Electronic Transactions on Numerical Analysis.
Volume 45, pp. 33–57, 2016.
Copyright © 2016, Kent State University.
ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

DOUBLE ANGLE THEOREMS FOR DEFINITE MATRIX PAIRS*

LUKA GRUBIŠIĆ†, SUZANA MIODRAGOVIĆ‡, AND NINOSLAV TRUHAR‡

Abstract. In this paper we present new double angle theorems for the rotation of the eigenspaces of Hermitian matrix pairs (H, M), where H is a non-singular matrix which can be factorized as $H = GJG^*$, $J = \mathrm{diag}(\pm 1)$, and M is non-singular. The rotation of the eigenspaces is measured in the matrix-dependent scalar product, and the bounds belong to relative perturbation theory. The quality of the new bounds are illustrated in the numerical examples.

Key words. matrix pairs, perturbation of eigenvectors, $\sin 2\Theta$ theorems

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. Controlling the size of the rotation of an invariant subspace of a matrix or operator A under a perturbation V is one of the fundamental problems in operator theory or matrix analysis; see [1, 6, 18, 19, 20] for results in operator theory and [4, 7, 14, 16, 23, 25] for recent results in the context of matrix analysis.

Among the most prominent results in this field of research is the series of papers by C. Davis on the rotation of invariant subspaces under the influence of a perturbation; see, e.g., [6] and the references therein. A sequence of three papers culminated with the cornerstone paper of Davis and Kahan [6], which had both fundamental importance in operator theory (scattering theory in mathematical physics) as well as ramifications in matrix analysis. In particular, it influenced the development of mathematical software for highly accurate solutions of singular value and eigenvalue problems [15].

The main objective in these studies was to obtain a bound of a trigonometric function of the angle operator associated with spectral subspaces of the unperturbed and perturbed operators, respectively. In what follows, we use $\mathrm{Spec}(H)$ to denote the spectrum of a matrix H. Let us consider matrices H and $\widetilde{H} = H + \delta H$, and let the claims $\mathrm{Spec}(H) = \mathfrak{L}_1 \cup \mathfrak{L}_2$, $\mathfrak{L}_1 \cap \mathfrak{L}_2 = \emptyset$, and $\mathrm{Spec}(\widetilde{H}) = \widetilde{\mathfrak{L}_1} \cup \widetilde{\mathfrak{L}_2}$, $\widetilde{\mathfrak{L}_1} \cap \widetilde{\mathfrak{L}_2} = \emptyset$ hold for the spectra of H and $\widetilde{H}$. We use $E(\mathfrak{L}_1)$ and $\widetilde{E}(\widetilde{\mathfrak{L}}_1)$ to denote the spectral projection of H and $\widetilde{H}$ associated to the sets $\mathfrak{L}_1$ and $\widetilde{\mathfrak{L}}_1$, respectively. We define the angle operator Θ by spectral calculus as

$$\Theta := \arcsin\big(E(\mathfrak{L}_1) - \widetilde{E}(\widetilde{\mathfrak{L}_1})\big) \tag{1.1}$$

since orthogonal projections are Hermitian (self-adjoint) and idempotent matrices/operators. The eigenvalues of the Hermitian matrix Θ are called the canonical angles. For easier reference, we also introduce the notation for the associated subspaces $\mathfrak{I} = \mathrm{Ran}(E(\mathfrak{L}_1))$ and $\widetilde{\mathfrak{I}} = \mathrm{Ran}(\widetilde{E}(\widetilde{\mathfrak{L}_1}))$. We freely associate the angle operator with either the subspaces or the corresponding projections whatever is appropriate in a given context. We use $\mathrm{Ran}(\cdot)$ to denote the range of a matrix. A generic estimate can be formulated as the following $\sin 2\Theta$-bound, which is taken from a recent paper by Albeverio and Motovilov [1]:

$$\|\sin 2\Theta\| \leq \pi \frac{\|H - \widetilde{H}\|}{d}\ , \tag{1.2}$$

*Received October 4, 2014. Accepted February 10, 2016. Published online on March 18, 2016. Recommended by F. Dopico.

†Department of Mathematics, University of Zagreb, Bijenička ulica, HR-10000 Zagreb, Croatia (luka.grubisic@math.hr).

‡Department of Mathematics, J. J. Strossmayer University of Osijek, Trg Ljudevita Gaja 6, HR-31000 Osijek, Croatia ({ssusic,ntruhar}@mathos.hr)

ETNA
Kent State University
http://etna.math.kent.edu

where $d = \mathrm{Dist}(\mathfrak{L}_1, \mathfrak{L}_2)$ is the Hausdorff distance. In the case $d = 0$, we take $1/d = \infty$ and the bound trivially holds. Following [1, 2, 20], we call such estimates a priori since only the separation d between the wanted and unwanted components of the spectrum of the unperturbed matrix H is appearing in the estimates. Alternatively, we can obtain estimates which feature only a posteriori distances (those in the spectrum of $\widetilde{H}$) by reversing the roles of the matrices.

Our main contribution is to establish such an estimate for the definitizable generalized eigenvalue problem in a matrix-dependent scalar product. For a discussion of the geometry of an Euclidean space in a matrix-dependent scalar product, see [13].

A matrix pair (H, M) is called definitizable if there exist scalars α and β such that $\alpha H + \beta M$ is a positive definite matrix. Given definite Hermitian matrix pairs (H, M) and $(\widetilde{H}, \widetilde{M}) = (H + \delta H, M + \delta M)$, where H, $\widetilde{H}$, M, and $\widetilde{M}$ are non-singular matrices and their spectral subspaces $\mathfrak{I}$ and $\widetilde{\mathfrak{I}}$ are of the same dimension, we are interested in providing estimates for the size of the rotation which moves $\mathfrak{I}$ to $\widetilde{\mathfrak{I}}$. We shall do this by bounding the sines of the double canonical angles between $\mathfrak{I}$ and $\widetilde{\mathfrak{I}}$ in the scalar product $(x, y)_M = x^* M y$. Note that (1.1) depends on the scalar product since it is required that a projection be self-adjoint in the chosen Hilbert space structure. We denote this by adding the subscript M to the angle operator and write $\Theta_M(\mathfrak{I}, \widetilde{\mathfrak{I}})$.

It is obvious that the double angle theorems do not directly bound the difference between the old invariant subspace $\mathfrak{I}$ and the new one $\widetilde{\mathfrak{I}}$. One possibility to interpret them is using spectral calculus as is shown in (1.2). This is, however, technically quite involved when also allowing for perturbations of a matrix-dependent scalar product. Alternatively, recall that there is a direct geometric interpretation for the double angle formulas. Perturbation measures as given by $\sin 2\Theta$-theorems can be viewed as bounds for the difference between $\widetilde{\mathfrak{I}}$ and its reflection $S\widetilde{\mathfrak{I}}$. Here S is a reflection operator where the mirror hyperplane is $\mathfrak{I}$ and S reverses the orthogonal complement of $\mathfrak{I}$. A direct bound on the angle between the original subspaces can be obtained by the same argument used for (1.2).

In this paper, at first, we consider the case when the matrices M and $\widetilde{M}$ are positive definite. Truhar and Li [21] and Li [17] studied a similar perturbation problem for the standard eigenvalue problem, and our bounds contain their results as a special case. Let us point out that almost all the known theorems for the standard eigenvalue problem have been generalized to the generalized eigenvalue problem; see, e.g., Li [14], who studied the generalized eigenvalue problem of a diagonalizable matrix pencil $H - \lambda M$ with real spectrum. A comprehensive overview of results from the point of view of relative perturbation theory can be found in [14] while in [5] similar questions have been considered using the standard (absolute) perturbation theory.

In the context of known operator theoretic results, this work also extend [8, 9, 22]. An advantage of new $\sin 2\Theta$-theorems over the existing $\sin \Theta$-theorems given in [8, 22] is that the relative gaps do not depend on the eigenvalues of the matrix pairs (H, M), $(\widetilde{H}, M)$, and $(\widetilde{H}, \widetilde{M})$ but just on the eigenvalues of the perturbed ones $(\widetilde{H}, M)$ and $(\widetilde{H}, \widetilde{M})$. Our estimates in their simplest form read as

$$\begin{aligned}\|\sin 2\Theta_M(\mathfrak{I}, \widetilde{\mathfrak{I}})\|_F \le \frac{\nu_1 \|B\|_2^4}{\mathrm{RelGap}_1} \|H^{-1}\|_2 \|H - \widetilde{H}\|_F \\ + \frac{\nu_2 \,\mathrm{RelGap}_2 + \nu_3}{\mathrm{RelGap}_2} \|M^{-\frac{1}{2}}(M - \widetilde{M}) M^{-\frac{1}{2}}\|_F.\end{aligned}$$

Here the constants RelGap_i, $i = 1, 2$, measure the separation between the wanted and unwanted components of the spectrum of the matrix pairs $(\widetilde{H}, M)$ and $(\widetilde{H}, \widetilde{M})$. The constants ν_i, $i = 1, 2, 3$, measure the stability of the inertia of a matrix pair (H, M) under a perturbation in

ETNA
Kent State University
http://etna.math.kent.edu

both matrices. Note that $\nu_i > 1$ for $i = 1, 2, 3$. Compared with (1.2), we observe that in the setting of the generalized eigenvalue problem, we have two a posteriori gaps to consider. One of them, namely the gaps in the spectrum of $(\widetilde{H}, M)$, is a purely artificial measure of spectral stability. However, it cannot be avoided due to the technique from [8] that we use.

