Distance to nearest skew-symmetric matrix polynomials of bounded rank

Dedicated to Paul Van Dooren on the occasion of his 75th birthday

Andrii Dmytryshyn[a], Froilán M. Dopico[b], Rakel Hellberg[a]

[a]*Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden.*
[b]*Universidad Carlos III de Madrid, ROR: https://ror.org/03ths8210, Departamento de Matemáticas, Avenida de la Universidad 30, 28911, Leganés, Spain.*

Abstract

We propose an algorithm that approximates a given matrix polynomial of degree d by another skew-symmetric matrix polynomial of a specified rank and degree at most d. The algorithm is built on recent advances in the theory of generic eigenstructures and factorizations for skew-symmetric matrix polynomials of bounded rank and degree. Taking into account that the rank of a skew-symmetric matrix polynomial is even, the algorithm works for any prescribed even rank greater than or equal to 2 and produces a skew-symmetric matrix polynomial of that exact rank. We also adapt the algorithm for matrix pencils to achieve a better performance. Lastly, we present numerical experiments for testing our algorithms and for comparison to the previously known ones.

Keywords: complete eigenstructure, genericity, low rank approximation, matrix nearness problems, matrix polynomials, skew-symmetry, normal rank, orbits, pencils
2000 MSC: 65F55, 15A18, 15A21

1. Introduction

Matrix nearness problems—quantifying the minimal perturbation needed to endow a matrix or matrix-valued function with a desired property—are fundamental in numerical linear algebra [28, 30, 35]. Among these, determining the distance to the nearest singular matrix polynomial (including

---

Email addresses: `andrii@chalmers.se` (Andrii Dmytryshyn), `dopico@math.uc3m.es` (Froilán M. Dopico), `rakel.hellberg@outlook.com` (Rakel Hellberg)

pencils) has garnered significant attention [2, 3, 22, 24, 25, 27, 29] due to its relevance in control theory, system identification, and the analysis of differential-algebraic equations.

In this work, we focus on a structured and more general variant of the distance to the singularity problem, namely, on computing the distance of a given matrix polynomial of degree d to the nearest skew-symmetric matrix polynomial of a given bounded rank and degree at most d and on computing such a nearest polynomial. Note that we impose a stronger condition than just the singularity on such nearest polynomial and that the distance to the singularity problem is a particular case of the problem we consider. Skew-symmetric matrix polynomials arise naturally in Hamiltonian systems, gyroscopic dynamics, and other physical models where antisymmetric structure encodes conservation laws or symmetries. Moreover, canonical forms of skew-symmetric matrix pencils have been studied in [38, 40], low-rank perturbations in [1], miniversal deformations in [13], codimensions of their orbits were computed in [21], and their stratifications were developed in [4, 14, 20].

Two recent methodological advances provide powerful tools for addressing the distance to the singularity problem for matrix polynomials (among the other challenges they also address). The first, introduced by Gnazzo and Guglielmi in [25, 26], considers the distance to singularity for general nonlinear matrix-valued functions. Their approach reformulates the problem as a non-convex optimization task, encoding singularity as a constraint. Another approach, called the Riemann-Oracle [27], developed by Gnazzo, Noferini, Nyman, and Poloni, offers a general-purpose Riemannian optimization strategy for a broad class of matrix nearness problems.

In this paper, we follow the approach of [5] and build on recent advances in the theory of generic eigenstructures and factorizations for skew-symmetric matrix polynomials of bounded rank and degree [17, 18]. Notably, our approach is completely different from those in [25, 26, 27] and allows us to consider the distance to matrix polynomials of any specified rank in the same way as the distance to the singularity.

During recent years, the problem of determining the generic eigenstructures for sets of matrix polynomials (including pencils) with fixed grade and bounded rank, has been intensively investigated. The description of such sets as a union of closures of a few "generic sets" of matrix polynomials with simple eigenstructures has been given for general matrix polynomials [15] (see also [10] for the result on pencils), for skew-symmetric matrix polynomials [12, 16] (including pencils), as well as for symmetric matrix polynomials of odd grade [8] (see also [7] for the result on pencils). Moreover, the cases of matrix pencils that are T-palindromic and T-alternating [6], and Hermitian [9] are also known. In this paper we develop a parametrization of the generic

sets for skew-symmetric matrix polynomials with fixed grade and bounded rank in terms of factorizations with very specific properties, similarly to how it was done in [17] for general matrix polynomials. This parametrization allows us to develop a simple and efficient algorithm for solving the distance problem mentioned above.

The rest of the paper is organized as follows. We begin by providing, in Section 2, the necessary background on skew-symmetric matrix polynomials, introducing key definitions and properties that will be used throughout the paper. Section 3 presents the announced parametrization of the generic sets for skew-symmetric matrix polynomials with fixed grade and bounded rank in terms of factorizations. In Section 4, we present a general solution to the distance problem for skew-symmetric matrix polynomials. Then, in Section 5, we review the relevant background on skew-symmetric matrix pencils, i.e., degree-one skew-symmetric matrix polynomials. In Section 6, we adapt our algorithm from Section 4 to the case of skew-symmetric matrix pencils and improve its computational efficiency. Finally, Section 7 presents numerical experiments.

2. Skew-symmetric matrix polynomials

We consider skew-symmetric $m \times m$ matrix polynomials $P(\lambda)$ of grade d, i.e., of degree less than or equal to d, over the field of complex numbers $\mathbb{C}$:

$$P(\lambda) = \lambda^d A_d + \cdots + \lambda A_1 + A_0, \quad A_i^T = -A_i, \ A_i \in \mathbb{C}^{m \times m} \text{ for } i = 0, \ldots, d.$$

Recall that the degree of $P(\lambda)$, denoted as $\deg(P)$, is the largest index k such that $A_k \neq 0$. By $\mathrm{POL}_{d,m}^{ss}$ we denote the vector space of $m \times m$ skew-symmetric matrix polynomials of grade d. If there is no risk of confusion, we write POL instead of $\mathrm{POL}_{d,m}^{ss}$. Using the Frobenius matrix norm of complex matrices [31], a distance in $\mathrm{POL}_{d,m}^{ss}$ is defined as $d(P, P') = \left(\sum_{i=0}^{d} \|A_i - A_i'\|_F^2 \right)^{\frac{1}{2}}$, where $P(\lambda) = \sum_{i=0}^{d} \lambda^i A_i$ and $P'(\lambda) = \sum_{i=0}^{d} \lambda^i A_i'$, making $\mathrm{POL}_{d,m}^{ss}$ a metric space with the Euclidean topology induced by this distance. We define the Frobenius norm of the matrix polynomial P as $\|P(\lambda)\|_F = \left(\sum_{i=0}^{d} \|A_i\|_F^2 \right)^{\frac{1}{2}}$. The fact that $\mathrm{POL}_{d,m}^{ss}$ is a metric space allows us to consider the closure of any of its subsets $\mathcal{S}$, which is denoted by $\overline{\mathcal{S}}$.

Two matrix polynomials $P(\lambda)$ and $Q(\lambda)$ are called *unimodularly congruent* if $F(\lambda)^T P(\lambda) F(\lambda) = Q(\lambda)$ for some unimodular matrix polynomial $F(\lambda)$ (i.e. $\det F(\lambda) \in \mathbb{C}\backslash\{0\}$), see also [37].

Theorem 2.1. [37] *Let $P(\lambda)$ be a skew-symmetric $m \times m$ matrix polynomial. Then there exist $r \in \mathbb{N}$ with $2r \leqslant m$ and a unimodular matrix polynomial $F(\lambda)$*

such that

$$F(\lambda)^T P(\lambda) F(\lambda) = \begin{bmatrix} 0 & g_1(\lambda) \\ -g_1(\lambda) & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & g_r(\lambda) \\ -g_r(\lambda) & 0 \end{bmatrix} \oplus 0_{m-2r} =: S(\lambda),$$

where g_j is a monic polynomial, for $j = 1, \ldots, r$, and $g_j(\lambda)$ divides $g_{j+1}(\lambda)$, for $j = 1, \ldots, r - 1$. Moreover, the canonical form $S(\lambda)$ is unique.

The $2r$ monic scalar polynomials $g_1(\lambda), g_1(\lambda), g_2(\lambda), g_2(\lambda), \ldots, g_r(\lambda), g_r(\lambda)$ in Theorem 2.1 are called the *invariant polynomials* of $P(\lambda)$, and, for any $\alpha \in \mathbb{C}$, each of them can be uniquely factored as

$$g_j(\lambda) = (\lambda - \alpha)^{\sigma_j} p_j(\lambda), \quad \text{with } p_j(\alpha) \neq 0$$

and $\sigma_j \geqslant 0$ being an integer. The sequence $0 \leq \sigma_1 = \sigma_1 \leq \sigma_2 = \sigma_2 \leq \cdots \leq \sigma_r = \sigma_r$ is called *the sequence of partial multiplicities* of $P(\lambda)$ at α. The number α is a finite eigenvalue of $P(\lambda)$ if the partial multiplicity sequence at α contains at least two nonzero terms, or, equivalently, if α is a root of at least one invariant polynomial $g_j(\lambda)$. The *elementary divisors* of $P(\lambda)$ associated with a finite eigenvalue α is the collection of factors $(\lambda - \alpha)^{\sigma_1}, (\lambda - \alpha)^{\sigma_1}, (\lambda - \alpha)^{\sigma_2}, (\lambda - \alpha)^{\sigma_2}, \ldots, (\lambda - \alpha)^{\sigma_r}, (\lambda - \alpha)^{\sigma_r}$ for which $\sigma_j > 0$. The (normal) rank of the skew-symmetric matrix polynomial $P(\lambda)$ in Theorem 2.1 is equal to the nonnegative integer $2r$, i.e., the normal rank is always even. Note also that a square matrix polynomial $P(\lambda)$ is singular if and only if $\det(P(\lambda)) = 0$, for all λ.

For completion, we recall that $\lambda = \infty$ is an eigenvalue of the matrix polynomial $P(\lambda) \in \mathrm{POL}_{d,m}^{ss}$ if zero is an eigenvalue of $\mathrm{rev}_d P(\lambda) := \lambda^d P(1/\lambda)$. The elementary divisors and partial multiplicities for the zero eigenvalue of $\mathrm{rev}_d P(\lambda)$ are the elementary divisors and partial multiplicities associated with the infinite eigenvalue of $P(\lambda)$.

For an $m \times n$ matrix polynomial $P(\lambda)$, the left and right null-spaces, over the field of rational functions $\mathbb{C}(\lambda)$, are defined as follows:

$$\mathcal{N}_{\mathrm{left}}(P) := \{ y(\lambda)^T \in \mathbb{C}(\lambda)^{1 \times m} : y(\lambda)^T P(\lambda) = 0_{1 \times n} \},$$
$$\mathcal{N}_{\mathrm{right}}(P) := \{ x(\lambda) \in \mathbb{C}(\lambda)^{n \times 1} : P(\lambda) x(\lambda) = 0_{m \times 1} \}.$$

Each subspace $\mathcal{V}$ of $\mathbb{C}(\lambda)^n$ has bases consisting entirely of vector polynomials. A basis of $\mathcal{V}$ consisting of vector polynomials whose sum of degrees is minimal among all bases of $\mathcal{V}$ consisting of vector polynomials is called a *minimal basis* of $\mathcal{V}$. The ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{V}$ is always the same. These degrees are called the minimal indices of $\mathcal{V}$ [23, 34]. Thus we can define the left and right *minimal indices* of a matrix

polynomial $P(\lambda)$ as those of $\mathcal{N}_{\text{left}}(P)$ and $\mathcal{N}_{\text{right}}(P)$, respectively. Note that for a skew-symmetric matrix polynomial the left minimal indices are equal to the right ones.

