the matrices are dense and randomly generated as

```
rng(0)
A = randn(n) + 1i*randn(n); A = (A + A')/2;
B = randn(n) + 1i*randn(n); B = sqrtm(B*B');
C = randn(n) + 1i*randn(n); C = (C + C')/2;
P = randn(n) + 1i*randn(n); P = (P + P')/2;
Q = randn(n) + 1i*randn(n); Q = sqrtm(Q*Q');
```

in MATLAB. We choose random complex numbers as the entries for the matrix $R \in \mathbb{C}^{n \times (n-1)}$, and compute the matrices \hat{A} , \hat{B} and \hat{C} of dimension $2n - 1 = 199$, resulting in operator determinants Δ_i of size $2n^2 - n = 19900$. These delta matrices do not need to be computed explicitly since the linear system in Algorithm 4.1 is solved as a Sylvester equation using the MEP matrices. For this example we set σ equal to zero to search for the eigenvalues with smallest magnitude, and we take a random starting vector z_0 in \mathcal{Z} .

When the problem is solved using Algorithm 4.1, we get convergence to the eigenvalues as shown in the left figure of Figure 1 where the absolute error between the Ritz values of the current iteration and the reference eigenvalues are plotted in function of the iteration k . The reference eigenvalues are chosen as the Ritz values at the last iteration k_{max} , which explains the sudden vertical jump in the figures at $k = 100$. The blue convergence curves are of interest as they correspond to Ritz values that approximate an eigenvalue of the NEPv. Convergence to the first eigenvalue happens around iteration $k = 30$, while the second blue line has not converged yet in iteration $k = 100$. The right plot in Figure 1 compares the computed Ritz values in the last iteration with the eigenvalues of the NEPv.

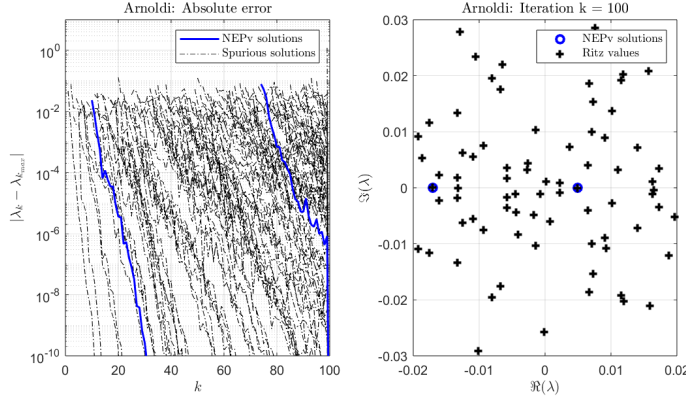


FIG. 1. *Left figure: Convergence plot for the example in Subsection 5.1. The blue lines represent convergence of Ritz values computed in Algorithm 4.1 towards eigenvalues of the NEPv, while the other curves represent convergence of Ritz values towards spurious solutions. Right figure: Comparison between the Ritz values and the NEPv eigenvalues at iteration $k = 100$.*

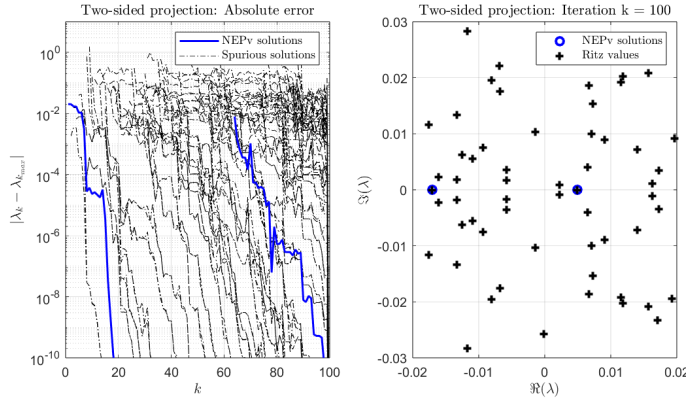


FIG. 2. *Similar to Figure 1 except now Algorithm 4.2 is used to compute the Ritz values.*