The second result of this paper are extensions of a $\sin\Theta$-theorem from [8] and new $\sin 2\Theta$-theorems for the matrix pairs (H, M) with M positive definite to the case when the matrix M is indefinite non-singular. This is done so that instead of the definite pairs (H, M), where both matrices H and M are non-singular indefinite, we consider matrix pairs $(H, H - \alpha M)$, where $\alpha \in \mathbb{R}$ is such that $H - \alpha M$ is a positive definite matrix. Our bounds are directly derived from the earlier results and are dependent on the parameter α. Also, the rotation between the unperturbed and perturbed subspaces is measured in the $H - \alpha M$-scalar product. One of the criteria for the choice of α is given by Veselić in [24, Theorem A1]. We will follow this approach in our discussion of the dependence of the bounds on the choice of α.

The paper has the following structure. In Section 2 we present new relative $\sin 2\Theta$-theorems for matrix pairs (H, M) with M positive definite. This is the main technical result of this paper. Section 3 contains relative $\sin\Theta$- and $\sin 2\Theta$-theorems which are generalizations of the known results given in [8] as well as extensions of the results from Section 2 to the case when M is an indefinite non-singular Hermitian matrix. All results are illustrated by numerical examples in Section 4.

Notations. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the spectral and Frobenius norms, respectively, and $\|\cdot\|$ denotes any unitary invariant norm. X^* is the conjugate transpose. I_n denotes the $n \times n$ identity matrix (we may simply write I instead if no confusion can arise).

2. Main results. Throughout the first part of this paper, we study the perturbation theory for spectral projections of Hermitian matrix pairs (H, M), where H is a non-singular Hermitian matrix which can be factorized as

$$H = GJG^*, \qquad J = \operatorname{diag}(\pm 1).$$

Here G is assumed to be non-singular, and M is positive definite. The corresponding perturbed pair $(\widetilde{H}, \widetilde{M}) = (H + \delta H, M + \delta M)$ has the form

$$\widetilde{H} = \widetilde{G}J\widetilde{G}^*, \qquad J = \operatorname{diag}(\pm 1),$$

with $\widetilde{G}$ non-singular and $\widetilde{M}$ positive definite. We emphasize that H and $\widetilde{H}$ are assumed, as it is customary in relative perturbation theory, to have the same inertia.

Under these assumptions, the matrix pairs (H, M) and $(\widetilde{H}, \widetilde{M})$ can be simultaneously diagonalized. That is, there exist non-singular matrices X and $\widetilde{X}$ such that

$$X^*HX = \Lambda, \quad X^*MX = I, \qquad \text{and} \qquad \widetilde{X}^*\widetilde{H}\widetilde{X} = \widetilde{\Lambda}, \quad \widetilde{X}^*\widetilde{M}\widetilde{X} = I, \tag{2.1}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, $\widetilde{\Lambda} = \operatorname{diag}(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n)$, $\lambda_i, \widetilde{\lambda}_i \in \mathbb{R}$, for $i = 1, \ldots, n$.

Given k, $1 \le k < n$, let us partition the matrices X and $\widetilde{X}$ as

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \qquad \text{and} \qquad \widetilde{X} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix},$$

where $X_1, \widetilde{X}_1 \in \mathbb{C}^{n\times k}$ and $X_2, \widetilde{X}_2 \in \mathbb{C}^{n\times(n-k)}$. The eigendecomposition (2.1) can now be written as

$$\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} H \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \qquad \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} M \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \end{bmatrix}. \tag{2.2}$$

From (2.2) it follows that

$$HX_1 = MX_1\Lambda_1, \qquad HX_2 = MX_2\Lambda_2,$$

where $\Lambda_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_k) \in \mathbb{C}^{k\times k}$, $\Lambda_2 = \operatorname{diag}(\lambda_{k+1}, \ldots, \lambda_n) \in \mathbb{C}^{(n-k)\times(n-k)}$ such that $\operatorname{Spec}(\Lambda_1) \cap \operatorname{Spec}(\Lambda_2) = \emptyset$ and similarly for $\widetilde{X}$. We are interested in bounding the change in the subspaces $\operatorname{Ran}(X_1)$ spanned by the X_1's columns. We shall do this by estimating the sines of the double canonical angles between the subspaces $\operatorname{Ran}(X_1)$ and $\operatorname{Ran}(\widetilde{X}_1)$ in the M-dependent scalar product; for more details, see [9].

A bound for these will be obtained using the simple triangle inequality for the angle function as given by [9, Lemma 2.2]

$$\begin{aligned} \left\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}(\widetilde{X}_1))\right\| \le{}& \left\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}(\widehat{X}_1))\right\| \\ &+ \left\|\sin 2\Theta_M(\operatorname{Ran}(\widehat{X}_1), \operatorname{Ran}(\widetilde{X}_1))\right\| . \end{aligned} \tag{2.3}$$

Here $\widehat{X} = [\widehat{X}_1 \quad \widehat{X}_2] \in \mathbb{C}^{n\times n}$ is a non-singular matrix which simultaneously diagonalizes the matrix pair $(\widetilde{H}, M) = (H + \delta H, M)$. We would like to emphasize that the bounds are given in the Frobenius norm since then they can be stated without any restrictions on the position of the spectra of the matrix pairs (H, M), $(\widetilde{H}, M)$, and $(\widetilde{H}, \widetilde{M})$. After imposing additional assumptions on these spectra, similar bounds are derived for any unitary invariant norm, which we generically denote by $\|\cdot\|$. Such estimates are presented as corollaries to the main results below. As suggested by inequality (2.3), the bound for $\left\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}(\widetilde{X}_1))\right\|$ is obtained using a two-step procedure.

2.1. The first step. First, we estimate $\left\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}(\widehat{X}_1))\right\|$. Before we formulate the main results, we present some notational conventions for the block-matrix calculus that we use extensively in the proofs of the main results.

Let (H, M) be a Hermitian pair defined by (2.1) and $(\widetilde{H}, M)$ the corresponding perturbed pair. Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ be the non-singular matrix from (2.2). Assume that k $(1 \le k < n)$ is given as in (2.2), and let $\widehat{X} = \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix}$ be a non-singular matrix such that

$$\begin{aligned} \begin{bmatrix} \widehat{X}_1^* \\ \widehat{X}_2^* \end{bmatrix} (H + \delta H) \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix} &= \begin{bmatrix} \widehat{\Lambda}_1 & 0 \\ 0 & \widehat{\Lambda}_2 \end{bmatrix}, \\ \begin{bmatrix} \widehat{X}_1^* \\ \widehat{X}_2^* \end{bmatrix} M \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix} &= \begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \end{bmatrix}, \end{aligned} \tag{2.4}$$

where $\widehat{\Lambda}_1 = \operatorname{diag}(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_k)$, $\widehat{\Lambda}_2 = \operatorname{diag}(\widehat{\lambda}_{k+1}, \ldots, \widehat{\lambda}_n)$, $\widehat{\lambda}_i \in \mathbb{R}$, for $i = 1, \ldots, n$, and where we assume that $\operatorname{Spec}(\widehat{\Lambda}_1) \cap \operatorname{Spec}(\widehat{\Lambda}_2) = \emptyset$.

Define

$$S_r := X \begin{bmatrix} I_k & \\ & -I_{n-k} \end{bmatrix} X^{-1} = X \begin{bmatrix} I_k & \\ & -I_{n-k} \end{bmatrix} X^* M, \tag{2.5}$$

and note that

$$S_r^2 = I_n, \quad \|S_r\|_2 \le \kappa(X) \qquad \text{and} \qquad S_r^* H S_r = H, \quad S_r^* M S_r = M.$$

ETNA
Kent State University
http://etna.math.kent.edu

Now we define the auxiliary matrix $\widehat{H} = S_r^* \widetilde{H} S_r = S_r^* \widehat{X}^{-*} \widehat{\Lambda} \widehat{X}^{-1} S_r = \widehat{Y}^{-*} \widehat{\Lambda} \widehat{Y}^{-1}$, where $\widehat{Y} := S_r \widehat{X}$ and

$$\widehat{Y}^* \widehat{H} \widehat{Y} = \widehat{\Lambda} \quad \text{and} \quad \widehat{Y}^* M \widehat{Y} = I. \tag{2.6}$$

For a given k $(1 \le k < n)$, let us partition the matrix $\widehat{Y}$ such that

$$\widehat{Y} = \begin{bmatrix} \widehat{Y}_1 & \widehat{Y}_2 \end{bmatrix}, \qquad \widehat{Y}_1 = S_r \widehat{X}_1 \in \mathbb{C}^{n\times k}, \qquad \widehat{Y}_2 = S_r \widehat{X}_2 \in \mathbb{C}^{n\times(n-k)}. \tag{2.7}$$

The norm of the sines of the double angle between the subspaces $\mathrm{Ran}(X_1)$ and $\mathrm{Ran}(\widehat{X}_1)$ is the same as the norm of the sines of the single angle between the subspaces $\mathrm{Ran}(\widehat{Y}_1)$ and $\mathrm{Ran}(\widehat{X}_1)$ as described in the following lemma.