Following [12, 14, 16], we define the *complete eigenstructure* of a skew-symmetric matrix polynomial $P(\lambda) \in \text{POL}_{d,m}^{ss}$ *of grade d* to be the collection of all finite and infinite eigenvalues, the corresponding elementary divisors (or, equivalently, the corresponding sequences of partial multiplicities), and the left and right minimal indices of $P(\lambda)$. For convenience, we also define the notion of orbit of a skew-symmetric matrix polynomial $P(\lambda) \in \text{POL}_{d,m}^{ss}$, denoted $\text{O}(P(\lambda))$, as the subset of matrix polynomials in $\text{POL}_{d,m}^{ss}$ with the same complete eigenstructure as $P(\lambda)$. Observe that all the polynomials in $\text{O}(P)$ have the same *degree* and the same rank. Note that the definition of orbit for general polynomials is analogous and given in, e.g., [15, 19].

The subset of matrix polynomials in $\text{POL}_{d,m}^{ss}$ with normal rank at most $2r$ will be denoted by $\text{POL}_{d,2r,m}^{ss}$, or simply by POL_{2r} when there is no need to specify the other parameters or they are clear from the context. A simple description of $\text{POL}_{d,2r,m}^{ss}$—as the closure of a certain "generic set" characterized by its eigenstructure—was obtained in [12, 16] and is recalled in Theorem 2.2. Such descriptions are also known for matrix polynomials with other types of structures [8, 15].

Theorem 2.2. [12] *Let m, r, and d be integers such that $m > 2$, $d \geq 1$, and $2 \leq 2r \leq (m-1)$. The set of $m \times m$ complex skew-symmetric matrix polynomials of grade d with rank at most $2r$ is a closed subset of $\text{POL}_{d,m}^{ss}$ equal to $\overline{\text{O}(W)}$, where $W \in \text{POL}_{d,m}^{ss}$ is an $m \times m$ complex skew-symmetric matrix polynomial of degree exactly d and rank exactly $2r$ with no elementary divisors at all, with t left minimal indices equal to $(\beta + 1)$ and with $(m - 2r - t)$ left minimal indices equal to β, where $\beta = \lfloor rd/(m-2r) \rfloor$ and $t = rd \mod (m-2r)$, and with the right minimal indices equal to the left minimal indices.*

Theorem 2.2 is one of the main tools for obtaining another simple description of $\text{POL}_{d,2r,m}^{ss}$ as the closure of a certain "generic set" of skew-symmetric matrix polynomials characterized by the simple way in which they can be factorized. This simple description is obtained in Section 3 and is the key result on which the algorithm we propose is based on.

3. Generic factorizations of skew-symmetric matrix polynomials with bounded rank

Theorem 3.1 is a key result for the development of the algorithm presented in this paper. It is one of the results included in [18]. Although the statement

of Theorem 3.1 is very simple, its proof is long, relies on several preliminary lemmas, and is based on results in [17]. Therefore, the proof is included in Subsection 3.1, which can be ommitted without affecting the understanding of the rest of the paper. In Theorem 3.1, as well as in other results, some general matrix polynomials (i.e., not necessarily skew-symmetric) are used, which motivates the introduction of the following notation: $\mathrm{POL}_d^{p \times q}$ is the vector space of $p \times q$ matrix polynomials of grade d over $\mathbb{C}$.

Theorem 3.1. *Let m, r and d be integers such that $d \geq 1$ and $2 \leq 2r \leq (m-1)$. Let us define the following subset of $\mathrm{POL}_{d,2r,m}^{ss}$*

$$\mathcal{G}_{d,2r,m}^{ss} := \left\{ U(\lambda)V(\lambda)^T - V(\lambda)U(\lambda)^T : \begin{array}{l} V(\lambda) \in \mathrm{POL}_{\lceil d/2 \rceil}^{m \times r}, \\[2mm] U(\lambda) \in \mathrm{POL}_{\lfloor d/2 \rfloor}^{m \times r} \end{array} \right\}.$$

Then,

$$\overline{\mathcal{G}_{d,2r,m}^{ss}} = \mathrm{POL}_{d,2r,m}^{ss}.$$

3.1. Proof of Theorem 3.1

The proof uses several definitions and results that can be found in [17, Sections 2 & 3], which are not repeated here for the sake of brevity. We use often the i-th column of a matrix polynomial $L(\lambda)$, which is denoted by $L_{*i}(\lambda)$ of simply by L_{*i}. In addition to the notations used above for different sets of matrix polynomials, we use in this subsection $\mathrm{POL}^{p \times q}$ to denote the set of $p \times q$ matrix polynomials over $\mathbb{C}$, without specifying the grade. The (normal) rank of a matrix polynomial $L(\lambda)$ is denoted by $\mathrm{rank}(L)$. For $k = 1, 2, \ldots$, we define the following $k \times k$ matrix

$$\Delta_k := \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \cdot^{\cdot^{\cdot}} & & \\ 1 & & & \end{bmatrix}.$$

With this definition, it can be shown that the set $\mathcal{G}_{d,2r,m}^{ss}$ in Theorem 3.1 can be expressed as follows

$$\mathcal{G}_{d,2r,m}^{ss} = \left\{ L(\lambda) \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} L(\lambda)^T : \begin{array}{l} L(\lambda) \in \mathrm{POL}^{m \times 2r}, \\[2mm] \deg(L_{*i}) \leq \left\lceil \dfrac{d}{2} \right\rceil, \quad \text{for } i = 1, \ldots, r, \\[3mm] \deg(L_{*i}) \leq \left\lfloor \dfrac{d}{2} \right\rfloor, \quad \text{for } i = r+1, \ldots, 2r \end{array} \right\},$$

(1)

6

which is more adequate for the purposes of this subsection. To see that the set in the right-hand side of (1) is indeed $\mathcal{G}^{ss}_{d,2r,m}$, it is enough to express $L(\lambda)$ as $L(\lambda) = \begin{bmatrix} -V(\lambda)\Delta_r & U(\lambda) \end{bmatrix}$.

If $P(\lambda) \in \mathrm{POL}^{ss}_{d,m}$, then its rational column space, as defined in [17, Definition 2.6], satisfies $\mathcal{C}ol(P) = \mathcal{C}ol(P^T)$. Lemma 3.2 is the first auxiliary result we need for proving Theorem 3.1. It is a consequence of [17, Theorems 3.11 and 3.19] and uses the well-known concept of *minimal basis* of a rational subspace [23] (see also [17, pp. 742-743]). Following a standard practice, we often say that a matrix polynomial is a minimal basis when its columns are a minimal basis of the rational subspace they span.

Lemma 3.2.   (i) *If $P(\lambda) \in \mathrm{POL}^{ss}_{d,2r,m}$ and $\mathrm{rank}(P) = 2r$, then $P(\lambda) = L(\lambda)E(\lambda)L(\lambda)^T$, where the columns of $L(\lambda) \in \mathrm{POL}^{m \times 2r}$ form a minimal basis of $\mathcal{C}ol(P)$ and $E(\lambda) \in \mathrm{POL}^{2r \times 2r}$ is skew-symmetric with $\mathrm{rank}(E) = 2r$.*

   (ii) *If $P(\lambda) \in \mathrm{POL}^{ss}_{d,2r,m}$, $\mathrm{rank}(P) = 2r$, $\deg(P) = d$, $P(\lambda)$ has not eigenvalues (finite or infinite) and the columns of $L(\lambda)$ in the factorization in item* (i) *are ordered in such a way that $\deg(L_{*1}) \geq \deg(L_{*2}) \geq \cdots \geq \deg(L_{*2r})$, then*

$$\deg(L_{*i}) + \deg(L_{*(2r-i+1)}) = d, \quad for\ i = 1, \ldots, 2r.$$

Proof. Item (i). Except for the skew-symmetry of $E(\lambda)$, the existence of a factorization $P(\lambda) = L(\lambda)E(\lambda)L(\lambda)^T$ with the stated properties follows immediately from combining Theorem 3.11-(i) and Remark 3.13 in [17] with the fact that $\mathcal{C}ol(P) = \mathcal{C}ol(P^T)$. For proving the skew-symmetry of $E(\lambda)$, we use that since $L(\lambda)$ is a minimal basis, it has a polynomial left inverse $M(\lambda) \in \mathrm{POL}^{2r \times m}$ such that $M(\lambda)L(\lambda) = I_{2r}$ [23, Theorem 4]. So, $E(\lambda) = M(\lambda)P(\lambda)M(\lambda)^T$ and the skew-symmetry of $E(\lambda)$ follows from that of $P(\lambda)$.

Item (ii). Theorem 3.19 in [17] implies that $P(\lambda) = L(\lambda)U(\lambda)$, where $L(\lambda), U(\lambda)^T \in \mathrm{POL}^{m \times 2r}$ are two minimal bases of $\mathcal{C}ol(P)$ and

$$\deg(L_{*i}) + \deg((U^T)_{*i}) = d, \quad \text{for } i = 1, \ldots, 2r. \tag{2}$$

Since, $L(\lambda)$ and $U(\lambda)^T$ are minimal bases of the same rational subspace, there exists a reordering of the columns of $U(\lambda)^T$ that makes the degrees of its columns equal to the degrees of the columns of $L(\lambda)$. Moreover, (2) guarantees that if the degrees of the columns of $L(\lambda)$ are decreasingly ordered, then those of the columns of $U(\lambda)^T$ are increasingly ordered. $\qquad \square$

Motivated by the properties of the matrix polynomials in the orbit $\mathrm{O}(W)$ in Theorem 2.2, we introduce the following subset of $\mathrm{POL}^{ss}_{d,2r,m}$, which includes $\mathrm{O}(W)$.

Definition 3.3.

$$\text{POL}_{d,\,2r,\,m}^{ss,neig} := \left\{ P(\lambda) \in \text{POL}_{d,\,2r,\,m}^{ss} : \begin{array}{l} \text{rank}(P) = 2r, \\ \deg(P) = d, \\ P(\lambda) \text{ has no eigenvalues} \end{array} \right\}.$$

We emphasize that when we say that a matrix polynomial has no eigenvalues, we refer to both finite and infinite eigenvalues.

The next lemma is a corollary of Theorem 2.2 and of $O(W) \subset \text{POL}_{d,\,2r,\,m}^{ss,neig} \subset \text{POL}_{d,\,2r,\,m}^{ss}$. One just need to take closures in this chain of inclusions. Note that Theorem 2.2 states, in particular, that $\text{POL}_{d,\,2r,\,m}^{ss}$ is a closed subset of $\text{POL}_{d,\,m}^{ss}$, which, on the other hand, is obvious from the definition of (normal) rank. Lemma 3.4 is important in the proof of Theorem 3.1.

Lemma 3.4. $\overline{\text{POL}_{d,\,2r,\,m}^{ss,neig}} = \text{POL}_{d,\,2r,\,m}^{ss}$.

Theorem 3.5 describes the set $\text{POL}_{d,\,2r,\,m}^{ss,neig}$ in terms of minimal rank factorizations.