Alternatively, if Algorithm 4.2 is used where we also compute the two-sided projection to obtain the pencil $H_1 - \lambda H_0$, then convergence to the real-valued eigenvalues is reached within fewer iterations ($k = 18$ and $k = 97$) as shown in Figure 2. Comparing the right plots of Figure 1 and Figure 2, we see fewer Ritz values for this second method due to the improved filtering. The drawback of this approach is the extra computation needed to obtain the projected pencil, and the fact that this pencil becomes singular after a certain number of iterations. This is illustrated in Figure 3 by computing the smallest singular values of $H_0 + \rho H_1$ for some randomly chosen value ρ . Note that the smallest singular value drops to machine precision around $k = 12$, meaning the projected pencil is numerically rank deficient at this point.

5.2. Example 2. Consider the scalar function $u(x) \in \mathbb{R}$ in function of $x \in \mathbb{R}$ such that

$$-\nabla^2 u(x) + f(u)c(x)u(x) = \lambda u(x).$$

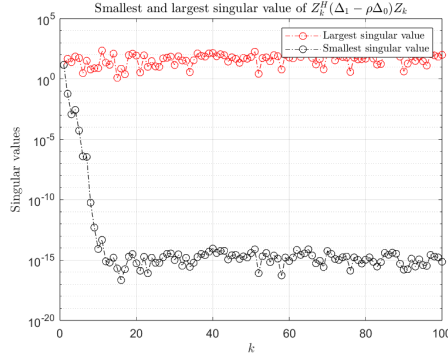


FIG. 3. Illustrates after which iteration the projected pencil $H_1 - \lambda H_0$ becomes numerically rank deficient for the example in Subsection 5.1.

In this example we solve a large sparse system that arises by discretizing this wave equation using finite differences on the domain $[-L/2, L/2]$ with Dirichlet boundary conditions $u(-L/2) = u(L/2) = 0$. The goal is to find the smallest eigenvalues λ and their corresponding wave functions $u(x)$ such that $\int_{-L/2}^{L/2} |u(x)|^2 dx = 1$. This equation is similar to the problem in [7], except now the nonlinearity is defined as

$$f(u) = \int_{-L/2}^{L/2} p(x) |\nabla u(x)|^2 dx, \quad \text{with} \quad p(x) = 5 \cos\left(\frac{\pi}{L}x\right),$$

and we take

$$c(x) = 1 - \exp\left(-\frac{(10x-1)^2}{10}\right).$$

The equation is discretized using central differences and an equidistant mesh of $n+2$ nodes $\{x_k\}_{k=0}^{n+1}$ with $x_0 = -L/2$ and $x_{n+1} = L/2$, where the distance between two nodes is $h = L/(n+1)$. The discrete wavefunction is stored in the vector $v \in \mathbb{C}^n$ such that v_i corresponds to $u(x_i)$ for $i = 1, \dots, n$. The resulting discretized problem is

$$Av = \lambda v + \frac{v^H P v}{v^H v} C v,$$

with banded matrices computed as

$$\begin{aligned} A_{i,i} &= 2/h^2, \quad A_{i,i+1} = A_{i+1,i} = -1/h^2, \\ C &= -\text{diag}([c(x_1) \quad c(x_2) \quad \dots \quad c(x_n)]), \\ P_{i,i} &= \frac{p(x_{i-1}) + p(x_{i+1})}{4h^2}, \quad P_{i,i+2} = P_{i+2,i} = \frac{-p(x_{i+1})}{4h^2}. \end{aligned}$$

For this example, we choose the values $L = 2$ and $n = 256$ and use the shift $\sigma = 50$. The matrix R is chosen as a horizontal stack of the identity matrix and a row vector with random complex elements. After 150 iterations of both Algorithm 4.1 and Algorithm 4.2, convergence has been reached to 4 and 5 eigenvalues, respectively. The approximated absolute errors of the computed eigenvalues are plotted in Figure 4 and Figure 5. In this case, Algorithm 4.2 has converged to the smallest eigenvalue