LEMMA 2.1. *Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and $\widehat{X} = \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix}$ with $X_1, \widehat{X}_1 \in \mathbb{C}^{n\times k}$ and $X_2, \widehat{X}_2 \in \mathbb{C}^{n\times(n-k)}$ be non-singular matrices which simultaneously diagonalize the Hermitian matrix pairs (H, M) and $(\widetilde{H}, M)$ as in (2.2) and (2.4), where $H, \widetilde{H} \in \mathbb{C}^{n\times n}$ are indefinite and $M \in \mathbb{C}^{n\times n}$ is positive definite. Let $\widehat{Y} = S_r \widehat{X} = \begin{bmatrix} \widehat{Y}_1 & \widehat{Y}_2 \end{bmatrix}$ be an M-orthogonal matrix where $S_r \in \mathbb{C}^{n\times n}$ is defined in (2.5). Then*

$$\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widehat{X}_1))\| = \|\sin \Theta_M(\mathrm{Ran}(\widehat{Y}_1), \mathrm{Ran}(\widehat{X}_1))\| = \|\widehat{Y}_2^* M \widehat{X}_1\|.$$

Proof. The matrix $X^* M \widehat{X} = \begin{bmatrix} X_1^* M \widehat{X}_1 & X_1^* M \widehat{X}_2 \\ X_2^* M \widehat{X}_1 & X_2^* M \widehat{X}_2 \end{bmatrix}$ is unitary. Using a CS decomposition of $X^* M \widehat{X}$, there exists unitary matrices $U_1, V_1 \in \mathbb{C}^{k\times k}$ and $U_2, V_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ such that

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^* X^* M \widehat{X} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \left[\begin{array}{c|c|c} C & 0 & -S \\ \hline 0 & I_{n-k} & 0 \\ S & 0 & C \end{array}\right] \qquad \text{when } k < \frac{n}{2},$$

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^* X^* M \widehat{X} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \qquad \text{when } k = \frac{n}{2},$$

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^* X^* M \widehat{X} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \left[\begin{array}{cc|c} I_{2k-n} & 0 & 0 \\ 0 & C & -S \\ \hline 0 & S & C \end{array}\right] \qquad \text{when } k \ge \frac{n}{2}.$$

Here $C = \mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_p)$ and $S = \mathrm{diag}(\sin\theta_1, \ldots, \sin\theta_p)$, and $\theta_1, \ldots, \theta_n$ are the canonical angles between subspaces measured in the M-inner product.

Without loss of generality let us assume that $k = \frac{n}{2}$. Then note that

$$\widehat{Y}^* M \widehat{X} = \begin{bmatrix} \widehat{Y}_1^* M \widehat{X}_1 & \widehat{Y}_1^* M \widehat{X}_2 \\ \widehat{Y}_2^* M \widehat{X}_1 & \widehat{Y}_2^* M \widehat{X}_2 \end{bmatrix}$$

is a unitary matrix, and its CS decomposition states—recall that $k = \frac{n}{2}$ is assumed—that there exist unitary matrices $W_1 \in \mathbb{C}^{k\times k}$, $W_2 \in \mathbb{C}^{(n-k)\times(n-k)}$, $Z_1 \in \mathbb{C}^{k\times k}$ and $Z_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ such that

$$\begin{bmatrix} W_1 & \\ & W_2 \end{bmatrix}^* \widehat{Y}^* M \widehat{X} \begin{bmatrix} Z_1 & \\ & Z_2 \end{bmatrix} = \begin{bmatrix} C_1 & -S_1 \\ S_1 & C_1 \end{bmatrix}. \tag{2.8}$$

ETNA
Kent State University
http://etna.math.kent.edu

From (2.5) and (2.7) we have

$$\begin{aligned} \widehat{Y}_1 &= X_1X_1^*M\widehat{X}_1 - X_2X_2^*M\widehat{X}_1, \\ \widehat{Y}_2 &= X_1X_1^*M\widehat{X}_2 - X_2X_2^*M\widehat{X}_2 \,. \end{aligned} \tag{2.9}$$

Inserting (2.9) into (2.8) it is easy to see that

$$\begin{bmatrix} W_1 & \\ & W_2 \end{bmatrix}^* \widehat{Y}^*M\widehat{X} \begin{bmatrix} Z_1 & \\ & Z_2 \end{bmatrix} = \begin{bmatrix} C^2+S^2 & -(C\cdot S + S\cdot C) \\ S\cdot C + C\cdot S & S^2 - C^2 \end{bmatrix}. \tag{2.10}$$

Now the proof simply follows by equating the right-hand side blocks $(1,2)$ of (2.8) and (2.10) and taking the norm of both sides. ∎

Using the previous lemma, we observe that in order to establish the estimate of the first term in the inequality (2.3), we have to bound $\|\widehat{Y}_2^*M\widehat{X}_1\|_F$. This can be done using a similar technique as in [8]. Before we formulate the main result, we recall the following remark which we need later; see [21].

REMARK 2.2. Let us assume that the matrix $H = GJG^*$ is perturbed such that $\widetilde{H} = G(J+E)G^*$. Since $\|H^{-1}\|_2\|\delta H\|_2 < 1$, it follows that $\|E\|_2 < 1$. From this we conclude that $\|EJ\|_2 < 1$, and hence, we can define $N := (I+EJ)^{1/2}$. Recall the following series expansion from [11, Theorem 6.2.8]:

$$N = (I+EJ)^{1/2} = I + \sum_{i=1}^{\infty}(-1)^{n-1}\frac{(2n-1)!!}{2^n n!}(EJ)^n,$$

where $(2n-1)!! = 1\cdot 3\cdot 5\cdots(2n-1)$. Since $\|EJ\|_2 < 1$, the series obviously converges, and it can be verified that $N = JN^*J$. Subsequently,

$$J + E = NJN^*,$$

and so $\|E\|_2 < 1$ implies that $H = GJG^*$, $J+E$, and $\widetilde{H} = G(J+E)G^*$ all have the same inertia as J.

Now we can state our first theorem:

THEOREM 2.3. *Let $H = GJG^*$, $\widetilde{H} = \widetilde{G}J\widetilde{G}^*$, and M be positive definite, and let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and $\widehat{X} = \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix}$ be non-singular matrices from* (2.2) *and* (2.4) *which simultaneously diagonalize the pairs (H,M) and $(\widetilde{H},M)$, respectively. Further, let B and $\widetilde{B}$ be J-unitary matrices which simultaneously diagonalize the pairs (G^*G,J) and $(\widetilde{G}^*\widetilde{G},J)$, respectively. If $\|H^{-1}\|_2\|\delta H\|_2 < 2/3$, then*

$$\frac{1}{2}\|\sin 2\Theta_M(\mathrm{Ran}(X_1),\mathrm{Ran}(\widehat{X}_1))\|_F \le \frac{\|B\|_2^2\|\widehat{B}_2\|_2\|\widetilde{B}_1\|_2\,\nu_{\mathrm{fr}}}{\mathrm{RelGap}_1}\|H^{-1}\|_2\|\delta H\|_F, \tag{2.11}$$

where

$$\mathrm{RelGap}_1 = \min_{\substack{i=k+1,\dots,n\\ j=1,\dots,k}} \frac{|\widehat{\lambda}_i - \widehat{\lambda}_j|}{\sqrt{|\widehat{\lambda}_i||\widehat{\lambda}_j|}} \qquad \textit{and}$$

$$\nu_{\mathrm{fr}} = \frac{2-\|H^{-1}\|_2\|\delta H\|_2}{(1-\|H^{-1}\|_2\|\delta H\|_2)(2-3\|H^{-1}\|_2\|\delta H\|_2)}. \tag{2.12}$$

ETNA
Kent State University
http://etna.math.kent.edu

Proof. As it has been shown in Lemma 2.1, it is enough to bound $\|\widehat{Y}_2^* M \widehat{X}_1\|_F$. From (2.4) and (2.6) we have

$$\widetilde{H}\widehat{X}_1 = M\widehat{X}_1\widehat{\Lambda}_1, \qquad \widehat{H}\widehat{Y}_2 = M\widehat{Y}_2\widehat{\Lambda}_2 .$$

Multiply the first equation from the left by $\widehat{Y}_2^*$ and the second by $\widehat{X}_1^*$ to get

$$\widehat{Y}_2^*\widetilde{H}\widehat{X}_1 = \widehat{Y}_2^* M\widehat{X}_1\widehat{\Lambda}_1, \qquad \widehat{X}_1^*\widehat{H}\widehat{Y}_2 = \widehat{X}_1^* M\widehat{Y}_2\widehat{\Lambda}_2 . \tag{2.13}$$

Transposing the second equation in (2.13) and subtracting them, we obtain

$$\widehat{\Lambda}_2\widehat{Y}_2^* M\widehat{X}_1 - \widehat{Y}_2^* M\widehat{X}_1\widehat{\Lambda}_1 = \widehat{Y}_2^*(\widehat{H} - \widetilde{H})\widehat{X}_1.$$