Theorem 3.5.

$$\text{POL}_{d,\,2r,\,m}^{ss,neig} = \left\{ L(\lambda) \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} L(\lambda)^T : \begin{array}{l} L(\lambda) \in \text{POL}^{m \times 2r}, \\ L(\lambda) \text{ is a minimal basis}, \\ \deg(L_{*1}) \geq \cdots \geq \deg(L_{*2r}), \\ \deg(L_{*i}) + \deg(L_{*(2r-i+1)}) = d, \\ \text{for } i = 1, \ldots, 2r \end{array} \right\}.$$

Proof. Let $\mathcal{S}$ be the set in the right-hand side of the equality in Theorem 3.5. The inclusion $\mathcal{S} \subseteq \text{POL}_{d,\,2r,\,m}^{ss,neig}$ is a direct consequence of [17, Theorem 3.19]. Therefore, we focus on proving that if $P(\lambda) \in \text{POL}_{d,\,2r,\,m}^{ss,neig}$, then $P(\lambda) \in \mathcal{S}$. By Lemma 3.2, any $P(\lambda) \in \text{POL}_{d,\,2r,\,m}^{ss,neig}$ can be factorized as

$$P(\lambda) = \widetilde{L}(\lambda)\widetilde{E}(\lambda)\widetilde{L}(\lambda)^T, \tag{3}$$

where the columns of $\widetilde{L}(\lambda) \in \text{POL}^{m \times 2r}$ are a minimal basis of $\mathcal{C}ol(P)$ and $\widetilde{E}(\lambda) \in \text{POL}^{2r \times 2r}$ is skew-symmetric with $\text{rank}(\widetilde{E}) = 2r$. Moreover, $\deg(\widetilde{L}_{*1}) \geq \deg(\widetilde{L}_{*2}) \geq \cdots \geq \deg(\widetilde{L}_{*2r})$, and

$$\deg(\widetilde{L}_{*i}) + \deg(\widetilde{L}_{*(2r-i+1)}) = d, \quad \text{for } i = 1, \ldots, 2r. \tag{4}$$

Let us assume that the first r columns of $\widetilde{L}(\lambda)$ have p different degrees $\widetilde{d}_1 > \cdots > \widetilde{d}_p$ and that there are r_i columns with degree $\widetilde{d}_i$ for $i = 1, \ldots, p$. Note that $r = r_1 + \cdots + r_p$. From (4), we get that the last r columns of $\widetilde{L}(\lambda)$ also

8

have p different degrees $\widetilde{d}_{p+1} > \cdots > \widetilde{d}_{2p}$ and that there are r_i columns among the last r columns with degree

$$\widetilde{d}_{2p-i+1} = d - \widetilde{d}_i, \quad \text{for } i = 1, \ldots, p, \tag{5}$$

which are in the "mirror" positions to the corresponding ones in the first r columns. These considerations induce the following partition of $\widetilde{L}(\lambda)$

$$\widetilde{L}(\lambda) = \begin{bmatrix} \widetilde{L}_1 & \cdots & \widetilde{L}_p & \widetilde{L}_{p+1} & \cdots & \widetilde{L}_{2p} \end{bmatrix}, \tag{6}$$

where we omit the dependence in λ in the blocks $\widetilde{L}_i$,

$$\begin{aligned}
&\widetilde{L}_i \in \mathrm{POL}^{m \times r_i}, \text{ for } i = 1, \ldots, p, \\
&\widetilde{L}_i \in \mathrm{POL}^{m \times r_{2p-i+1}}, \text{ for } i = p+1, \ldots, 2p, \\
&\deg(\widetilde{L}_i) = \widetilde{d}_i, \text{ for } i = 1, \ldots, 2p, \\
&\widetilde{d}_1 > \cdots > \widetilde{d}_p \geq \widetilde{d}_{p+1} > \cdots > \widetilde{d}_{2p},
\end{aligned} \tag{7}$$

and all the columns in the same block $\widetilde{L}_i$ have the same degree $\widetilde{d}_i$. Note that $\widetilde{d}_p \geq d/2$ and $\widetilde{d}_{p+1} \leq d/2$, because $\widetilde{d}_p \geq \widetilde{d}_{p+1}$ and $\widetilde{d}_p + \widetilde{d}_{p+1} = d$ by (5). So,

$$\widetilde{d}_1 > \cdots > \widetilde{d}_p \geq d/2 \geq \widetilde{d}_{p+1} > \cdots > \widetilde{d}_{2p}. \tag{8}$$

The partition of $\widetilde{L}(\lambda)$ induces a conformable partition in the skew-symmetric matrix polynomial $\widetilde{E}(\lambda)$. By [17, Corollary 3.17-(ii)] such a partition satisfies the following

$$\widetilde{E}(\lambda) = \left(\widetilde{E}_{ij}\right)_{1 \leq i,j \leq 2p} \text{ with } \widetilde{E}_{ij} = -\widetilde{E}_{ji}^T, \quad \deg(\widetilde{E}_{ij}) \leq d - \widetilde{d}_i - \widetilde{d}_j, \tag{9}$$

and where we omit the dependence in λ in the blocks $\widetilde{E}_{ij}$.

The next step in the proof is to show that $\widetilde{E}(\lambda)$ can be chosen to be lower block anti-triangular, i.e., $\widetilde{E}_{ij} = 0$ for $i + j \leq 2p$. For that purpose, we consider two cases. First, if $\widetilde{d}_p > \widetilde{d}_{p+1}$, then (8) becomes $\widetilde{d}_1 > \cdots > \widetilde{d}_p > d/2 > \widetilde{d}_{p+1} > \cdots > \widetilde{d}_{2p}$, which, combined with the degree bound in (9) and (5), implies that $\widetilde{E}(\lambda)$ is indeed lower block anti-triangular. Second, if $\widetilde{d}_p = d/2 = \widetilde{d}_{p+1}$, then the degree bound in (9) and (5) only guarantee $\widetilde{E}_{ij} = 0$ for $i + j \leq 2p$ and $(i,j) \neq (p,p)$. In this case, the skew-symmetric submatrix $\left(\widetilde{E}_{ij}\right)_{p \leq i,j \leq p+1}$ containing the only potential non-zero block above the main block anti-diagonal must be invertible, since $\mathrm{rank}(\widetilde{E}) = 2r$, and must have degree zero by the degree bound in (9) and $\widetilde{d}_p = d/2 = \widetilde{d}_{p+1}$. That is, it is

a constant invertible matrix. Therefore by [33, Corollary 2.6.6-(b)], there exists an invertible constant complex matrix Q such that

$$Q \begin{bmatrix} \widetilde{E}_{p,p} & \widetilde{E}_{p,p+1} \\ \widetilde{E}_{p+1,p} & \widetilde{E}_{p+1,p+1} \end{bmatrix} Q^T = \begin{bmatrix} 0 & H \\ -H^T & 0 \end{bmatrix},$$

with $H \in \mathbb{C}^{r_p \times r_p}$ invertible. Define $\widetilde{Q} = \mathrm{diag}(I_\ell, Q, I_\ell)$, where $\ell = r - r_p$, and use (3) to express $P(\lambda)$ as

$$P(\lambda) = (\widetilde{L}(\lambda)\widetilde{Q}^{-1})\,(\widetilde{Q}\widetilde{E}(\lambda)\widetilde{Q}^T)\,(\widetilde{L}(\lambda)\widetilde{Q}^{-1})^T.$$

Observe that $(\widetilde{Q}\widetilde{E}(\lambda)\widetilde{Q}^T)$ is now skew-symmetric, lower block anti-triangular, and the properties of its blocks are still those stated in (9). On the other hand, the polynomial matrix $(\widetilde{L}(\lambda)\widetilde{Q}^{-1})$ is a minimal basis of $\mathcal{C}ol(P)$, because has full column rank $2r$ and the degrees of their columns are the same as those of $\widetilde{L}(\lambda)$, and admits a partition as the one in (6) with the properties stated in (7). To avoid additional notation, we actualize the definitions of $\widetilde{L}(\lambda) \leftarrow (\widetilde{L}(\lambda)\widetilde{Q}^{-1})$ and of $\widetilde{E}(\lambda) \leftarrow (\widetilde{Q}\widetilde{E}(\lambda)\widetilde{Q}^T)$. That is, in the rest of the proof, we work with a factorization of $P(\lambda)$ as in (3) assuming that $\widetilde{E}(\lambda)$ is block lower anti-triangular. Observe that the blocks $\widetilde{E}_{i,2p-i+1}$, $i = 1, \dots, 2p$, on the main block anti-diagonal must be constant matrices, due to (5) and (9), and invertible, because $\mathrm{rank}(\widetilde{E}) = 2r$.

The lower block anti-triangularity and the skew-symmetry of $\widetilde{E}(\lambda)$ allow us to establish, through a direct multiplication, the following factorization

$$\widetilde{E}(\lambda) = X(\lambda)\,\Lambda\,X(\lambda)^T, \tag{10}$$

where

$$X(\lambda) := \begin{bmatrix} \widetilde{E}_{1,2p}\Delta_{r_1} & & & & & \\ \widetilde{E}_{2,2p}\Delta_{r_1} & \widetilde{E}_{2,2p-1}\Delta_{r_2} & & & & \\ \widetilde{E}_{3,2p}\Delta_{r_1} & \widetilde{E}_{3,2p-1}\Delta_{r_2} & \ddots & & & \\ \vdots & \vdots & & \widetilde{E}_{p,p+1}\Delta_{r_p} & & \\ \vdots & \vdots & & \frac{\widetilde{E}_{p+1,p+1}}{2}\Delta_{r_p} & I_{r_p} & \\ \widetilde{E}_{2p-2,2p}\Delta_{r_1} & \widetilde{E}_{2p-2,2p-1}\Delta_{r_2} & \cdot^{\cdot^{\cdot}} & & & \ddots \\ \widetilde{E}_{2p-1,2p}\Delta_{r_1} & \frac{\widetilde{E}_{2p-1,2p-1}}{2}\Delta_{r_2} & & & & I_{r_2} \\ \frac{\widetilde{E}_{2p,2p}}{2}\Delta_{r_1} & & & & & I_{r_1} \end{bmatrix}$$

10

and

$$\Lambda := \begin{bmatrix} & & & & & \Delta_{r_1} \\ & & & & \iddots & \\ & & & \Delta_{r_p} & & \\ & & -\Delta_{r_p} & & & \\ & \iddots & & & & \\ -\Delta_{r_1} & & & & & \end{bmatrix} = \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} \in \mathbb{C}^{2r \times 2r}.$$

Note that the diagonal blocks of $X(\lambda)$ are constant invertible matrices and, so, the matrix polynomial $X(\lambda)$ is invertible (in fact, it is unimodular). Combining the factorizations (3) and (10), we get

$$P(\lambda) = (\widetilde{L}(\lambda)X(\lambda)) \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} (\widetilde{L}(\lambda)X(\lambda))^T, \tag{11}$$

which imply that the columns of $L(\lambda) := \widetilde{L}(\lambda)X(\lambda)$ form a polynomial basis of $\mathcal{C}ol(P)$. Taking into account (5), (6), (7), and (9), we can partition $L(\lambda)$ as

$$L(\lambda) = \begin{bmatrix} L_1 & \cdots & L_p & L_{p+1} & \cdots & L_{2p} \end{bmatrix}, \tag{12}$$

where $L_j = \widetilde{L}_j$ for $j = p+1, \ldots, 2p$, and $\deg(L_j) \leq \deg(\widetilde{L}_j) = \widetilde{d}_j$ for $j = 1, \ldots, p$. But if some column of some block L_j had degree strictly smaller than $\widetilde{d}_j$, we would have a polynomial basis of $\mathcal{C}ol(P)$ with the sum of all the degrees of its vectors strictly smaller than the sum of all the degrees of the column vectors of the minimal basis $\widetilde{L}(\lambda)$ of $\mathcal{C}ol(P)$, which is impossible. So, $L(\lambda)$ is a minimal basis with the degrees of its columns exactly equal to those of $\widetilde{L}(\lambda)$, which concludes the proof. $\qquad\square$

The factorizations in Theorem 3.5 satisfy

$$\deg(L_{*1}) \geq \cdots \geq \deg(L_{*r}) \geq d/2 \geq \deg(L_{*(r+1)}) \geq \cdots \geq \deg(L_{*2r}),$$

because otherwise $\deg(L_{*r}) + \deg(L_{*(r+1)}) = d$ would not hold. This motivates the following definition of certain subsets of $\mathrm{POL}^{ss}_{d,2r,m}$.