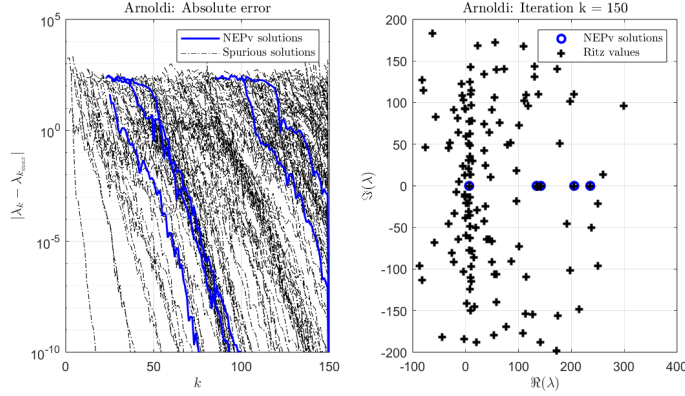


FIG. 4. *Left figure: Convergence plot for the example in Subsection 5.2. The blue lines represent convergence of Ritz values computed in Algorithm 4.1 towards eigenvalues of the NEPv, while the other curves represent convergence of Ritz values towards spurious solutions. Right figure: Comparison between the Ritz values and the NEPv eigenvalues at iteration $k = 150$.*

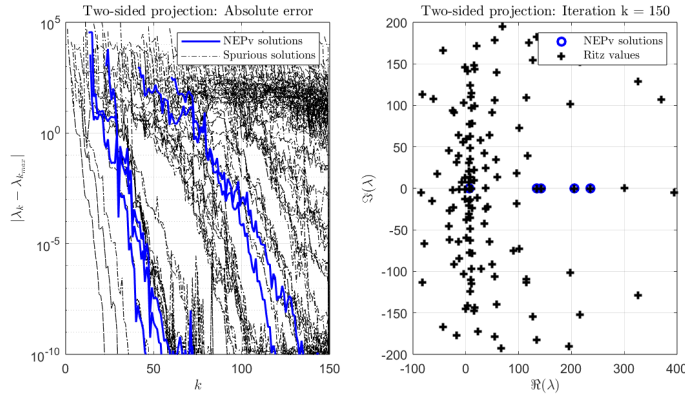


FIG. 5. *Similar to Figure 4 except now Algorithm 4.2 is used to compute the Ritz values.*

$\lambda = 6.67$ in 45 iterations, showing its faster convergence rate to the NEPv eigenvalues in comparison with Algorithm 4.1 which converges to the same eigenvalue after 75 iterations. The entries of the eigenvectors corresponding to the 5 computed eigenvalues are the discretization of the eigenfunctions $u(x)$ and are plotted in Figure 6.

5.3. Example 3. Consider the NEPv (2.1) with a rank r matrix C generated in MATLAB as

```
C1 = orth(randn(n, r) + 1i*randn(n, r));
C = C1*diag(randn(r, 1))*C1';
```

and where the other matrices are generated as in Subsection 5.1. For this example we take $n = 5$, $r = 2$, a shift $\sigma = 0$ and a complex randomly generated matrix R . To retrieve the eigenvalues, we solve the pencil (Δ_1, Δ_0) using the standard Arnoldi method, i.e., Algorithm 4.1 without the projection step onto the vector space \mathcal{Z} in line 10. The singular system in line 5 is solved using the Bartels-Steward method [11], in which the smaller singular systems are solved by using an LU factorization and

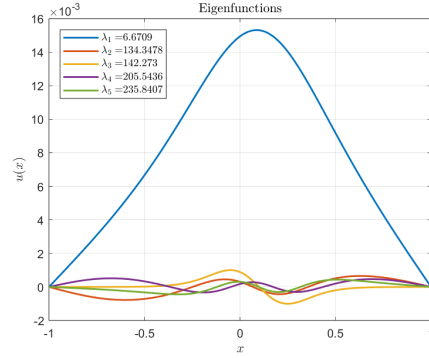


FIG. 6. Eigenfunctions $u(x)$, scaled with its eigenvalue, corresponding to the 5 computed eigenvalues in function of x for the example in Subsection 5.2.