Also, using hyperbolic singular value decomposition of the matrices $\widetilde{G}$ and $\widehat{G}$, we have that

$$\begin{aligned} \widetilde{G}^*\widehat{X} &= \widetilde{B}|\widehat{\Lambda}|^{1/2} \\ \widehat{G}^*\widehat{Y} &= \widehat{B}|\widehat{\Lambda}|^{1/2} , \end{aligned} \tag{2.14}$$

where $\widetilde{B}$ and $\widehat{B}$ are J-unitary matrices which simultaneously diagonalize the matrix pairs $(\widetilde{G}^*\widetilde{G}, J)$ and $(\widehat{G}^*\widehat{G}, J)$, respectively. Starting from

$$\widehat{\Lambda}_2\widehat{Y}_2^* M\widehat{X}_1 - \widehat{Y}_2^* M\widehat{X}_1\widehat{\Lambda}_1 = \widehat{Y}_2^*\widehat{G}\widehat{G}^{-1}(\widehat{H} - \widetilde{H})\widetilde{G}^{-*}\widetilde{G}^*\widehat{X}_1$$

and (2.14), we establish a structured Sylvester equation

$$\widehat{\Lambda}_2\widehat{Y}_2^* M\widehat{X}_1 - \widehat{Y}_2^* M\widehat{X}_1\widehat{\Lambda}_1 = |\widehat{\Lambda}_2|^{1/2}\widehat{B}_2^*\widehat{G}^{-1}(\widehat{H} - \widetilde{H})\widetilde{G}^{-*}\widetilde{B}_1|\widehat{\Lambda}_1|^{1/2} . \tag{2.15}$$

Using [16, Lemma 2.4] for (2.15), we find the bound

$$\|\widehat{Y}_2^* M\widehat{X}_1\|_F \le \frac{\|\widehat{B}_2\|_2\|\widetilde{B}_1\|_2\|\widehat{G}^{-1}(\widehat{H} - \widetilde{H})\widetilde{G}^{-*}\|_F}{\mathrm{RelGap}_1} , \tag{2.16}$$

where

$$\mathrm{RelGap}_1 = \min_{\substack{i=k+1,\ldots,n \\ j=1,\ldots,k}} \frac{|\widehat{\lambda}_i - \widehat{\lambda}_j|}{\sqrt{|\widehat{\lambda}_i||\widehat{\lambda}_j|}} .$$

We have now reduced the problem to that of estimating $\|\widehat{G}^{-1}(\widehat{H} - \widetilde{H})\widetilde{G}^{-*}\|_F$ in (2.16).

Towards this end, set

$$H = GJG^*, \qquad \widetilde{H} = GNJN^*G^* \equiv \widetilde{G}J\widetilde{G}^*, \qquad \widetilde{G} = GN, \tag{2.17}$$

and define

$$W := G^{-1}S_r^*G = B^{-*}\begin{bmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{bmatrix} B^*. \tag{2.18}$$

First, note that

$$WJW^* = G^{-1}S_r^*GJG^*S_rG^{-*} = G^{-1}HG^{-*} = J,$$

and so W is J-unitary. Further, we have that

$$W^2 = I, \qquad W^*J = JW, \qquad \|W\|_2 \le \|B\|_2^2.$$

ETNA
Kent State University
http://etna.math.kent.edu

We use the identity $GW = S_r^* G$, which follows from (2.18), to get

$$\begin{aligned}\widehat{H} &= S_r^* \widetilde{H} S_r = S_r^* \widetilde{G} J \widetilde{G}^* S_r = S_r^* G N J N^* G^* S_r \\ &= G W N J N^* W^* G^* = G N N^{-1} W N J N^* W^* N^{-*} N^* G^*,\end{aligned}$$

where $\widetilde{G}$ is given as in (2.17),

$$\widetilde{N} = N^{-1} W N, \qquad \widehat{G} = \widetilde{G}\widetilde{N},$$

and, hence,

$$\widehat{G}\widehat{Y} = \widehat{B}|\widehat{\Lambda}|^{1/2}.$$

Now, it is easy to see that

$$\|\widehat{G}^{-1}(\widehat{H} - \widetilde{H})\widetilde{G}^{-*}\|_F = \|J\widehat{G}^*\widetilde{G}^{-*} - \widehat{G}^{-1}\widetilde{G}J\|_F = \|J\widetilde{N}^* - \widetilde{N}^{-1}J\|_F \,.$$

We express a bound for $\|J\widetilde{N}^* - \widetilde{N}^{-1}J\|_F$ in terms of the norm of E (and thus of $\|\delta H\|_F$). Similarly as in [21, Section 3.1] and Remark 2.2, one can obtain the following inequality:

$$\|J\widetilde{N}^* - \widetilde{N}^{-1}J\|_F \le 2\|B\|_2^2\, \nu_{\mathrm{fr}}\, \|H^{-1}\|_2 \|\delta H\|_F \,, \tag{2.19}$$

where

$$\nu_{\mathrm{fr}} = \frac{2 - \|H^{-1}\|_2\|\delta H\|_2}{(1 - \|H^{-1}\|_2\|\delta H\|_2)(2 - 3\|H^{-1}\|_2\|\delta H\|_2)}. \tag{2.20}$$

The proof now simply follows from (2.16) and (2.19). □

The bound (2.11) given in the previous theorem depends on the norm of the J-unitary matrices $\widetilde{B}$ and $\widehat{B}$. Here we want to emphasize that $\|\widetilde{B}\|_2$ and $\|\widehat{B}\|_2$ can be bounded in terms of $\|B\|_2$ as it is done in [21, Section 3.2]. Using these results, we can state the following corollary.

COROLLARY 2.4. *To the conditions of Theorem 2.3 add the following:*

$$\gamma := \frac{\|H^{-1}\|_2\|\delta H\|_F}{2 - 3\|H^{-1}\|_2\|\delta H\|_2} \le \frac{1}{4\|B\|_2^2}\,.$$

Then

$$\frac{1}{2}\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widehat{X}_1))\|_F \le \frac{\|B\|_2^4\, \nu_{\mathrm{fr}}}{1 - 4\gamma\|B\|_2^2}\, \frac{\|H^{-1}\|_2\|\delta H\|_F}{\mathrm{RelGap}_1}\,. \tag{2.21}$$

The bounds (2.11) and (2.21) can be also given in any unitary invariant norm $\|\cdot\|$ by imposing additional assumptions on the spectra of the matrix pairs $(\widetilde{H}, M)$. This is stated in the next corollary.

COROLLARY 2.5. *Let the same assumptions as in Theorem 2.3 and Corollary 2.4 hold. If there exists $a \ge 0$ and $\delta > 0$ such that*

$$\||\widehat{\Lambda}_1|\|_2 \le a \qquad \textit{and} \qquad \||\widehat{\Lambda}_2|^{-1}\|_2^{-1} \ge a + \delta$$

or

$$\||\widehat{\Lambda}_1|^{-1}\|_2^{-1} \ge a + \delta \qquad \textit{and} \qquad \||\widehat{\Lambda}_2|\|_2 \le a\,,$$

ETNA
Kent State University
http://etna.math.kent.edu

then

$$\frac{1}{2}\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widehat{X}_1))\| \le \frac{\|B\|_2^4\,\nu_{\mathrm{fr}}}{1-4\gamma\|B\|_2^2}\,\frac{\|H^{-1}\|_2\|\delta H\|}{\frac{a}{\sqrt{a(a+\delta)}}}. \tag{2.22}$$

Proof. The proof follows by applying [8, Lemma 2.1] to the structured Sylvester equation (2.15). ◻

The estimates given here contain an additional factor which depends on the J-unitary matrix B from (2.14), whose norm may be large for the case when the matrix H in the matrix pair (H, M) is any indefinite Hermitian matrix. There exist several different estimates for $\|B\|_2$ which are given for certain classes of matrices $H = GJG^*$. For example, in [23] and [21], one can find such estimates for the case when $H = GJG^*$ belongs to the class of so-called "well-behaved matrices" as defined by Demmel and Barlow [3]. This class of matrices contains scaled diagonal dominant matrices, block-scaled diagonally dominant (BSDD) matrices, and *quasi-definite* matrices. For basic properties of the J-unitary matrices and for some other bounds for $\|B\|_2$, see, e.g., [24].

In [8] one can find a new sharp estimate for the condition number of all J-unitary matrices which diagonalize a pair (G^*G, J), where $H = GJG^*$ is a *quasi-definite* block matrix. A matrix H is a *quasi-definite* matrix if there exists a permutation matrix P such that

$$H_{qd} \equiv P^*HP = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^* & -H_{22} \end{bmatrix},$$

where H_{11} and $H_{22} + H_{12}^*H_{11}^{-1}H_{12}$ are positive definite. Such matrices can be factorized in the form $H = GJG^*$, where

$$G = \begin{bmatrix} I & 0 \\ H_{12}^*H_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} H_{11}^{1/2} & 0 \\ 0 & (H_{22} + H_{12}^*H_{11}^{-1}H_{12})^{1/2} \end{bmatrix}, \qquad J = \begin{bmatrix} I & \\ & -I \end{bmatrix}.$$

The bound given in [8, Theorem 4.1] has the following form:

$$\|B\|_2 \le \frac{\|H_{12}^*H_{11}^{-1}\|_2}{2} + \sqrt{1 + \left(\frac{\|H_{12}^*H_{11}^{-1}\|_2}{2}\right)^2}.$$

2.2. The second step. In a second step we consider upper bounds for the expression $\left\|\sin 2\Theta_M(\mathrm{Ran}(\widehat{X}_1), \mathrm{Ran}(\widetilde{X}_1))\right\|$. The corresponding estimate will be stated in our second theorem. As in Section 2.1, we start by specifying the notation for the block-matrix representation of the matrices involved. This matrix calculus will be used extensively in the proofs that follow.