Definition 3.6. *Let m, r and d be integers such that $d \geq 1$ and $2 \leq 2r \leq (m-1)$. Let $d_1, d_2, \ldots, d_r$ be integers such that $d \geq d_i \geq d/2$ for $i = 1, \ldots, r$. Then*

$$\mathcal{A}(d_1, \ldots, d_r) := \left\{ L(\lambda) \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} L(\lambda)^T : \begin{array}{l} L(\lambda) \in \mathrm{POL}^{m \times 2r}, \\ \deg(L_{*i}) = d_i, \\ \deg(L_{*(2r-i+1)}) = d - d_i, \\ \quad for\ i = 1, \ldots, r \end{array} \right\}.$$

If Definition 3.6 is compared with Theorem 3.5, we see that in the matrix polynomials in $\mathcal{A}(d_1, \ldots, d_r)$, the factor $L(\lambda)$ is not required to be a minimal basis and that the columns of $L(\lambda)$ are not necessarily arranged in decreasing degree order. On the other hand, it is obvious that $\mathcal{A}(d_1, \ldots, d_r) \subset \mathrm{POL}_{d, 2r, m}^{ss}$ and that any polynomial in $\mathrm{POL}_{d, 2r, m}^{ss, neig}$ belongs to one of these sets $\mathcal{A}(d_1, \ldots, d_r)$.

The key property of the sets in Definition 3.6 in the context of this paper is stated in the next theorem. We observe that if in the description of the set $\mathcal{G}_{d, 2r, m}^{ss}$ in (1) the degree inequalities were replaced by strict equalities, then we would obtain the set $\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)$ that appears in Theorem 3.7.

Theorem 3.7. *If $d_1, d_2, \ldots, d_r$ are any integers such that $d \geq d_i \geq d/2$ for $i = 1, \ldots, r$, then*

$$\mathcal{A}(d_1, \ldots, d_r) \subseteq \overline{\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)}.$$

Proof. Let us assume that there exists one integer $d_j > \left\lceil \frac{d}{2} \right\rceil$. Otherwise, there is nothing to prove. Along the proof, for any index $i = 1, \ldots, r$, we define $i' := 2r - i + 1$ and $d_{i'} := d - d_i$. Let $P(\lambda) \in \mathcal{A}(d_1, \ldots, d_r)$. Then,

$$P(\lambda) = \sum_{i=1}^{r} P_i(\lambda), \quad \text{with } P_i(\lambda) := \begin{bmatrix} L_{*i} & L_{*i'} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} L_{*i}^T \\ L_{*i'}^T \end{bmatrix},$$

where we omit the dependence in λ in the columns of $L(\lambda)$. Let us focus on the summand $P_j(\lambda)$ corresponding to $d_j > \left\lceil \frac{d}{2} \right\rceil$, which implies $d_{j'} < \left\lfloor \frac{d}{2} \right\rfloor$ and $d_j - d_{j'} \geq 2$. Then

$$L_{*j} = v_j \lambda^{d_j} + v(\lambda) \quad \text{and} \quad L_{*j'} = w_{j'} \lambda^{d_{j'}} + w(\lambda),$$

where $v_j, w_{j'} \in \mathbb{C}^m$ are nonzero constant vectors, and $v(\lambda)$ and $w(\lambda)$ are polynomial vectors such that $\deg(v(\lambda)) < d_j$ and $\deg(w(\lambda)) < d_{j'}$. For $\varepsilon > 0$ define

$$L_{*j'}^{(\varepsilon)} := \varepsilon \lambda^{d_{j'}+1} v_j + w_{j'} \lambda^{d_{j'}} + w(\lambda) \text{ and } P_j^{(\varepsilon)}(\lambda) := \begin{bmatrix} L_{*j} & L_{*j'}^{(\varepsilon)} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} L_{*j}^T \\ (L_{*j'}^{(\varepsilon)})^T \end{bmatrix}.$$

Observe that $\lim_{\varepsilon \to 0} P_j^{(\varepsilon)}(\lambda) = P_j(\lambda)$. Since

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{\lambda^{d_j - d_{j'} - 1}}{\varepsilon} & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{\lambda^{d_j - d_{j'} - 1}}{\varepsilon} & 1 \end{bmatrix}^T,$$

we have that

$$P_j^{(\varepsilon)}(\lambda) := \left(\begin{bmatrix} L_{*j} & L_{*j'}^{(\varepsilon)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{\lambda^{d_j - d_{j'} - 1}}{\varepsilon} & 1 \end{bmatrix} \right) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \left(\begin{bmatrix} L_{*j} & L_{*j'}^{(\varepsilon)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{\lambda^{d_j - d_{j'} - 1}}{\varepsilon} & 1 \end{bmatrix} \right)^T.$$

Note that the columns of

$$
\begin{bmatrix} L_{*j} & L_{*j'}^{(\varepsilon)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{\lambda^{d_j-d_{j'}-1}}{\varepsilon} & 1 \end{bmatrix} = \begin{bmatrix} v(\lambda) - w_{j'}\frac{\lambda^{d_j-1}}{\varepsilon} - w(\lambda)\frac{\lambda^{d_j-d_{j'}-1}}{\varepsilon} & L_{*j'}^{(\varepsilon)} \end{bmatrix}
$$

have degrees $d_j - 1$ and $d_{j'} + 1$, respectively, if ε is sufficiently small. Then

$$
P^{(\varepsilon)}(\lambda) := P_j^{(\varepsilon)}(\lambda) + \sum_{i=1, i \neq j}^{r} P_i(\lambda) \in \mathcal{A}(d_1, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_r)
$$

for all ε sufficiently small. This, together with $\lim_{\varepsilon \to 0} P^{(\varepsilon)}(\lambda) = P(\lambda)$, implies $P(\lambda) \in \overline{\mathcal{A}(d_1, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_r)}$ and

$$
\mathcal{A}(d_1, \ldots, d_j, \ldots, d_r) \subseteq \overline{\mathcal{A}(d_1, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_r)}.
$$

If $d_1 = \cdots = d_{j-1} = d_j - 1 = d_{j+1} = \cdots = d_r = \lceil \frac{d}{2} \rceil$, we are done. If not, we repeat the process above with the set $\mathcal{A}(d_1, \ldots, d_{j-1}, d_j - 1, d_{j+1}, \ldots, d_r)$. $\qquad \square$

We are now in the position to prove Theorem 3.8, which includes the result in Theorem 3.1.

Theorem 3.8.

$$
\overline{\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)} = \overline{\mathcal{G}_{d, 2r, m}^{ss}} = \text{POL}_{d, 2r, m}^{ss}.
$$

Proof. Note first that, since $\text{POL}_{d, 2r, m}^{ss}$ is a closed subset of $\text{POL}_{d, m}^{ss}$ and $\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right) \subseteq \text{POL}_{d, 2r, m}^{ss}$, we have that $\overline{\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)} \subseteq \text{POL}_{d, 2r, m}^{ss}$. On the other hand, if $P(\lambda) \in \text{POL}_{d, 2r, m}^{ss,neig}$, then Theorem 3.5 implies that

$$
P(\lambda) = L(\lambda) \begin{bmatrix} 0 & \Delta_r \\ -\Delta_r & 0 \end{bmatrix} L(\lambda)^T \in \mathcal{A}(\deg(L_{*1}), \ldots, \deg(L_{*r})).
$$

Therefore, $\text{POL}_{d, 2r, m}^{ss,neig} \subseteq \overline{\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)} \subseteq \text{POL}_{d, 2r, m}^{ss}$ follows from Theorem 3.7, and the equality $\overline{\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right)} = \text{POL}_{d, 2r, m}^{ss}$ from Lemma 3.4. Finally, (1) and Definition 3.6 yield $\mathcal{A}\left(\left\lceil \frac{d}{2} \right\rceil, \ldots, \left\lceil \frac{d}{2} \right\rceil\right) \subseteq \mathcal{G}_{d, 2r, m}^{ss} \subseteq \text{POL}_{d, 2r, m}^{ss}$. Taking the closures in this chain of inclusions completes the proof. $\qquad \square$

4. Distance to rank-deficient skew-symmetric matrix polynomials

Let $P(\lambda) = \sum_{i=0}^{d} \lambda^i P_i$, $P_i \in \mathbb{C}^{m \times m}$, $i = 0, \ldots, d$, be a given matrix polynomial of degree d. Our goal is to find a matrix polynomial in $\text{POL}_{d, 2r, m}^{ss} =$

13

POL$_{2r}$ that is as near as possible to $P(\lambda)$, i.e., to minimize the distance from $P(\lambda)$ to POL$_{2r}$,

$$\text{dist}_{\text{POL}_{2r}}\left(\sum_{i=0}^{d}\lambda^i P_i\right) = \min_{\sum_{i=0}^{d}\lambda^i S_i \in \text{POL}_{2r}}\left(\sum_{i=0}^{d}\|P_i - S_i\|_F^2\right)^{1/2}. \qquad (13)$$

This problem includes the well-known problem of finding the *distance to singularity*[1] when m is even

$$\text{dist}_{\text{POL}_{m-2}}(P) = \text{dist}_{sing}(P) := \min\{\|\Delta P\|_F : (P + \Delta P)(\lambda) \text{ is singular}\}. \qquad (14)$$

Using Theorem 3.1 we can substitute the minimization problem (13) with the following minimization problem (for brevity, we write $\mathcal{G}_{2r}$ instead of $\mathcal{G}_{d,2r,m}^{ss}$)

$$\text{dist}_{\text{POL}_{2r}}\left(\sum_{i=0}^{d}\lambda^i P_i\right) = \min_{\sum_{i=0}^{d}\lambda^i S_i \in \mathcal{G}_{2r}}\left(\sum_{i=0}^{d}\|P_i - S_i\|_F^2\right)^{1/2}$$

$$= \min_{\sum_{i=0}^{d}\lambda^i S_i \in \mathcal{G}_{2r}}\left\|\begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_d \end{bmatrix} - \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_d \end{bmatrix}\right\|_F = \min_{\sum_{i=0}^{d}\lambda^i S_i \in \mathcal{G}_{2r}}\left\|\begin{bmatrix} \text{vec}(P_0) \\ \text{vec}(P_1) \\ \vdots \\ \text{vec}(P_d) \end{bmatrix} - \begin{bmatrix} \text{vec}(S_0) \\ \text{vec}(S_1) \\ \vdots \\ \text{vec}(S_d) \end{bmatrix}\right\|_F, \qquad (15)$$

where $\text{vec}(\cdot)$ denotes the operator that stacks the columns of a matrix into one long column vector [32, Def. 4.2.9]. The reason for such reformulation is that the matrix polynomials $S(\lambda) \in \mathcal{G}_{2r}$ have a very special structure that we can use to develop an algorithm for solving the minimization problem. In the sequel, for simplicity, given a matrix polynomial $S(\lambda) = \sum_{i=0}^{d}\lambda^i S_i$, we define

$$\text{vec}(S(\lambda)) := \begin{bmatrix} \text{vec}(S_0) \\ \text{vec}(S_1) \\ \vdots \\ \text{vec}(S_d) \end{bmatrix}. \qquad (16)$$

From Theorem 3.1 every element in $\mathcal{G}_{2r}$ can be written as

$$S(\lambda) = U(\lambda)V(\lambda)^T - V(\lambda)U(\lambda)^T,$$

---

[1]Recall that any skew-symmetric matrix polynomial of odd size is automatically singular. Therefore, if m is odd the nearest singular skew-symmetric matrix polynomial to $P(\lambda)$ is just the nearest skew-symmetric matrix polynomial, which can be easily found by finding the nearest skew-symmetric matrix in the Frobenius norm to each of the matrix coefficients of $P(\lambda)$.