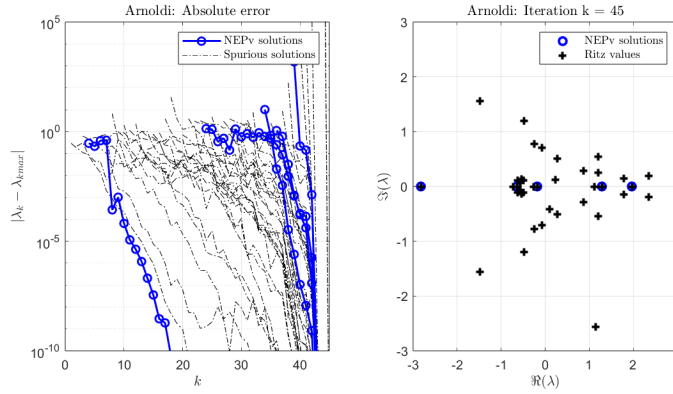


FIG. 7. Left figure: Convergence plot for the example in Subsection 5.3. The blue lines represent convergence of Ritz values computed using standard Arnoldi towards eigenvalues of the NEPv, while the other curves represent convergence of Ritz values towards spurious solutions. Right figure: Comparison between the Ritz values and the NEPv eigenvalues at iteration $k = 42$.

discarding the zero rows. The convergence behavior for all Ritz values is plotted in Figure 7. Convergence to the smallest eigenvalue is reached around iteration $k = 18$.

6. Conclusion and outlook. The NEPv considered in this paper contains one scalar nonlinearity in which the eigenvector appears quadratically. We show how the solutions of this problem are related to the real-valued solutions of a system of polynomial equations, and that the upper bound on the number of isolated solutions is n^2 where n is the dimension of the NEPv. The first main contribution of this paper is the conversion to a GEP in the sense that the spectrum of the resulting linear problem contains all the eigenvalues of the NEPv. We stated under which conditions the resulting GEP may become singular. As a second contribution, we show how the Arnoldi method can be used to efficiently obtain the eigenvalues of the NEPv by exploiting the structure of the resulting linear problem and by filtering a portion of the spurious solutions. Using the alternative two-sided projection results in an improved filtering of the spurious solutions at the cost of explicitly having to compute the projection. In case the linear pencil is singular because of a low rank matrix C , the methods can still be used since system (4.6) is solvable whenever the shift is not an

eigenvalue, however the filtering approach is not applicable anymore. The theoretical results are supported by numerical examples.

The paper only discussed the NEPv with one scalar nonlinearity, but this can be generalized to multiple terms in future research. A problem that will occur however is the exponential increase in the size of the resulting linear problem as the number of scalar nonlinearities is increased.

Appendix A. Technical Lemmas.

LEMMA A.1. *Assume Δ_0 is nonsingular and that (3.3) has $2n^2 - n$ distinct eigentuples $(\lambda, \mu) \in \mathbb{C} \times \mathbb{C}$. Let (λ, μ) and $z \in \mathbb{C}^{2n^2 - n}$ be such an eigentuple with its corresponding eigenvector, then at least one of the following statements must be true:*

- *The eigenvector z is an element of \mathcal{Z} .*
- *There exists a nonzero vector $y \in \mathbb{C}^{n-1}$ such that $y^H R^H M(\lambda, \mu) = 0$.*

Proof. Because Δ_0 is nonsingular, Theorem 3.5 says there must exist a v and w such that (λ, μ, v, w) solves the MEP (3.1) with $z = v \otimes w$, i.e.

$$(A.1) \quad \begin{cases} Mv = 0, \\ R^H M w_2 = 0, \\ MRw_1 + Sw_2 = 0, \end{cases}$$

where $w = [w_1^T \ w_2^T]^T$. Note that if $\text{rank}(M) < n - 1$, the eigenpair has a geometric multiplicity greater than one which contradicts with the assumption that all of them are distinct, thus $\text{rank}(M) = n - 1$. We consider the following two cases:

- **Case I:** $\text{rank}(R^H M) < n - 1$. The matrix $R^H M$ is rank deficient and there must exist a nonzero vector y in its left null space, i.e., $y^H R^H M = 0$.
- **Case II:** $\text{rank}(R^H M) = n - 1$. In this case, M and $R^H M$ share the same one-dimensional (right) null space and from (A.1) both v and w_2 lie in this space, therefore there must exist an $\alpha \in \mathbb{C}$ such that $w_2 = \alpha v$. Now we have

$$z = v \otimes w = v \otimes \begin{bmatrix} w_1 \\ \alpha v \end{bmatrix} = \text{vec} \left(\begin{bmatrix} w_1 v^T \\ \alpha v v^T \end{bmatrix} \right) \in \mathcal{Z}. \quad \square$$