Assume k $(1 \le k < n)$ as in (2.2) and (2.4), and let $\widetilde{X} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix}$ be a non-singular matrix such that

$$\begin{bmatrix} \widetilde{X}_1^* \\ \widetilde{X}_2^* \end{bmatrix} \widetilde{H} \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix} = \begin{bmatrix} \widetilde{\Lambda}_1 & 0 \\ 0 & \widetilde{\Lambda}_2 \end{bmatrix}, \qquad \begin{bmatrix} \widetilde{X}_1^* \\ \widetilde{X}_2^* \end{bmatrix} \widetilde{M} \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \end{bmatrix}, \tag{2.23}$$

where $\widetilde{\Lambda}_1 = \mathrm{diag}(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_k)$ and $\widetilde{\Lambda}_2 = \mathrm{diag}(\widetilde{\lambda}_{k+1}, \ldots, \widetilde{\lambda}_n)$, $\widetilde{\lambda}_i \in \mathbb{R}$, for $i = 1, \ldots, n$, and $\mathrm{Spec}(\widetilde{\Lambda}_1) \cap \mathrm{Spec}(\widetilde{\Lambda}_2) = \emptyset$. Also, recall that $\widehat{X} = \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix}$ satisfies (2.4).

Similarly as in the previous section, we define the matrix

$$T_r = \widehat{X} \begin{bmatrix} I_k & \\ & -I_{n-k} \end{bmatrix} \widehat{X}^{-1} = \widehat{X} \begin{bmatrix} I_k & \\ & -I_{n-k} \end{bmatrix} \widehat{X}^* M. \tag{2.24}$$

Note that

$$T_r^2 = I_n, \quad \|T_r\|_2 \le \kappa(\widehat{X}) \qquad \text{and} \qquad T_r^* \widetilde{H} T_r = \widetilde{H}, \quad T_r^* M T_r = M.$$

Also, we define an auxiliary matrix $\widehat{M}$ as

$$\widehat{M} = T_r^* \widetilde{M} T_r = \widetilde{Y}^{-*} I_n \widetilde{Y}^{-1},$$

where $\widetilde{Y} = T_r \widetilde{X}$. The matrix $\widetilde{Y}$ is $\widehat{M}$-orthogonal and

$$\widetilde{Y}^* \widetilde{H} \widetilde{Y} = \widetilde{X}^* T_r^* \widetilde{H} T_r \widetilde{X} = \widetilde{X}^* \widetilde{H} \widetilde{X} = \widetilde{\Lambda}.$$

For a given k $(1 \le k < n)$, let us partition the matrix $\widetilde{Y}$ such that

$$\widetilde{Y} = \begin{bmatrix} \widetilde{Y}_1 & \widetilde{Y}_2 \end{bmatrix}, \qquad Y_1 \in \mathbb{C}^{n \times k}, \qquad Y_2 \in \mathbb{C}^{n \times (n-k)}.$$

Using a similar argument as in the previous section, the norm of the sines of the double angles between the subspaces $\mathrm{Ran}(\widehat{X}_1)$ and $\mathrm{Ran}(\widetilde{X}_1)$ is the same as the norm of the sines of the single angles between the subspaces $\mathrm{Ran}(\widetilde{X}_1)$ and $\mathrm{Ran}(\widetilde{Y}_1)$.

First, let us define the sines of the angles between the subspaces $\mathrm{Ran}(\widetilde{X}_1)$ and $\mathrm{Ran}(\widetilde{Y}_1)$ in the M-inner product space. The matrix $\widetilde{Y}$ is $\widehat{M}$-orthogonal, and $\widetilde{X}$ is $\widetilde{M}$-orthogonal. The relationship between matrices which are orthogonal in the $\widetilde{M}$ $(\widetilde{M} = M + \delta M)$- and $\widehat{M}$ $(\widehat{M} = M + T_r^* \delta M T_r)$-scalar product is important for our perturbation theory. For $\widetilde{X}$ such that $\widetilde{X}^* \widetilde{M} \widetilde{X} = I$, we compute

$$\widetilde{X}^* M \widetilde{X} = I - \widetilde{X}^* \delta M \widetilde{X}.$$

Assume that $I - \widetilde{X}^* \delta M \widetilde{X}$ is positive definite. Then it has the block-Cholesky decomposition

$$KK^* = I - \widetilde{X}^* \delta M \widetilde{X}, \tag{2.25}$$

where

$$K = \begin{bmatrix} K_{11} & 0 \\ K_{21} & K_{22} \end{bmatrix}$$

and $K_{11} \in \mathbb{C}^{k \times k}$, $K_{21} \in \mathbb{C}^{(n-k) \times k}$, $K_{22} \in \mathbb{C}^{(n-k) \times (n-k)}$. A direct computation proves that the matrix $\widetilde{X} K^{-*}$ is M-orthogonal. Similarly, we conclude that the columns of $\widetilde{X}_1 K_{11}^{-*}$, where $K_{11} K_{11}^* = I_k - \widetilde{X}_1^* \delta M \widetilde{X}_1$, are M-orthogonal. Also, for $\widetilde{Y}$ such that

$$\widetilde{Y}^* \widehat{M} \widetilde{Y} = \widetilde{Y}^* T_r^* (M + \delta M) T_r \widetilde{Y} = I,$$

using the fact that $\widetilde{X} = T_r \widetilde{Y}$, it follows that

$$\widetilde{Y}^* M \widetilde{Y} = I - \widetilde{Y}^* T_r^* \delta M T_r \widetilde{Y} = I - \widetilde{X}^* \delta M \widetilde{X} = KK^*.$$

It is easy to see now that the matrix $\widetilde{Y} K^{-*}$ is an M-orthogonal matrix, and by direct computation we conclude that the columns of the matrix $-\widetilde{Y}_1 K_{11}^{-*} K_{21}^* K_{22}^{-*} + \widetilde{Y}_2 K_{22}^{-*}$ are M-orthogonal.

Finally, we obtain

$$\begin{aligned} &\| \sin \Theta_M(\mathrm{Ran}(\widetilde{X}_1), \mathrm{Ran}(\widetilde{Y}_1)) \| \\ &\qquad = \| - K_{22}^{-1} K_{21} K_{11}^{-1} \widetilde{Y}_1^* M \widetilde{X}_1 K_{11}^{-*} + K_{22}^{-1} \widetilde{Y}_2^* M \widetilde{X}_1 K_{11}^{-*} \| \,, \end{aligned} \tag{2.26}$$

ETNA
Kent State University
http://etna.math.kent.edu

which will be used to provide estimates of the angle operator in the next lemma.

LEMMA 2.6. *Let* $\widehat{X} = \begin{bmatrix}\widehat{X}_1 & \widehat{X}_2\end{bmatrix}$ *and* $\widetilde{X} = \begin{bmatrix}\widetilde{X}_1 & \widetilde{X}_2\end{bmatrix}$, *with* $\widehat{X}_1, \widetilde{X}_1 \in \mathbb{C}^{n\times k}$ *and* $\widehat{X}_2, \widetilde{X}_2 \in \mathbb{C}^{n\times(n-k)}$, *be non-singular matrices which simultaneously diagonalize the Hermitian matrix pairs* $(\widetilde{H}, M)$ *and* $(\widetilde{H}, \widetilde{M})$ *as in* (2.4) *and* (2.23)*, respectively, where* $\widetilde{H} \in \mathbb{C}^{n\times n}$ *is indefinite and* M, $\widetilde{M} \in \mathbb{C}^{n\times n}$ *are positive definite matrices. Let* $\widetilde{Y} = T_r\widetilde{X} = \begin{bmatrix}\widetilde{Y}_1 & \widetilde{Y}_2\end{bmatrix}$ *be an* $\widehat{M}$*-orthogonal matrix, where* $T_r \in \mathbb{C}^{n\times n}$ *is defined in* (2.24) $(\widehat{M} = T_r^*\widetilde{M}T_r)$*. Then*

$$\|\sin 2\Theta_M(\mathrm{Ran}(\widehat{X}_1), \mathrm{Ran}(\widetilde{X}_1))\| = \|\sin\Theta_M(\mathrm{Ran}(\widetilde{X}_1), \mathrm{Ran}(\widetilde{Y}_1))\|. \tag{2.27}$$

Proof. It is easy to see that

$$K^{-1}\widetilde{X}^*M\widehat{X} = \begin{bmatrix} K_{11}^{-1}\widetilde{X}_1^*M\widehat{X}_1 & K_{11}^{-1}\widetilde{X}_1^*M\widehat{X}_2 \\ (K^{-1}\widetilde{X}^*M\widehat{X})_{21} & (K^{-1}\widetilde{X}^*M\widehat{X})_{22}\end{bmatrix} \tag{2.28}$$

is a unitary matrix, where

$$(K^{-1}\widetilde{X}^*M\widehat{X})_{21} = -K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{X}_1^*M\widehat{X}_1 + K_{22}^{-1}\widetilde{X}_2^*M\widehat{X}_1,$$
$$(K^{-1}\widetilde{X}^*M\widehat{X})_{22} = -K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{X}_1^*M\widehat{X}_2 + K_{22}^{-1}\widetilde{X}_2^*M\widehat{X}_2.$$

Similarly as in the proof of Lemma 2.1, without loss of generality, it can be assumed that $k = \frac{n}{2}$. Then by a CS decomposition, there exists unitary matrices $U_1, V_1 \in \mathbb{C}^{k\times k}$ and $U_2, V_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ such that

$$\begin{bmatrix}U_1 & \\ & U_2\end{bmatrix}^* K^{-1}\widetilde{X}^*M\widehat{X}\begin{bmatrix}V_1 & \\ & V_2\end{bmatrix} = \begin{bmatrix}C & -S\\ S & C\end{bmatrix}. \tag{2.29}$$

Here $C = \mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_p)$ and $S = \mathrm{diag}(\sin\theta_1, \ldots, \sin\theta_p)$, and $\theta_1, \ldots, \theta_n$ are canonical angles between subspaces measured in the M-inner product.