with $V(\lambda) \in \text{POL}^{m \times r}_{\lceil d/2 \rceil}$ and $U(\lambda) \in \text{POL}^{m \times r}_{\lfloor d/2 \rfloor}$, and, vice versa, every matrix polynomial of this form is in $\mathcal{G}_{2r}$. Equivalently,

$$
\begin{aligned}
\text{vec}(S(\lambda)) &= \text{vec}(U(\lambda)V(\lambda)^T - V(\lambda)U(\lambda)^T) \\
&= -\mathbf{M}(U(\lambda))\,\text{vec}(V(\lambda)) = \mathbf{M}(V(\lambda))\,\text{vec}(U(\lambda)),
\end{aligned} \tag{17}
$$

where an $(d+1)m^2 \times (d-t+1)mr$ matrix $\mathbf{M}(W(\lambda))$ is defined for given matrix polynomials $W(\lambda) = \sum_{i=0}^{t} \lambda^i W_i$ of size $m \times r$ and grade t, as follows

$$
\begin{aligned}
&\mathbf{M} = \mathbf{M}(W(\lambda)) \\
&= \begin{bmatrix}
(W_0 \otimes I) - (I \otimes W_0)N & 0 & 0 \\
\vdots & (W_0 \otimes I) - (I \otimes W_0)N & \vdots \\
(W_{t-1} \otimes I) - (I \otimes W_{t-1})N & \vdots & 0 \\
(W_t \otimes I) - (I \otimes W_t)N & (W_{t-1} \otimes I) - (I \otimes W_{t-1})N & (W_0 \otimes I) - (I \otimes W_0)N \\
0 & (W_t \otimes I) - (I \otimes W_t)N & \ddots & \vdots \\
\vdots & 0 & (W_{t-1} \otimes I) - (I \otimes W_{t-1})N \\
0 & \vdots & (W_t \otimes I) - (I \otimes W_t)N
\end{bmatrix},
\end{aligned} \tag{18}
$$

where N is the $mr \times mr$ permutation matrix that can "transpose" vectorized $m \times r$ matrices, i.e., $\text{vec}(X^T) = N\,\text{vec}(X)$ for any $m \times r$ matrix X [32, Th. 4.3.8], and the identity matrix has size $m \times m$. Now we can rewrite (15) using (17) in two different ways as follows:

$$
\text{dist}_{\text{POL}_{2r}}\left(\sum_{i=0}^{d} \lambda^i P_i\right) = \min_{\substack{V(\lambda) \in \text{POL}^{m \times r}_{\lceil d/2 \rceil}, \\ U(\lambda) \in \text{POL}^{m \times r}_{\lfloor d/2 \rfloor}}} \left\| \text{vec}(P(\lambda)) + \mathbf{M}(U(\lambda))\,\text{vec}(V(\lambda)) \right\|_F \tag{19}
$$

$$
= \min_{\substack{V(\lambda) \in \text{POL}^{m \times r}_{\lceil d/2 \rceil}, \\ U(\lambda) \in \text{POL}^{m \times r}_{\lfloor d/2 \rfloor}}} \left\| \text{vec}(P(\lambda)) - \mathbf{M}(V(\lambda))\,\text{vec}(U(\lambda)) \right\|_F. \tag{20}
$$

Following [5], we solve (19)-(20) using alternating least squares, see, e.g., [36, 39]. To be exact, we fix $U(\lambda)$ and use formulation (19) to find $V(\lambda)$ by solving

$$
\min_{\text{vec}(V(\lambda)) \in \mathbb{C}^{mr(\lceil d/2 \rceil + 1)}} \left\| \text{vec}(P(\lambda)) + \mathbf{M}(U(\lambda))\,\text{vec}(V(\lambda)) \right\|_F,
$$

then use the obtained $V(\lambda)$ in formulation (20) to find $U(\lambda)$ by solving

$$
\min_{\text{vec}(U(\lambda)) \in \mathbb{C}^{mr(\lfloor d/2 \rfloor + 1)}} \left\| \text{vec}(P(\lambda)) - \mathbf{M}(V(\lambda))\,\text{vec}(U(\lambda)) \right\|,
$$

and repeat the two steps with the computed $U(\lambda)$ until convergence. Moreover, solutions to these minimization problems are found simply by using

15

Matlab's backslash command, i.e., $\operatorname{vec}(V(\lambda)) := -\mathbf{M}(U(\lambda))\backslash \operatorname{vec}(P(\lambda))$ and $\operatorname{vec}(U(\lambda)) := \mathbf{M}(V(\lambda))\backslash \operatorname{vec}(P(\lambda))$.

Note that the sequence of distances $\rho_k := \left(\sum_{i=0}^{d} \|P_i - S_i^{(k)}\|_F^2\right)^{1/2}$, where $S^{(k)} := U^{(k)}(\lambda)V^{(k)}(\lambda)^T - V^{(k)}(\lambda)U^{(k)}(\lambda)^T$, generated by this alternating least squares approach is always convergent, since ρ_k is a monotonic non-increasing sequence bounded below by zero.

5. Skew-symmetric matrix pencils

This section is devoted to grade-one matrix polynomials, also known as matrix pencils. Here, we recall some definitions and results for skew-symmetric matrix pencils, in a similar way as was done for skew-symmetric matrix polynomials in Section 2. The reason for treating pencils separately is that our approach to solving the distance problems allows a significant improvement in this case; see Section 6 and Example 7.2.

We start by recalling the Kronecker-type canonical form of skew-symmetric matrix pencils under congruence. For each $k = 1, 2, \ldots$, and for each $\mu \in \mathbb{C}$ define the $k \times k$ matrices

$$J_k(\mu) := \begin{bmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{bmatrix}, \qquad I_k := \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

and for each $k = 0, 1, \ldots$, define the $k \times (k+1)$ matrices

$$F_k := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix}, \qquad G_k := \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}.$$

All non-specified entries of $J_k(\mu), I_k, F_k$, and G_k are zeros.

An $m \times m$ matrix pencil $A - \lambda B$ is called *congruent* to $C - \lambda D$ if there is a non-singular matrix S such that $S^T A S = C$ and $S^T B S = D$. We also define the *congruence orbit* of a skew-symmetric matrix pencil $A - \lambda B$ under the action of the group $GL_m(\mathbb{C})$ on the space of all skew-symmetric matrix pencils by congruence as follows:

$$O^c(A - \lambda B) = \{S^T(A - \lambda B)S \ : \ S \in GL_m(\mathbb{C})\}.$$

In Theorem 5.1 we recall the canonical form under congruence of skew-symmetric matrix pencils.

Theorem 5.1 ([40]). *Each skew-symmetric $m \times m$ matrix pencil $A - \lambda B$ is congruent to a direct sum, uniquely determined up to permutation of summands, of pencils of the form*

$$\mathcal{H}_h(\mu) := \begin{bmatrix} 0 & J_h(\mu) \\ -J_h(\mu)^T & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & I_h \\ -I_h & 0 \end{bmatrix}, \quad \mu \in \mathbb{C},$$

$$\mathcal{K}_k := \begin{bmatrix} 0 & I_k \\ -I_k & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & J_k(0) \\ -J_k(0)^T & 0 \end{bmatrix},$$

$$\mathcal{M}_m := \begin{bmatrix} 0 & F_m \\ -F_m^T & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & G_m \\ -G_m^T & 0 \end{bmatrix}.$$

The blocks $\mathcal{H}_h(\mu)$ and $\mathcal{K}_k$ correspond to the finite and infinite eigenvalues, respectively, and all together form the *regular part* of $A - \lambda B$. The blocks $\mathcal{M}_m$ correspond to the right and left minimal indices, which are equal in the case of skew-symmetric matrix pencils, and form the *singular part* of $A - \lambda B$. Using this canonical form, we can now formulate Theorem 5.2, which is a special case of Theorem 2.2. Define $\mathrm{PEN}_{2r,m}^{ss}$ as the set of $m \times m$ skew-symmetric matrix pencils of rank $\leq 2r$, i.e., $\mathrm{PEN}_{2r,m}^{ss} = \mathrm{POL}_{1,2r,m}^{ss}$.

Theorem 5.2 ([16]). *Let m and r be integers such that $2 \leq 2r \leq m-1$. Then $\mathrm{PEN}_{2r,m}^{ss} = \overline{\mathrm{O}^c(\mathcal{W}_{2r}^m)}$, where*

$$\mathcal{W}_{2r}^m = \mathrm{diag}(\underbrace{\mathcal{M}_{\alpha+1}, \ldots, \mathcal{M}_{\alpha+1}}_{t}, \underbrace{\mathcal{M}_\alpha, \ldots, \mathcal{M}_\alpha}_{m-2r-t}),$$

with $\alpha = \lfloor \frac{r}{m-2r} \rfloor$ and $t = r \mod (m-2r)$.

Similarly, the following theorem was proven earlier in [11]. Note that, with the notation of Theorem 3.1, an immediate consequence of Theorem 5.3 is that $\mathrm{PEN}_{2r,m}^{ss} = \mathcal{G}_{1,2r,m}^{ss}$, *without involving the closure of $\mathcal{G}_{1,2r,m}^{ss}$*. Thus, the result for skew-symmetric pencils is stronger than for skew-symmetric matrix polynomials of grade larger than 1. It is possible to prove that the use of the closure is necessary in Theorem 3.1.

Theorem 5.3 (Rank-1 decomposition of skew-symmetric matrix pencils from [11]). *Let $S_0 - \lambda S_1 \in \mathrm{PEN}_{2r,m}^{ss}$. Then*

$$S_0 - \lambda S_1 = \sum_{i=1}^{r} u_i v_i^T - v_i u_i^T,$$

for some vectors $u_1, \ldots, u_r \in \mathbb{C}^m$ and polynomial vectors $v_1, \ldots, v_r \in \mathbb{C}[\lambda]^m$ of degree ≤ 1.

17

The decomposition in Theorem 5.3 is achieved by decomposing each block in the canonical form in Theorem 5.1 into a sum of rank-1 matrix pencils (a tool which is not available for skew-symmetric matrix polynomials of grade larger than 1). Also note that for a pencil $A - \lambda B$ *not* in the canonical form, this decomposition is still possible. Let S be a matrix such that $S^{-1}(A - \lambda B)S^{-T}$ is in the canonical form. Then, a rank-1 decomposition of $A - \lambda B$ is done by taking $v_i = S\tilde{v}_i$ and $u_i = S\tilde{u}_i$, where $\tilde{v}_i$ and $\tilde{u}_i$ are the vectors in the decomposition of the matrix pencil in the canonical form.

Note also that since skew-symmetric pencils always have even rank, the rank-1 pencils in Theorem 5.3 come in pairs ($u_i v_i^T$ and $v_i u_i^T$) that form skew-symmetric pencils ($u_i v_i^T - v_i u_i^T$) of rank 2, when the rank of $S_0 - \lambda S_1$ is exactly $2r$.