LEMMA A.2. *Suppose Δ_0 is nonsingular and that (3.3) has $2n^2 - n$ distinct eigentuples $(\lambda, \mu) \in \mathbb{C} \times \mathbb{C}$. Denote $\{(\lambda_i, \mu_i, z_i)\}_{i=1}^N$ as the set of eigenpairs and corresponding eigenvectors of (3.3) for which $z_i \in \mathcal{Z}$, then this set is a basis for \mathcal{Z} .*

Proof. The system of GEPs (3.3) has distinct eigentuples, so the set of N eigenvectors $\{z_i\}_{i=1}^N$ must be linearly independent and $N \leq \dim(\mathcal{Z})$. Denote the set $\{(\hat{\lambda}_i, \hat{\mu}_i, \hat{z}_i)\}_{i=1}^{2n^2 - n - N}$ as the other solutions of (3.3) for which the eigenvector \hat{z}_i does not lie in \mathcal{Z} , then according to Lemma A.1 there must exist a set of nonzero vectors $\{y_i\}_{i=1}^{2n^2 - n - N}$ such that $(\hat{\lambda}_i, \hat{\mu}_i, y_i)$ solves the rMEP $y_i^H R^H M(\hat{\lambda}_i, \hat{\mu}_i) = 0$. From Theorem 3.7 and the assumption that Δ_0 is nonsingular, we have that $\text{rank}(CR - \sigma BR) = n - 1$ for any $\sigma \in \mathbb{C}$ and if this fact is combined with [15, Lemma 1], we must deduce that the number of solutions of this rMEP precisely equals $\ell = \frac{1}{2}n(n - 1)$ counting multiplicities, hence

$$\begin{aligned} 2n^2 - n - N &\leq \ell, \\ \iff N &\geq n^2 + \ell = \dim(\mathcal{Z}). \end{aligned}$$

Consequently, N must be equal to $\dim(\mathcal{Z})$ and therefore the set $\{z_i\}_{i=1}^N$ forms a basis for \mathcal{Z} . \square

Appendix B. Proof of Theorem 4.7.

The proof consists of three parts: In the first two parts the Newton polytopes of both polynomials are determined. In the third part, the Bernstein-Khovanskii-Kushnirenko (BKK) theorem [9, Theorem 5.4, Page 346] is used to obtain an upper bound on the number of isolated solutions.

Define $\Gamma = A - \lambda B$ and let $C = GDG^H$ be the reduced eigenvalue decomposition such that $D \in \mathbb{C}^{r \times r}$ contains the r nonzero eigenvalues of C , and $G \in \mathbb{C}^{n \times r}$ the corresponding eigenvectors. Assume Γ is invertible, then from the matrix determinant lemma (see also [17, Equation (0.8.5.1)]) we have,

$$f(\lambda, \mu) = \det(D) \det(\Gamma) \det(D^{-1} - \mu G^H \Gamma^{-1} G).$$

The variable μ only appears in the last determinant of this expression and since $G^H \Gamma^{-1} G$ is a matrix function of λ of size r , $f(\lambda, \mu)$ can be seen as a function of degree r in μ with coefficients in function of λ . Thus if $\lambda^i \mu^j$ is a monomial of f , then $j \leq r$. From Theorem 2.4, f has a maximal degree of n and therefore $i + j \leq n$. As a result, the Newton polytope of $f(\lambda, \mu)$ is $P_f = \text{Conv}((0, 0), (n, 0), (n - r, r), (0, r))$. This remains true in case Γ is singular since this occurs for a finite number of values λ and $f(\lambda, \mu)$ is continuous.