Since

$$\begin{aligned}&\|\sin\Theta_M(\mathrm{Ran}(\widetilde{X}_1), \mathrm{Ran}(\widetilde{Y}_1))\|_F\\ &\quad = \|-K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{Y}_1^*M\widetilde{X}_1K_{11}^{-*} + K_{22}^{-1}\widetilde{Y}_2^*M\widetilde{X}_1K_{11}^{-*}\|_F\end{aligned} \tag{2.30}$$

and

$$\begin{aligned}\widetilde{Y}_1 &= T_r\widetilde{X}_1 = \widehat{X}_1\widehat{X}_1^*M\widetilde{X}_1 - \widehat{X}_2\widehat{X}_2^*M\widetilde{X}_1\\ \widetilde{Y}_2 &= T_r\widetilde{X}_1 = \widehat{X}_1\widehat{X}_1^*M\widetilde{X}_2 - \widehat{X}_2\widehat{X}_2^*M\widetilde{X}_2\,,\end{aligned} \tag{2.31}$$

and by inserting (2.31) into (2.30) and using (2.29), we obtain

$$\|\sin\Theta_M(\mathrm{Ran}(\widetilde{X}_1), \mathrm{Ran}(\widetilde{Y}_1))\| = \|2SC\|. \tag{2.32}$$

As C and S are diagonal matrices with the cosine and sine of the canonical angles between the subspaces $\mathrm{Ran}(\widetilde{X}_1)$ and $\mathrm{Ran}(\widehat{X}_1)$ in (2.27) on its diagonal, the proof follows from (2.32). ∎

We collect these observations and results in the next theorem, which will be proved in Appendix A.

ETNA
Kent State University
http://etna.math.kent.edu

THEOREM 2.7. *Let* $\widehat{X} = \begin{bmatrix} \widehat{X}_1 & \widehat{X}_2 \end{bmatrix}$ *and* $\widetilde{X} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix}$ *be non-singular matrices which simultaneously diagonalize the pairs* $(H+\delta H, M)$ *and* $(H+\delta H, M+\delta M)$*, respectively, as in* (2.4) *and* (2.23)*. If*

$$\eta_M := \|M^{-1/2}\delta M M^{-1/2}\|_2 < \frac{1}{2}, \tag{2.33}$$

then

$$\begin{aligned} \|\sin 2\Theta_M(\operatorname{Ran}\widehat{X}_1, \operatorname{Ran}\widetilde{X}_1)\|_F \le \nu_{\text{sec}1}\, &\|\widetilde{M}^{-1/2}\delta M\widetilde{M}^{-1/2}\|_F \\ &+ \frac{\nu_{\text{sec}2}}{\operatorname{RelGap}_2}\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|_F, \end{aligned} \tag{2.34}$$

where

$$\operatorname{RelGap}_2 := \min_{\substack{i=k+1,\dots,n\\ j=1,\dots,k}} \frac{|\widetilde{\lambda}_i - \widetilde{\lambda}_j|}{\sqrt{|\widetilde{\lambda}_i|^2 + |\widetilde{\lambda}_j|^2}}, \tag{2.35}$$

and

$$\nu_{\text{sec}1} := \frac{1-\eta_M}{1-2\eta_M}, \qquad \nu_{\text{sec}2} := \frac{(2-3\eta_M)(1-\eta_M)}{\sqrt{2}(1-2\,\eta_M)^2}.$$

By additional assumptions on the spectra of the matrix pairs $(\widetilde{H}, \widetilde{M})$, the bound (2.34) can be derived for any unitary invariant norm which we generically denote by $\|\cdot\|$. This is shown in the next corollary.

COROLLARY 2.8. *Let the same assumptions as in Theorem 2.7 hold. If there exists* $a \ge 0$ *and* $\delta > 0$ *such that*

$$\||\widetilde{\Lambda}_1|\|_2 \le a \qquad \textit{and} \qquad \||\widetilde{\Lambda}_2|^{-1}\|_2^{-1} \ge a+\delta \tag{2.36}$$

or

$$\||\widetilde{\Lambda}_1|^{-1}\|_2^{-1} \ge a+\delta \qquad \textit{and} \qquad \||\widetilde{\Lambda}_2|\|_2 \le a,$$

then

$$\begin{aligned} &\|\sin 2\Theta_M(\operatorname{Ran}\widehat{X}_1, \operatorname{Ran}\widetilde{X}_1)\| \\ &\qquad \le \nu_{\text{sec}1}\,\|\widetilde{M}^{1/2}\delta M\widetilde{M}^{-1/2}\| + \frac{\nu_{\text{sec}2}\,\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|}{\frac{\delta}{\sqrt{|a|+|a+\delta|}}}, \end{aligned} \tag{2.37}$$

and

$$\nu_{\text{sec}1} = \frac{1-\eta_M}{1-2\eta_M}, \qquad \nu_{\text{sec}2} = \frac{(2-3\eta_M)(1-\eta_M)}{\sqrt{2}(1-2\,\eta_M)^2}.$$

Proof. The proof follows by applying [15, Lemma 2.3] to the structured Sylvester equation (A.10). □

2.3. The main result. As indicated earlier (see (2.3)), without any additional assumptions on the location of the spectra, we obtain an upper bound for

$$\|\sin 2\Theta_M(\mathfrak{I}, \widetilde{\mathfrak{I}})\|_F = \|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}(\widetilde{X}_1))\|_F$$

ETNA
Kent State University
http://etna.math.kent.edu

as the sum of the bounds for

$$\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widehat{X}_1))\|_F \qquad \text{and} \qquad \|\sin 2\Theta_M(\mathrm{Ran}(\widehat{X}_1), \mathrm{Ran}(\widetilde{X}_1))\|_F.$$

The following is our main result:

THEOREM 2.9. *Let (H, M) be a Hermitian pair, and let $(\widetilde{H}, \widetilde{M})$ be the perturbed pair. Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and $\widetilde{X} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix}$ be non-singular matrices which simultaneously diagonalize the pairs (H, M) and $(H + \delta H, M + \delta M)$ as in (2.2) and (2.23), respectively. Let B be the J-unitary matrix from Theorem 2.3. If*

$$\eta_M := \|M^{-1/2}\delta M M^{-1/2}\|_2 < \frac{1}{2},$$

then

$$\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \leq \frac{2\|B\|_2^4}{1 - 4\gamma\|B\|_2^2} \frac{\nu_{\mathrm{fr}}\, \|H^{-1}\|_2 \|\delta H\|_F}{\mathrm{RelGap}_1} + \frac{\nu_{\mathrm{sec1}}\, \mathrm{RelGap}_2 + \nu_{\mathrm{sec2}}}{\mathrm{RelGap}_2} \|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|_F, \tag{2.38}$$

where

$$\nu_{\mathrm{sec1}} = \frac{1 - \eta_M}{1 - 2\eta_M}, \qquad \nu_{\mathrm{sec2}} := \frac{(2 - 3\eta_M)(1 - \eta_M)}{\sqrt{2}(1 - 2\,\eta_M)^2}, \qquad \gamma = \frac{\|H^{-1}\|_2 \|\delta H\|_F}{2 - 3\|H^{-1}\|_2 \|\delta H\|_2},$$

RelGap_1, ν_{fr} *are as in* (2.12)*, and* RelGap_2 *is as in* (2.35)*.*

Proof. The proof follows by inserting (2.21) and (2.34) into (2.3). □

Using the estimates (2.22) and (2.37), we can derive a bound similar to (2.38) for any unitary invariant norm $\|\cdot\|$.