6. Distance to rank-deficient skew-symmetric matrix pencils

The fact that the vectors u_i in Theorem 5.3 are constant vectors allows us to improve considerably the algorithm described in Section 4 in the pencil case. Further improvements are achieved if the input pencil is skew-symmetric. Therefore, in this section we specifically deal with skew-symmetric input pencils. Given an $m \times m$ complex skew-symmetric matrix pencil $A - \lambda B$, we find a nearby skew-symmetric matrix pencil $S_0 - \lambda S_1$ of rank $\leq 2r$, with $2 \leq 2r < m$. By Theorem 5.3, we can assume that $S_0 - \lambda S_1$ is in the form

$$S_0 = UV_0^T - V_0 U^T \ \text{ and } \ S_1 = UV_1^T - V_1 U^T,$$

where $U = [u_1, \ldots, u_r]$, and $V_0 - \lambda V_1 = [v_1, \ldots, v_r]$. Denoting for simplicity $\mathrm{PEN}_{2r,m}^{ss}$ just by PEN_{2r}, our distance problem is reformulated in the pencil case as follows

$$
\begin{aligned}
\mathrm{dist}_{\mathrm{PEN}_{2r}}(A - \lambda B) &= \min_{S_0 - \lambda S_1 \in \mathrm{PEN}_{2r}} \left(\|A - S_0\|_F^2 + \|B - S_1\|_F^2 \right)^{1/2} \\
&= \min_{U, V_0, V_1 \in \mathbb{C}^{m \times r}} \left\| \begin{bmatrix} A \\ B \end{bmatrix} - \begin{bmatrix} UV_0^T - V_0 U^T \\ UV_1^T - V_1 U^T \end{bmatrix} \right\|_F .
\end{aligned}
\tag{21}
$$

Similarly to Section 4, we solve (21) using a method of alternating directions. As before, we fix U to find V_0 and V_1, i.e., we solve

$$
\min_{V_0, V_1 \in \mathbb{C}^{m \times r}} \left\| \begin{bmatrix} A \\ B \end{bmatrix} - \begin{bmatrix} UV_0^T - V_0 U^T \\ UV_1^T - V_1 U^T \end{bmatrix} \right\|_F ,
$$

but now this problem is equivalent to solving the following two minimization problems independently

$$
\min_{V_0 \in \mathbb{C}^{m \times r}} \left\| A - (UV_0^T - V_0 U^T) \right\|_F \quad \text{ and } \quad \min_{V_1 \in \mathbb{C}^{m \times r}} \left\| B - (UV_1^T - V_1 U^T) \right\|_F . \tag{22}
$$

Then, we use the obtained V_0 and V_1 to find U, i.e., we solve

$$\min_{U \in \mathbb{C}^{m \times r}} \left\| \begin{bmatrix} A \\ B \end{bmatrix} - \begin{bmatrix} UV_0^T - V_0 U^T \\ UV_1^T - V_1 U^T \end{bmatrix} \right\|_F. \tag{23}$$

Now we explain how to solve each of the problems in (22).

We assume for this purpose that U has full rank r, which is the generic situation in the algorithm, but we will make some comments at the end on how to proceed when U is rank deficient. Note that, taking the SVD of $U = RST^H$, with $R \in \mathbb{C}^{m \times m}$ and $T \in \mathbb{C}^{r \times r}$ unitary (orthogonal in the real case), yields the following minimization problem equivalent to the first one in (22)

$$\min_{V_0 \in \mathbb{C}^{m \times r}} \left\| R^H A \overline{R} - (S(R^H V_0 \overline{T})^T - (R^H V_0 \overline{T})S^T) \right\|_F. \tag{24}$$

An analogous problem can be obtained for the second minimization problem for V_1 in (22) with $R^H B \overline{R}$ instead of $R^H A \overline{R}$. Let $X := R^H V_0 \overline{T}$ and $E := R^H A \overline{R}$, and note that E is skew-symmetric. Denote by S_1 the $r \times r$ diagonal block of S and by X_1 and X_2 the r first and $m - r$ last rows of X, that is,

$$S = \begin{bmatrix} S_1 \\ 0 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

Note that S_1 is invertible if U has rank r. By making the corresponding partition of $E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}$, we obtain from (24) the following equivalent minimization problem

$$\min_{X_1 \in \mathbb{C}^{r \times r}, X_2 \in \mathbb{C}^{(m-r) \times r}} \left\| \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} - \begin{bmatrix} S_1 X_1^T - X_1 S_1^T & S_1 X_2^T \\ -X_2 S_1^T & 0 \end{bmatrix} \right\|_F. \tag{25}$$

The problems for X_1 and X_2 decouple in (25) and can be solved very efficiently by computing the exact solutions of two linear systems as we explain in the following. Since $E_{21} = -E_{12}^T$, the two off-diagonal blocks that contain X_2 are equivalent and $X_2 = E_{12}^T S_1^{-1}$, which taking into account that S_1 is diagonal can be computed very efficiently. As for X_1, we note that since S_1 is diagonal and E_{11} is skew-symmetric, the (i,j)-element in $S_1 X_1^T - X_1 S_1^T = E_{11}$ is $s_{ii} x_{ji} - s_{jj} x_{ij} = e_{ij}$, while the (j,i)-element is $s_{jj} x_{ij} - s_{ii} x_{ji} = e_{ji} = -e_{ij}$. So, both lead to the same equation for x_{ij} and x_{ji}, when $i \neq j$, which can be chosen in infinitely many ways (as well as the diagonal elements of X_1 since there are no restriction over them). In particular, we can choose X_1 to be skew-symmetric. Then, the equations simplify to $x_{ji} = \frac{e_{ij}}{s_{ii}+s_{jj}}$. When X is calculated, V_0 is given by

$$V_0 = R \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} T^T.$$

19

In the same fashion, we obtain V_1. If the rank of U is less than r, then some diagonal entries of S_1 are zero, but (25) is still valid. In that scenario, we can use the Moore-Penrose inverse, $S_1^\dagger$, of S_1, which is a diagonal $r \times r$ matrix with $(S_1^\dagger)_{ii} = 1/s_{ii}$, when $s_{ii} \neq 0$, and zeros elsewhere. Then, $X_2 = E_{12}^T S_1^\dagger$ is a least squares solution with minimum norm for X_2. A least squares minimal norm skew-symmetric solution for X_1 is obtained taking $x_{ji} = 0$, when $s_{ii} = s_{jj} = 0$, and proceeding as in the full rank case for the other entries.

Now, fixing V_0 and V_1, we take another step of the alternating directions. We solve for U in exactly the same way as in the case of polynomials of any degree derived in Section 4, i.e.,

$$\text{vec}(U) := \begin{bmatrix} (V_0 \otimes I_m) - (I_m \otimes V_0)\, N \\ (V_1 \otimes I_m) - (I_m \otimes V_1)\, N \end{bmatrix} \setminus \begin{bmatrix} \text{vec}(A) \\ \text{vec}(B) \end{bmatrix}.$$

7. Numerical experiments

In this section, we compare the developed algorithms with one another and, when possible, with other available algorithms. Specifically, we compute the distance to singularity for matrix polynomials using three methods (including ours), and we compute the distance to the set of matrix pencils of a specified bounded rank using two of our algorithms. The reason for this distinction is that the other available algorithms can compute only the distance to singularity, and not the distance to the nearest matrix polynomial of rank $2r$ (for arbitrary r). The experiments were run using MATLAB (R2024b Update 4) on MacBook Air (Apple M3, 16GB). The code is available on GitHub.[2]

Our algorithm uses Generic Eigenstructures and Alternating direction method for Rank $2r$ approximation of Skew-symmetric matrix polynomials, thus we abbreviate it to GEARS. We also use GEARS-SVD for the version with svd-decomposition for matrix pencils.

Example 7.1. *We calculate a distance to singularity for skew-symmetric matrix polynomials using three different algorithms. Namely Riemman-Oracle (RO) [27] and Matrix-Valued Function (MVF) [25, 26], as well as GEARS, the algorithm developed in this paper. We run each algorithm on 72 random skew-symmetric matrix polynomials, i.e., 4 random matrix polynomials for each combination of the size (4×4, 6×6, or 8×8), degree (1, 2, or 3), and the underlying field (real or complex). The results are summarized in Figures 1,2, and 3. The graphs in Figure 1 show that all algorithms find a skew-symmetric matrix polynomial at essentially the same distance from a*

---

[2]`https://github.com/rakeljh/dist_sing_skew_pol`

given skew-symmetric matrix polynomial. Nevertheless, our algorithm produces the answers that are "the most singular"[3], see Figure 2 (recall that the output of GEARS is singular by construction). Moreover, Figure 3 shows that GEARS is also the fastest out of the tested algorithms.
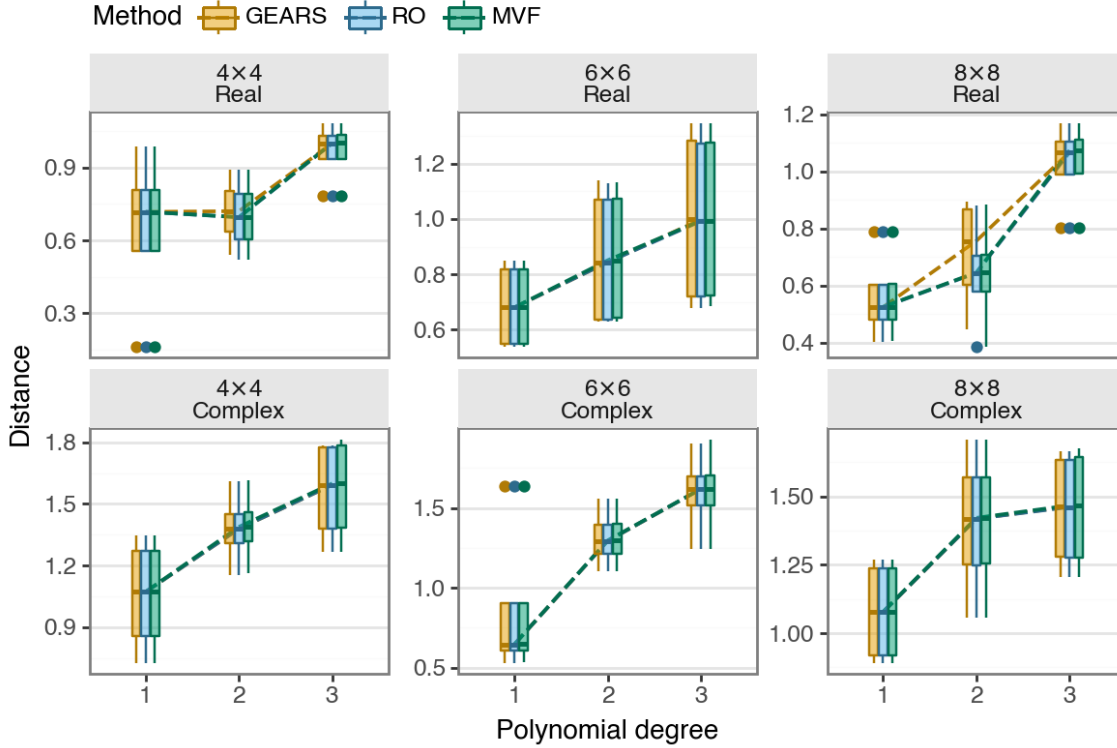

Figure 1: Comparison of the computed minimal distances to singularity for random skew-symmetric matrix polynomials of given size, degree, and with the entries in the underlying field (real or complex). All the algorithms find a skew-symmetric matrix polynomial at essentially the same distance from a given skew-symmetric matrix polynomial.