The second function $g(\lambda, \mu)$ can be seen as a function of degree $n - 1$ in λ with coefficients in function of μ . Alternatively, using the Woodbury matrix identity and the matrix determinant lemma, the adjugate in the expression of g can be written as

$$\text{adj}(\Gamma - \mu C) = f(\lambda, \mu) \Gamma^{-1} + \mu \det(D) \det(\Gamma) \Gamma^{-1} G \text{adj}(D^{-1} - \mu G^H \Gamma^{-1} G) G^H \Gamma^{-1},$$

assuming both Γ and $M = \Gamma - \mu C$ are nonsingular, but because this expression is continuous in μ it also holds in case M is singular. The expression can be seen as a matrix polynomial in μ of degree r with coefficients in function of λ , meaning $g(\lambda, \mu)$ has a degree of $r + 1$ in μ . Additionally, from Theorem 2.4, this polynomial has degree n , thus if $\lambda^i \mu^j$ is a monomial in $g(\lambda, \mu)$, then $i \leq n - 1$, $j \leq r + 1$ and $i + j \leq n$. The corresponding Newton polytope is $P_g = \text{Conv}((0, 0), (n - 1, 0), (n - 1, 1), (n - r - 1, r + 1), (0, r + 1))$. This also remains true in case Γ is singular due to the continuity of $g(\lambda, \mu)$.

The BKK theorem [9, Theorem 5.4, Page 346] states that the number of isolated solutions of the system of polynomial equations $f(\lambda, \mu) = g(\lambda, \mu) = 0$ is bounded by the mixed volume $MV_2(P_f, P_g)$, where P_f and P_g are the Newton polytopes of f and g , respectively. This mixed volume is the coefficient of the monomial $\lambda_f \lambda_g$ in

$$\begin{aligned} \text{Vol}_2(\lambda_f P_f + \lambda_g P_g) &= \left(rn - \frac{r^2}{2} \right) \lambda_f^2 + \left((r + 1)(n - 1) - \frac{r^2}{2} \right) \lambda_g^2 \\ &\quad + (n + r(2n - r - 1)) \lambda_f \lambda_g, \end{aligned}$$

and therefore the number of isolated solutions is bounded by $n + r(2n - r - 1)$.

Acknowledgments. The authors would like to thank Tom Kaiser for his expertise and insightful discussions, which contributed to the proof of Theorem 4.7.

REFERENCES

- [1] W. E. ARNOLDI, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quarterly of applied mathematics, 9 (1951), pp. 17–29.