COROLLARY 2.10. *Assuming the conditions of Theorem 2.9 and Corollaries 2.5 and 2.8 hold. Using the same notation, we have*

$$\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\| \leq \frac{2\|B\|_2^4\, \nu_{\mathrm{fr}}}{1 - 4\gamma\|B\|_2^2} \frac{\|H^{-1}\|_2 \|\delta H\|}{\frac{\delta}{\sqrt{a(a+\delta)}}} + \frac{\nu_{\mathrm{sec1}}\, \frac{\delta}{\sqrt{|a|+|a+\delta|}} + \nu_{\mathrm{sec2}}}{\frac{\delta}{\sqrt{|a|+|a+\delta|}}} \|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|\,.$$

Proof. The proof can be derived in a similar way as that of Theorem 2.9. □

Since the estimate (2.38) in Theorem 2.9 is rather technical, we propose a simplification which gives a reasonable upper bound for (2.38) under certain restrictions.

Starting from (2.38), we define the function

$$\nu_{sec2}(\eta_M) := \frac{(2 - 3\eta_M)(1 - \eta_M)}{\sqrt{2}(1 - 2\,\eta_M)^2}.$$

It is an increasing function for $\eta_M \in [0, \frac{1}{2}[$. Also, $\nu_{sec2}(0) = \sqrt{2}$, which means that the minimal value of that function on the interval $[0, \frac{1}{2}[$ is $\sqrt{2}$. Since

$$\lim_{\eta_M \to \frac{1}{2}} \nu_{sec2}(\eta_M) = \infty,$$

ETNA
Kent State University
http://etna.math.kent.edu

we may reasonably further restrict the range of η_M. For example, letting $\eta_M < \frac{1}{7}$ yields $\nu_{sec2}(\eta_M) < \frac{33\sqrt{2}}{25}$. It means that $\nu_{sec2}(\eta_M) \in [\sqrt{2}, \frac{33\sqrt{2}}{25}[$ for $\eta_M \in [0, \frac{1}{7}[$, and so if we substitute a constant for $\nu_{sec2}(\eta_M)$, it will not significantly affect the estimate in (2.38). Using similar consideration as above, we state the simpler upper bound for (2.38) in the following corollary.

COROLLARY 2.11. *Let all of the conditions from Theorem 2.9 hold. If in addition*

$$\gamma\|B\|_2^2 < \frac{1}{32} \qquad \textit{and} \qquad \eta_M = \|M^{-1/2}\delta M M^{-1/2}\|_2 < \frac{1}{7},$$

then

$$\|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \le \frac{32\|B\|_2^4}{7\,\mathrm{RelGap}_1}\|H^{-1}\|_2\|\delta H\|_F + \frac{36\,\mathrm{RelGap}_2 + 60}{35\,\mathrm{RelGap}_2}\|M^{-\frac{1}{2}}\delta M M^{-\frac{1}{2}}\|_F,$$

where RelGap_1 *and* RelGap_2 *are defined in* (2.12) *and* (2.35)*, respectively.*

Proof. Recall the relationship (2.20). It follows that

$$\frac{2 - \|H^{-1}\|_2\|\delta H\|_2}{(1 - \|H^{-1}\|_2\|\delta H\|_2)(2 - 3\|H^{-1}\|_2\|\delta H\|_2)}\|H^{-1}\|_2\|\delta H\|_F < 2\|H^{-1}\|_2\|\delta H\|_F\,.$$

Further, using the estimates from [10], it follows that

$$\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\| \le \frac{\|M^{-1/2}\delta M M^{-1/2}\|}{\sqrt{1-\eta_M}},$$
$$\|\widetilde{M}^{-1/2}\delta M \widetilde{M}^{-1/2}\| \le \frac{\|M^{-1/2}\delta M M^{-1/2}\|}{1-\eta_M},$$

and hence,

$$\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\| \le \sqrt{\tfrac{7}{6}}\|M^{-1/2}\delta M M^{-1/2}\|$$

and

$$\|\widetilde{M}^{-1/2}\delta M \widetilde{M}^{-1/2}\| \le \tfrac{7}{6}\|M^{-1/2}\delta M M^{-1/2}\|.$$

Analogously, it can be easily seen that

$$\frac{1-\eta_M}{1-2\eta_M} \le \frac{6}{5} \qquad \text{if} \qquad \eta_M < \frac{1}{7}.$$

Finally, from the above inequities, it follows that $\eta_M < 1/7$ implies $\nu_{\mathrm{sec2}} \le 2$. The bound from the statement of the theorem is now obtained from (2.38). ☐

3. sin Θ- and sin 2Θ-theorems for definite matrix pairs. In this section we present new $\sin\Theta$- and $\sin 2\Theta$-theorems which generalize results form [8] and from the previous section. Here we consider definite matrix pairs (H, M), where H is an indefinite Hermitian matrix which can be factorized as

$$H = GJG^*, \quad J = \mathrm{diag}(\pm 1),$$

G is assumed to be a non-singular matrix, and M is a non-singular indefinite Hermitian matrix. This means that there exists $\alpha \in \mathbb{R}$ such that $M_\alpha := H - \alpha M$ is positive definite. That is, there is a non-singular matrix X such that

$$X^*HX = D_H = \mathrm{diag}(a_1, \dots, a_n) \qquad \text{and} \qquad X^*MX = D_M = \mathrm{diag}(b_1, \dots, b_n).$$

ETNA
Kent State University
http://etna.math.kent.edu

More to the point, we consider the following generalized eigenvector problem

$$Hx = \lambda M x \tag{3.1}$$

and the corresponding perturbed one

$$(H + \delta H)\widetilde{x} = \widetilde{\lambda}(M + \delta M)\widetilde{x},$$

such that the matrix pair $(\widetilde{H}, \widetilde{M})$ is also definite, which means that there exists a non-singular matrix $\widetilde{X}$ such that

$$\widetilde{X}^*\widetilde{H}\widetilde{X} = \widetilde{D}_H = \operatorname{diag}(\widetilde{a}_1, \ldots, \widetilde{a}_n) \quad \text{and} \quad \widetilde{X}^*\widetilde{M}\widetilde{X} = \widetilde{D}_M = \operatorname{diag}(\widetilde{b}_1, \ldots, \widetilde{b}_n). \tag{3.2}$$

One of the criteria by which we can choose α so that $M_\alpha = H - \alpha M$ is positive definite is given in the next theorem.

THEOREM 3.1 ([24, Theorem A1]). *Suppose that the Hermitian matrix pair (H, M) is definite, where $M \in \mathbb{C}^{n\times n}$ is non-singular with m positive eigenvalues and $n - m$ negative eigenvalues. Then there exists a non-singular matrix X such that*

$$X^*HX = \begin{bmatrix} \Lambda_+ & \\ & -\Lambda_- \end{bmatrix} X^*MX = \begin{bmatrix} I_m^* & \\ & -I_{n-m} \end{bmatrix}, \tag{3.3}$$

$$\Lambda_+ = \operatorname{diag}(\lambda_1^+, \ldots, \lambda_m^+), \qquad \Lambda_- = \operatorname{diag}(\lambda_1^-, \ldots, \lambda_{n-m}^-),$$

$$\lambda_1^+ \geq \cdots \geq \lambda_m^+, \qquad \lambda_1^- \geq \cdots \geq \lambda_{n-m}^-.$$

Moreover for $\alpha \in \mathbb{R}$, $H - \alpha M$ is positive definite if and only if

$$\lambda_m^+ > \alpha > \lambda_1^-,$$

with the convention that $\lambda_1^- = -\infty$ when $m = n$ and $\lambda_m^+ = +\infty$ when $m = 0$.

We want to derive $\sin\Theta$- and $\sin 2\Theta$-theorems using results from [8, Theorem 3.4] and the previous section. For this purpose, note that the matrices X and $\widetilde{X}$ from (3.2) and (3.3) simultaneously diagonalize the matrix pairs (H, M_α) and $(\widetilde{H}, \widetilde{M}_\alpha) = (H + \delta H, M_\alpha + \delta M_\alpha)$, respectively, such that

$$X^*HX = D_H \qquad \text{and} \qquad X^*M_\alpha X = X^*HX - \alpha X^*MX = D_H - \alpha D_M =: D_{M_\alpha},$$

where $D_{M_\alpha} = \operatorname{diag}(a_1 - \alpha b_1, \ldots, a_n - \alpha b_n)$ is a diagonal matrix with positive diagonal entries, i.e., $a_i - \alpha b_i > 0$, for all $i = 1, \ldots, n$, and similarly for the perturbed matrix pairs. This means that the eigenvectors of the matrix pairs (H, M) and (H, M_α) span the same eigenspace and similarly for the perturbed matrix pairs.

Instead of the eigenproblem (3.1), we consider

$$Hx = \mu M_\alpha x$$

and the corresponding perturbed one

$$\widetilde{H}\widetilde{x} = \widetilde{\mu}\widetilde{M}_\alpha\widetilde{x},$$

where $\widetilde{H} = \widetilde{G}J\widetilde{G}^*$, $J = \operatorname{diag}(\pm 1)$, $\delta H = \widetilde{H} - H$, and $\delta M_\alpha = \delta H - \alpha\delta M$. Eigenvalues of the matrix pairs (H, M_α) and $(\widetilde{H}, \widetilde{M}_\alpha)$ are of the form $\mu_i = \frac{a_i}{a_i - \alpha b_i}$, $\widetilde{\mu}_i = \frac{\widetilde{a}_i}{\widetilde{a}_i - \alpha\widetilde{b}_i}$, $i = 1, \ldots, n$, respectively. The subspaces remain the same.