In addition, to investigate how the computational cost of GEARS scales with the input size, we extend our experiments to random skew-symmetric matrix polynomials of degree 2 with sizes ranging from 20×20 up to 50×50, over both the real and complex fields, Figures 4 and 5. We carefully controlled the quality of the computed output, roughly matching the accuracy reported in

---

[3]We note that the MVF algorithm requires an assumption [26, Assumption 4.8]. This assumption used in the outer iteration of the method is that the smallest singular value of $(F + \Delta F)(\mu_j)$ is simple and non-zero for $j = 1, \ldots, m$. In the case of skew-symmetric matrix polynomials, the singular values come in pairs and hence the smallest singular value is only simple if n is odd, in which case it is zero. This may affect the performance of this method when used to find the nearest singular skew-symmetric matrix pencil.
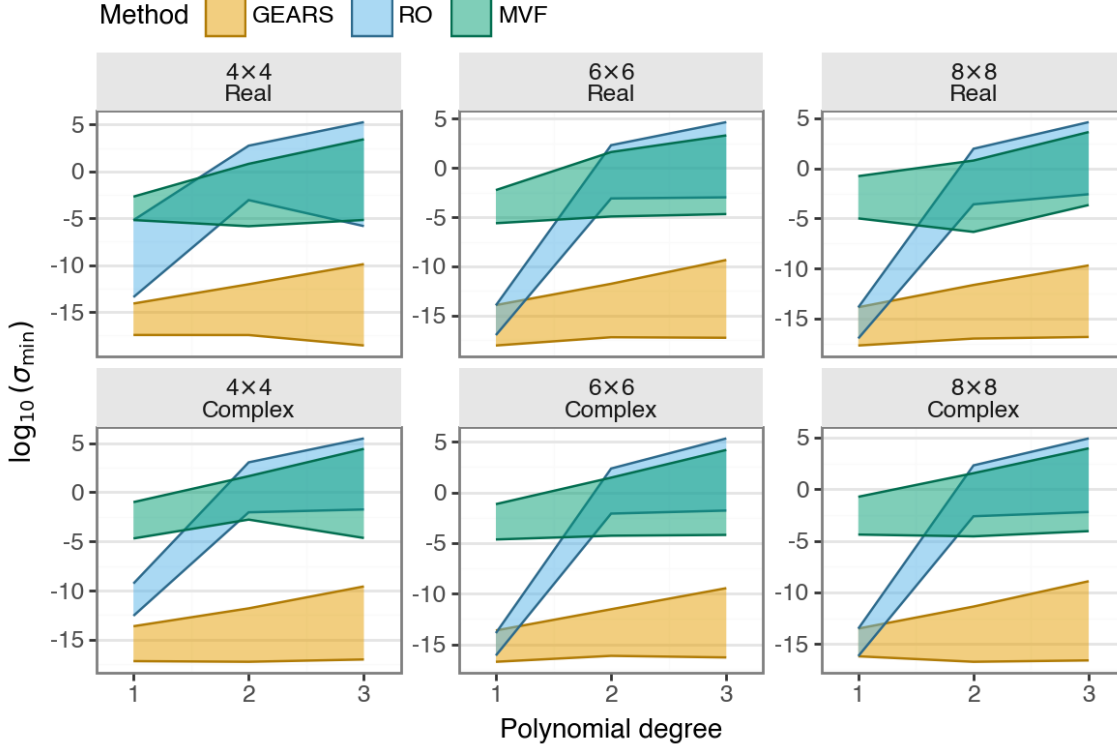
Figure 2: Comparison of the "singularity" of the output for random skew-symmetric matrix polynomials of given size, degree, and the entries in the underlying field (real or complex). To be exact, we follow [26] and plot smallest singular value $\sigma_{\min}$ of the constructed matrix polynomial evaluated at the points $x_j + \sqrt{-1}y_j$ given by $[X, Y] = meshgrid(-1000 : 40 : 1000)$. In all the performed experiments GEARS produces the outputs with significantly lower singular values.

[27] for 30×30 complex unstructured matrix polynomials of degree 2.[4] Note that [27] reports the runtime of around 300 seconds for 30×30 matrices, while our approach requires approximately 200 seconds for comparable output quality. In the real case, the algorithm is typically about one order of magnitude faster than in the complex setting when asked to maintain the same quality of the output.

In Example 7.1, we use the code for polynomials, which works for any degree, including degree one. Nevertheless, Section 5 and Theorem 5.3 show that the parametrization for matrix polynomials in Theorem 3.1 simplifies significantly when we are dealing with matrix pencils. This has enabled us

---

[4]To be exact we terminate the algorithm when the change in distance between consecutive iterations is less than $10^{-4} \cdot \|P\|_F$ for complex matrix polynomials and less than $10^{-3} \cdot \|P\|_F$ for real matrix polynomials.
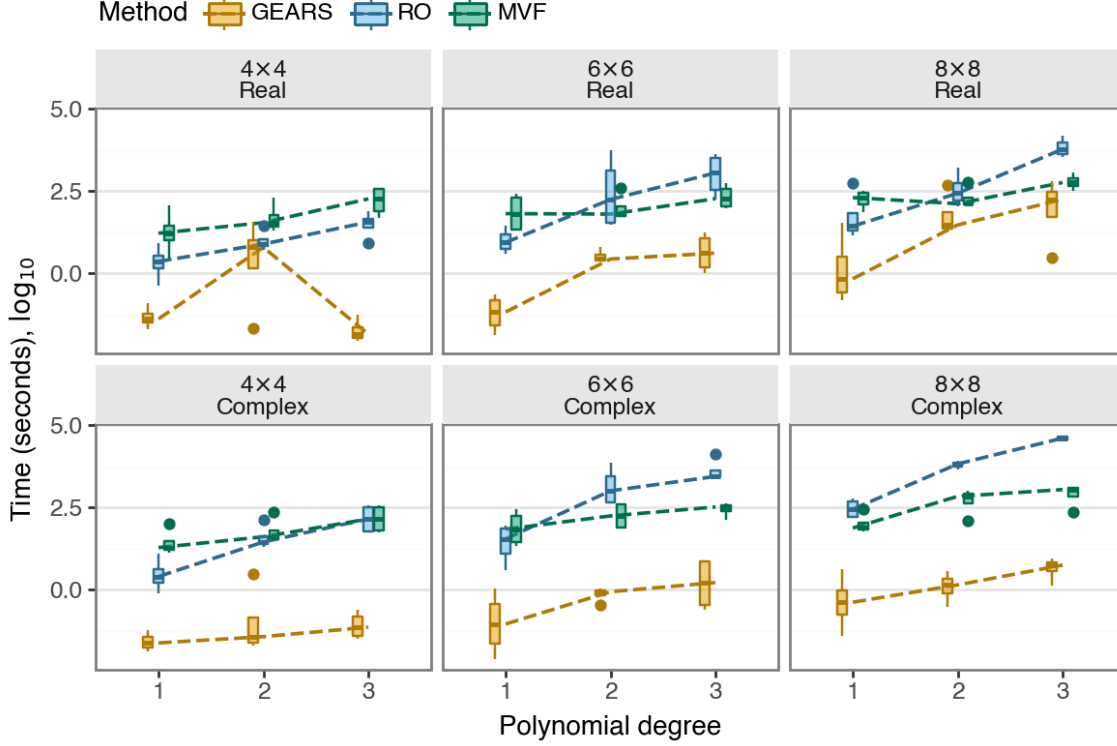
Figure 3: Comparison of the speed for random skew-symmetric matrix polynomials of given size, degree, and the entries in the underlying field (real or complex). In all the performed experiments GEARS is the fastest.

to develop the approach explained in Section 6, which for one of the alternating directions step solves matrix equations directly, without reformulating it using the Kronecker product as a large system of linear equations. Figures 6 and 7 illustrate the resulting time savings while maintaining accuracy for the skew-symmetric pencils generated in Example 7.2.

Example 7.2. *In this example we compare performance of the algorithms GEARS and GEARS-SVD developed in Sections 4 and 6, respectively, on the matrix pencil input. We run each algorithm on 50 random skew-symmetric matrix pencils, i.e., 10 random complex skew-symmetric matrix pencils for each size (from 6×6 to 30×30), and compute the result for ranks $n - 2$ for all sizes, and $\lfloor n/4 \rfloor$ for sizes 12×12 and up. The results are summarized in Figures 6 and 7.*

As seen in Figure 7, the computation time was improved when using GEARS-SVD compared to GEARS. The experiments were done with a maximum number of iterations, and since GEARS-SVD converges faster than
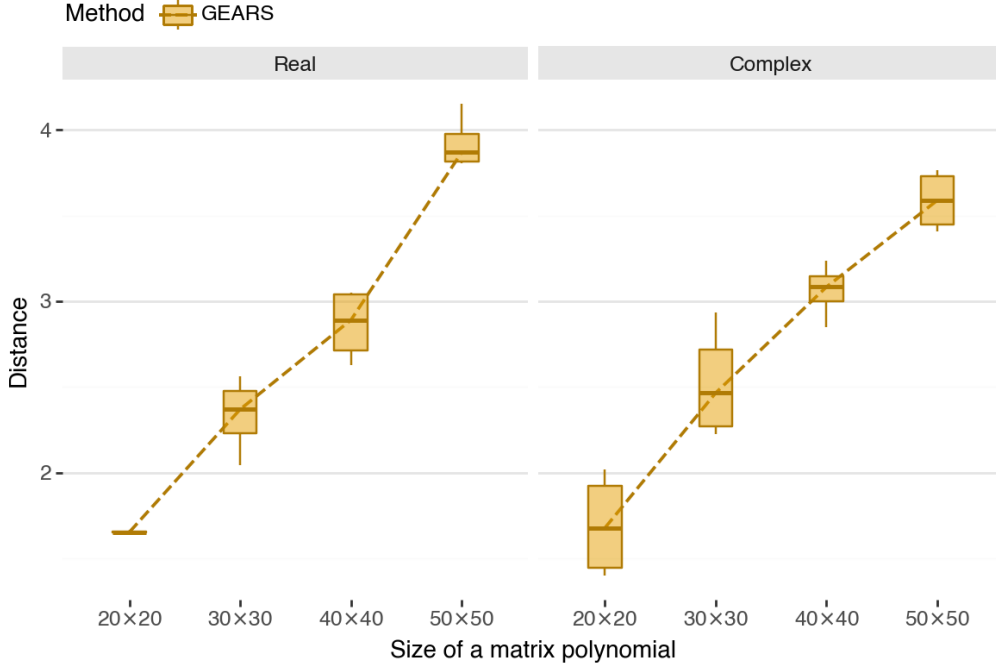
Figure 4: GEARS algorithm for computing the minimal distance to singularity for random skew-symmetric matrix polynomials of sizes 20×20, 30×30, 40×40, and 50×50, of degree 2, with entries over the real or complex field (4 examples for each setting).

GEARS, the distance results were also slightly better for large matrix pencil sizes when using GEARS-SVD. This is seen in Figure 6.

8. Conclusions and future work

We present an algorithm that approximates a given matrix polynomial by a skew-symmetric matrix polynomial of a specified rank. Our algorithm works for any rank $2r$ such that $2 \leq 2r \leq m - 1$, where $m \times m$ is the matrix size, and produces a skew-symmetric matrix polynomial that is of rank $2r$ up to rounding errors. In the tested settings, it is significantly faster than the other available algorithms, which in addition are only valid for the particular case when $2r$ in the largest even number smaller than m.

Extensions of this result to other classes of structured matrix polynomials, such as symmetric, symmetric/skew-symmetric, or palindromic, are part of our future work.
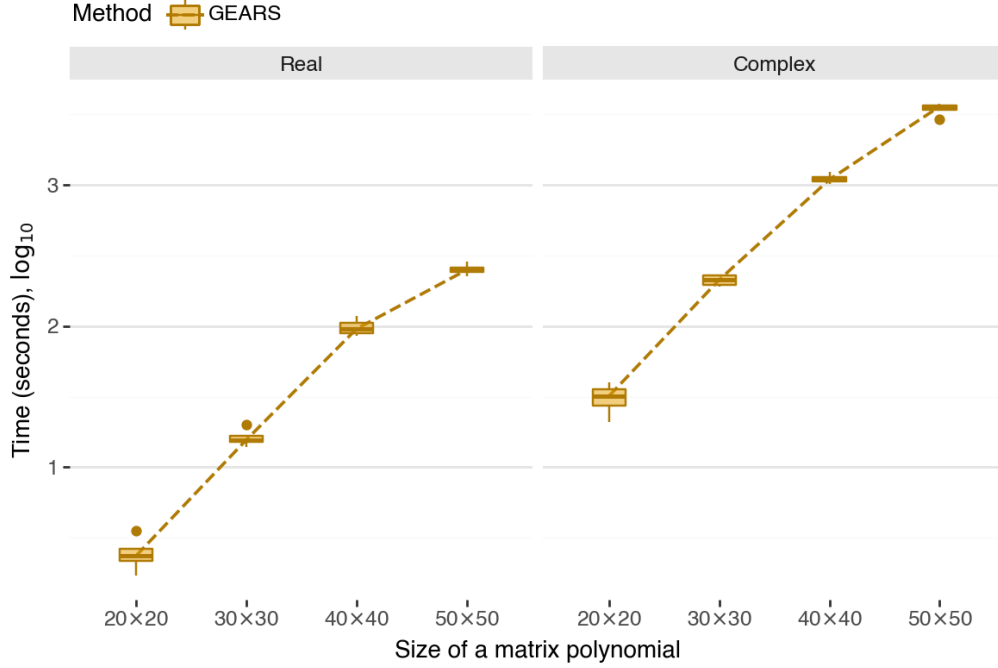
Figure 5: Illustration of the time scaling of the GEARS algorithm as the input size increases. The plot shows the runtime for random skew-symmetric matrix polynomials of given sizes and degree 2, with entries over the real or complex field (4 examples for each setting).