- [2] F. V. ATKINSON, *Multiparameter eigenvalue problems. Volume 1: Matrices and compact operators*, Academic Press, New York, 1972.
- [3] Z. BAI, R.-C. LI, AND D. LU, *Sharp estimation of convergence rate for self-consistent field iteration to solve eigenvector-dependent nonlinear eigenvalue problems*, SIAM Journal on Matrix Analysis and Applications, 43 (2022), pp. 301–327.
- [4] W. BAO, H. WANG, AND P. A. MARKOWICH, *Ground, symmetric and central vortex states in rotating Bose-Einstein condensates*, Communications in Mathematical Sciences, 3 (2005), pp. 57–88.
- [5] T. BÜHLER AND M. HEIN, *Spectral clustering based on the graph p -Laplacian*, in Proceedings of the 26th annual international conference on machine learning, 2009, pp. 81–88.
- [6] Y. CAI, L.-H. ZHANG, Z. BAI, AND R.-C. LI, *On an eigenvector-dependent nonlinear eigenvalue problem*, SIAM Journal on Matrix Analysis and Applications, 39 (2018), pp. 1360–1382.
- [7] R. CLAES, E. JARLEBRING, K. MEERBERGEN, AND P. UPADHYAYA, *Linearizable eigenvector nonlinearities*, SIAM journal on Matrix Analysis and Applications, 43 (2022), pp. 764–786.
- [8] R. CLAES, K. MEERBERGEN, AND S. TELEN, *Contour integration for eigenvector nonlinearities*, SIAM Journal on Matrix Analysis and Applications, 44 (2023), pp. 1619–1644.
- [9] D. A. COX, J. LITTLE, AND D. O’SHEA, *Using algebraic geometry*, Springer, 1998.
- [10] B. DONG, B. YU, AND Y. YU, *A homotopy method for finding all solutions of a multiparameter eigenvalue problem*, SIAM Journal on Matrix Analysis and Applications, 37 (2016), pp. 550–571.
- [11] J. D. GARDINER, A. J. LAUB, J. J. AMATO, AND C. B. MOLER, *Solution of the Sylvester matrix equation $AXB^T + CXD^T = E$* , ACM Transactions on Mathematical Software (TOMS), 18 (1992), pp. 223–231.
- [12] D. R. HARTREE, *The wave mechanics of an atom with a non-coulomb central field. Part II. Some results and discussion*, in Mathematical Proceedings of the Cambridge Philosophical Society, vol. 24, Cambridge University Press, 1928, pp. 111–132.
- [13] P. HENNING AND E. JARLEBRING, *The Gross–Pitaevskii equation and eigenvector nonlinearities: numerical methods and algorithms*, SIAM Review, 67 (2025), pp. 256–317.
- [14] M. E. HOCHSTENBACH, T. KOŠIR, AND B. PLESTENJAK, *A Jacobi–Davidson type method for the two-parameter eigenvalue problem*, SIAM Journal on Matrix Analysis and Applications, 26 (2004), pp. 477–497.
- [15] M. E. HOCHSTENBACH, T. KOŠIR, AND B. PLESTENJAK, *Numerical methods for rectangular multiparameter eigenvalue problems, with applications to finding optimal ARMA and LTI models*, Numerical linear algebra with applications, 31 (2024), p. e2540.
- [16] M. E. HOCHSTENBACH, C. MEHL, AND B. PLESTENJAK, *Solving singular generalized eigenvalue problems. Part II: Projection and augmentation*, SIAM Journal on Matrix Analysis and Applications, 44 (2023), pp. 1589–1618.
- [17] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, second ed., 2012.
- [18] E. JARLEBRING, S. KVAAL, AND W. MICHIELS, *An inverse iteration method for eigenvalue problems with eigenvector nonlinearities*, SIAM Journal on Scientific Computing, 36 (2014), pp. A1978–A2001.
- [19] E. JARLEBRING AND V. P. LITHELL, *From eigenvector nonlinearities to eigenvalue nonlinearities*, arXiv preprint arXiv:2506.16182, (2025).
- [20] W. KOHN AND L. J. SHAM, *Self-consistent equations including exchange and correlation effects*, Physical review, 140 (1965), p. A1133.
- [21] T. KOŠIR AND B. PLESTENJAK, *On the singular two-parameter eigenvalue problem II*, Linear algebra and its applications, 649 (2022), pp. 433–451.
- [22] T.-Y. LI, *Numerical solution of multivariate polynomial systems by homotopy continuation methods*, Acta numerica, 6 (1997), pp. 399–436.
- [23] K. MEERBERGEN AND Z. WANG, *The shift-and-invert arnoldi method for singular matrix pencils*, arXiv preprint arXiv:2411.02895, (2024).
- [24] T. T. NGO, M. BELLALIJ, AND Y. SAAD, *The trace ratio optimization problem*, SIAM review, 54 (2012), pp. 545–569.
- [25] P. PULAY, *Convergence acceleration of iterative sequences. The case of SCF iteration*, Chemical physics letters, 73 (1980), pp. 393–398.
- [26] E. RINGH AND E. JARLEBRING, *Nonlinearizing two-parameter eigenvalue problems*, SIAM Journal on Matrix Analysis and Applications, 42 (2021), pp. 775–799.
- [27] C. C. J. Roothaan, *New developments in molecular orbital theory*, Reviews of modern physics, 23 (1951), p. 69.
- [28] K. RUYMBEEK, K. MEERBERGEN, AND W. MICHIELS, *Subspace method for multiparameter-*

- eigenvalue problems based on tensor-train representations*, Numerical Linear Algebra with Applications, 29 (2022), p. e2439.
- [29] B. SHAPIRO AND M. SHAPIRO, *On eigenvalues of rectangular matrices*, Proceedings of the Steklov Institute of Mathematics, 267 (2009), pp. 248–255.
- [30] V. SIMONCINI, *Computational methods for linear matrix equations*, siam REVIEW, 58 (2016), pp. 377–441.