References

[1] L. Batzke. Generic rank-two perturbations of structured regular matrix pencils. *Oper. Matrices*, 10(1):83–112, 2016.

[2] R. Byers, C. He, and V. Mehrmann. Where is the nearest non-regular pencil? *Linear Algebra Appl.*, 285(1):81–105, 1998.

[3] B. Das and S. Bora. Nearest rank deficient matrix polynomials. *Linear Algebra Appl.*, 674:304–350, 2023.

[4] S. Das and A. Dmytryshyn. Minimal degenerations of orbits of skew-symmetric matrix pencils. *arXiv:2510.25482*, 2025.

[5] S. Das, A. Dmytryshyn, F. M. Dopico, and M. Gulliksson. Generic eigenstructures and alternating direction method for fixed rank approximation of matrix polynomials. *In preparation*, 2026.
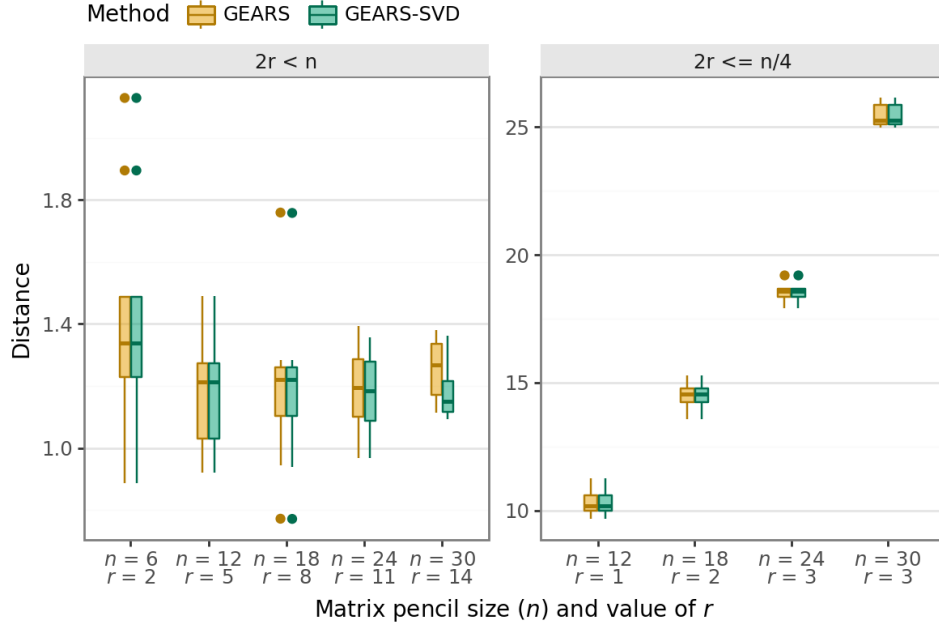
Figure 6: Comparison of the computed minimal distances for random complex skew-symmetric matrix pencils of size $n \times n$ to the set of skew-symmetric pencils with rank $\leq 2r$. The only noticeable difference is for the largest sizes.

[6] F. De Terán. A geometric description of the sets of palindromic and alternating matrix pencils with bounded rank. *SIAM J. Matrix Anal. Appl.*, 39(3):1116–1134, 2018.

[7] F. De Terán, A. Dmytryshyn, and F. M. Dopico. Generic symmetric matrix pencils with bounded rank. *J. Spectr. Theory*, 10(3):905–926, 2020.

[8] F. De Terán, A. Dmytryshyn, and F. M. Dopico. Generic symmetric matrix polynomials with bounded rank and fixed odd grade. *SIAM J. Matrix Anal. Appl.*, 41(3):1033–1058, 2020.

[9] F. De Terán, A. Dmytryshyn, and F. M. Dopico. Generic eigenstructures of Hermitian pencils. *SIAM J. Matrix Anal. Appl.*, 45(1):260–283, 2024.

[10] F. De Terán and F. M. Dopico. A note on generic Kronecker orbits of matrix pencils with fixed rank. *SIAM J. Matrix Anal. Appl.*, 30(2):491–496, 2008.

[11] F. De Terán, C. Mehl, and V. Mehrmann. Low-rank perturbation of regular matrix pencils with symmetry structures. *Found. Comput. Math.*, 22(1):257–311, 2022.

[12] F. De Terán, A. Dmytryshyn, and F. M. Dopico. Even grade generic skew-symmetric matrix polynomials with bounded rank. *Linear Algebra Appl.*, 702:218–239, 2024.

[13] A. Dmytryshyn. Miniversal deformations of pairs of skew-symmetric matrices under congruence. *Linear Algebra Appl.*, 506:506–534, 2016.
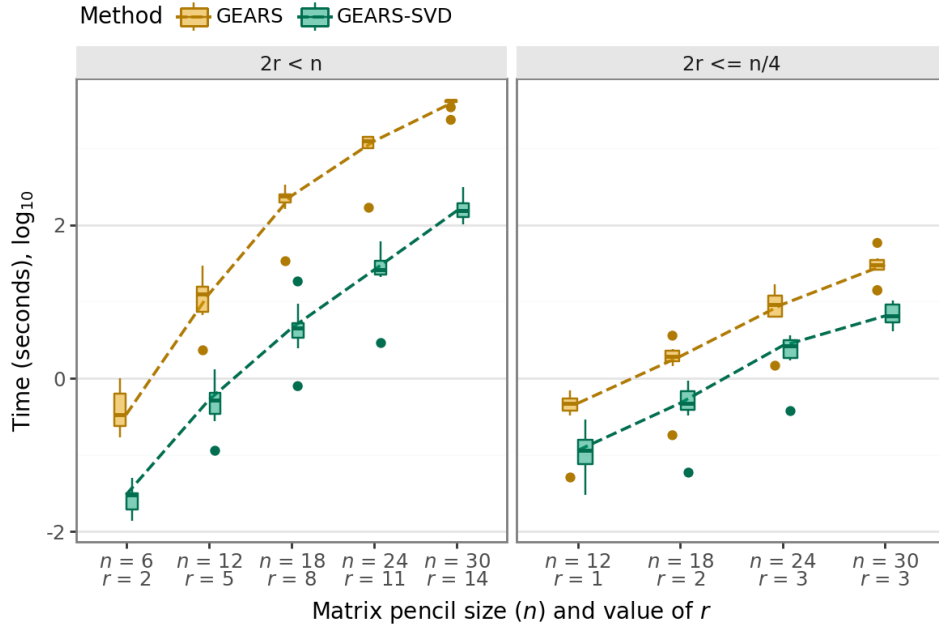
Figure 7: Comparison of the computation time for determining the distance to the set of skew-symmetric matrix pencils with rank $\leq 2r$, for random complex skew-symmetric matrix pencils of size $n \times n$. GEARS-SVD outperforms GEARS.

[14] A. Dmytryshyn. Structure preserving stratification of skew-symmetric matrix polynomials. *Linear Algebra Appl.*, 532:266–286, 2017.

[15] A. Dmytryshyn and F. M. Dopico. Generic complete eigenstructures for sets of matrix polynomials with bounded rank and degree. *Linear Algebra Appl.*, 535:213–230, 2017.

[16] A. Dmytryshyn and F. M. Dopico. Generic skew-symmetric matrix polynomials with fixed rank and fixed odd grade. *Linear Algebra Appl.*, 536:1–18, 2018.

[17] A. Dmytryshyn, F. M. Dopico, and P. Van Dooren. Minimal rank factorizations of polynomial matrices. *Linear Algebra Appl.*, 721:736–774, 2025.

[18] A. Dmytryshyn, F. M. Dopico, and P. Van Dooren. Minimal rank factorizations of structured polynomial matrices. *In preparation*, 2026.

[19] A. Dmytryshyn, S. Johansson, B. Kågström, and P. Van Dooren. Geometry of matrix polynomial spaces. *Found. Comput. Math.*, 20:423–450, 2020.

[20] A. Dmytryshyn and B. Kågström. Orbit closure hierarchies of skew-symmetric matrix pencils. *SIAM J. Matrix Anal. Appl.*, 35(4):1429–1443, 2014.

[21] A. Dmytryshyn, B. Kågström, and V. Sergeichuk. Skew-symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations. *Linear Algebra Appl.*, 438(8):3375–3396, 2013.

[22] F. M. Dopico, V. Noferini, and L. Nyman. A Riemannian optimization method to compute the nearest singular pencil. *SIAM J. Matrix Anal. Appl.*, 45(4):2007–2038, 2024.

[23] G. D. Forney. Minimal bases of rational vector spaces with applications to multivariable linear systems. *SIAM J. Control Optim.*, 13(3):493–520, 1975.

[24] M. Giesbrecht, J. Haraldson, and G. Labahn. Computing the nearest rank-deficient matrix polynomial. In *Proceedings of the 2017 ACM International Symposium on Symbolic and Algebraic Computation*, ISSAC '17, pages 181–188, New York, NY, USA, 2017. ACM.

[25] M. Gnazzo and N. Guglielmi. Approximating the closest structured singular matrix polynomial. *Linear and Multilinear Algebra*, 73(9):2222–2250, 2025.

[26] M. Gnazzo and N. Guglielmi. On the numerical approximation of the distance to singularity for matrix-valued functions. *SIAM J. Matrix Anal. Appl.*, 46(2):1484–1517, 2025.

[27] M. Gnazzo, V. Noferini, L. Nyman, and F. Poloni. Riemann-oracle: A general-purpose Riemannian optimizer to solve nearness problems in matrix theory. *Found. Comput. Math.*, pages 1–45, 2025.

[28] N. Guglielmi and C. Lubich. Matrix nearness problems and eigenvalue optimization. *arXiv preprint arXiv:2503.14750*, 2025.

[29] N. Guglielmi, C. Lubich, and V. Mehrmann. On the nearest singular matrix pencil. *SIAM J. Matrix Anal. Appl.*, 38:776–806, 2017.

[30] N. J. Higham. Matrix nearness problems and applications. In *Applications of matrix theory (Bradford, 1988)*, volume 22 of *Inst. Math. Appl. Conf. Ser. New Ser.*, pages 1–27. Oxford Univ. Press, New York, 1989.

[31] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.

[32] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[33] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[34] T. Kailath. *Linear Systems*. Prentice Hall, New Jersey, 1980.

[35] D. Kressner and M. Voigt. Distance problems for linear dynamical systems. *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory: Festschrift in Honor of Volker Mehrmann*, pages 559–583, 2015.

[36] Y. Li, Y. Liang, and A. Risteski. Recovery guarantee of weighted low-rank approximation via alternating minimization. In M. F. Balcan and K. Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2358–2367, New York, New York, USA, 20–22 Jun 2016. PMLR.

[37] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Skew-symmetric matrix polynomials and their Smith forms. *Linear Algebra Appl.*, 438(12):4625–4653, 2013.

[38] L. Rodman. Comparison of congruences and strict equivalences for real, complex, and quaternionic matrix pencils with symmetries. *Electron. J. Linear Algebra*, 16:248–283, 2007.

[39] N. Srebro and T. Jaakkola. Weighted low-rank approximations. In *Proceedings of the 20th international conference on machine learning (ICML-03)*, pages 720–727, 2003.

[40] R. C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.