Using the result given in [8, Theorem 3.4], we can state our first theorem.

THEOREM 3.2. *Let* (H, M_α) *be a Hermitian pair, and let* $(\widetilde{H}, \widetilde{M_\alpha})$ *be the perturbed pair. Let* $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ *and* $\widetilde{X} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 \end{bmatrix}$ *be non-singular matrices which simultaneously diagonalize* (H, M_α) *and* $(\widetilde{H}, \widetilde{M_\alpha})$*, respectively. If*

$$\eta_{M_\alpha}(\alpha) := \|M_\alpha^{-1/2} \delta M_\alpha M_\alpha^{-1/2}\|_2 < \frac{1}{2}\,,$$

then

$$\begin{aligned} \|\sin\Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F &\le \frac{\|B\|_2 \|\widetilde{B}\|_2}{\mathrm{RelGap}(\alpha)} \|G^{-1} \delta H \widetilde{G}^{-*}\|_F \\ &+ \frac{1}{\mathrm{RGap}(\alpha)} \frac{\sqrt{1-\eta_{M_\alpha}(\alpha)}}{\sqrt{1-2\,\eta_{M_\alpha}(\alpha)}} \|M_\alpha^{-1/2} \delta M_\alpha \widetilde{M}_\alpha^{-1/2}\|_F, \end{aligned} \tag{3.4}$$

where $\sin\Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(X_1))$ *is a diagonal matrix with the sines of the canonical angles in the weighted* M_α*-inner product space between* $\mathrm{Ran}(X_1)$ *and* $\mathrm{Ran}(\widetilde{X}_1)$ *on its diagonal, and where*

$$\mathrm{RelGap}(\alpha) = \min_{\substack{i=k+1,\dots,n \\ j=1,\dots,k}} \frac{|\mu_i - \widehat{\mu}_j|}{\sqrt{|\mu_i||\widehat{\mu}_j|}} = \min_{\substack{i=k+1,\dots,n \\ j=1,\dots,k}} \frac{\left|\frac{a_i}{a_i-\alpha b_i} - \frac{\widetilde{a}_j}{a_j-\alpha b_j}\right|}{\sqrt{\left|\frac{a_i}{a_i-\alpha b_i}\right|\left|\frac{\widetilde{a}_j}{a_j-\alpha b_j}\right|}}$$

and

$$\mathrm{RGap}(\alpha) = \min_{\substack{i=k+1,\dots,n \\ j=1,\dots,k}} \frac{|\widehat{\mu}_i - \widetilde{\mu}_j|}{|\widetilde{\mu}_j|} = \min_{\substack{i=k+1,\dots,n \\ j=1,\dots,k}} \frac{\left|\frac{\widetilde{a}_i}{a_i-\alpha b_i} - \frac{\widetilde{a}_j}{\widetilde{a}_j-\alpha\widetilde{b}_j}\right|}{\left|\frac{\widetilde{a}_j}{\widetilde{a}_j-\alpha\widetilde{b}_j}\right|}\,.$$

Proof. For a proof see that of [8, Theorem 3.4]. ◻

Using the same approach, by applying Theorem 2.9 to the matrix pairs (H, M_α) and $(\widetilde{H}, \widetilde{M_\alpha})$, one can state a bound for the sines of the double angle between the subspaces $\mathrm{Ran}(X_1)$ and $\mathrm{Ran}(\widetilde{X}_1)$. Here we only mention a simplified form of this theorem. Using a similar discussion as in the previous section, in the next corollary we present such estimates. Here it is particularly useful to obtain α-independent bounds for the stability constants $\nu_{\mathrm{sec}2}$ and $\nu_{\mathrm{sec}1}$.

COROLLARY 3.3. *Let all the conditions from Theorem 2.9 hold. If*

$$\gamma \|B\|_2^2 < \frac{1}{32} \qquad \text{and} \qquad \eta_{M_\alpha}(\alpha) < \frac{1}{7},$$

then

$$\begin{aligned} \|\sin 2\Theta_M(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F &\le \frac{32\|B\|_2^4}{7\,\mathrm{RelGap}_1(\alpha)} \|H^{-1}\|_2 \|\delta H\|_F \\ &+ \frac{36\,\mathrm{RelGap}_2(\alpha) + 60}{35\,\mathrm{RelGap}_2(\alpha)} \|M_\alpha^{-1/2} \delta M_\alpha M_\alpha^{-1/2}\|_F, \end{aligned} \tag{3.5}$$

ETNA
Kent State University
http://etna.math.kent.edu

where

$$\mathrm{RelGap}_1(\alpha) = \min_{\substack{i=k+1,\ldots,n \\ j=1,\ldots,k}} \frac{|\widehat{\mu}_i - \widehat{\mu}_j|}{\sqrt{|\widehat{\mu}_i||\widehat{\mu}_j|}} = \min_{\substack{i=k+1,\ldots,n \\ j=1,\ldots,k}} \frac{|\frac{\widetilde{a}_i}{a_i-\alpha b_i} - \frac{\widetilde{a}_j}{a_j-\alpha b_j}|}{\sqrt{|\frac{\widetilde{a}_i}{a_i-\alpha b_i}||\frac{\widetilde{a}_j}{a_j-\alpha b_j}|}},$$

$$\mathrm{RelGap}_2(\alpha) = \min_{\substack{i=k+1,\ldots,n \\ j=1,\ldots,k}} \frac{|\widetilde{\mu}_i - \widetilde{\mu}_j|}{\sqrt{|\widetilde{\mu}_i|^2 + |\widetilde{\mu}_j|^2}} = \min_{\substack{i=k+1,\ldots,n \\ j=1,\ldots,k}} \frac{|\frac{\widetilde{a}_i}{\widetilde{a}_i-\alpha\widetilde{b}_i} - \frac{\widetilde{a}_j}{\widetilde{a}_j-\alpha\widetilde{b}_j}|}{\sqrt{|\frac{\widetilde{a}_i}{\widetilde{a}_i-\alpha\widetilde{b}_i}|^2 + |\frac{\widetilde{a}_j}{\widetilde{a}_j-\alpha\widetilde{b}_j}|^2}}.$$

REMARK 3.4. Note that the bounds (3.4) and (3.5) depend on $\alpha \in (\lambda_1^-, \lambda_m^+)$ from Theorem 3.1. Since both inverses of the relative gaps tend to infinity when α is close to λ_1^- or λ_m^+ and our bound is pessimistic in that case, for α we choose $\alpha = \frac{\lambda_1^- + \lambda_m^+}{2}$. This choice is further justified since in applications one frequently observes that the inverses of the gaps decay very rapidly from the boundary towards the middle of the interval. In particular note that in (3.5), we see that both terms in the estimate essentially depend on the relative gaps. Only the bound for the perturbation in the factor M_α can be optimized. It is, however, assumed to be less than $1/7$. This is a reasonable assumption since for the asymptotic considerations we assume that $\delta H \to 0$ and $\delta M_\alpha \to 0$. This indicates that in optimizing the effectivity of the bounds, we should concentrate on optimizing the relative gaps $\mathrm{RelGap}_1(\alpha)$ and $\mathrm{RelGap}_2(\alpha)$ by increasing them. Our choice furthermore seems to be reasonable for this. This is illustrated in the numerical examples.

4. Numerical examples.

The purpose of this section is to experimentally compare the bounds of several $\sin\Theta$-theorems.

4.1. A family of random matrices.

In this example we consider a family of random perturbations of a *quasi-definite* matrix motivated by considerations of the abstract Bogoliubov-de-Gennes model from [12]. Our results can be applied to this problem, assuming that $M = I$ and $\delta M = 0$. Let

$$H = \begin{bmatrix} H_{11} & \\ & -H_{11} \end{bmatrix}$$

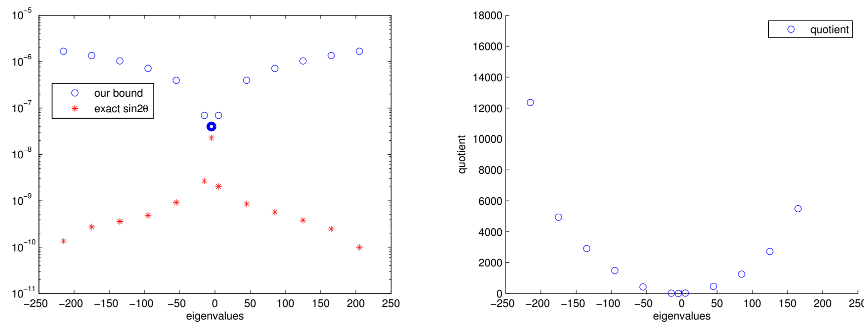
be a *quasi-definite* matrix, and let

$$\widetilde{H} = \begin{bmatrix} H_{11} & \Omega_\omega \\ \Omega_\omega & -H_{11} \end{bmatrix}$$

be a corresponding perturbed matrix. Here H_{11} is a fixed positive definite matrix given as (in MATLAB notation):

```
n=25;
h11=0.01:0.001:0.013;
h11=[h11,10*(1:n-4)];
[Qtmp,temp]=qr(rand(25));
    H11=Qtmp*diag(h11)*Qtmp'+ 5*eye(n);
    H11=1/2*(H11+H11');
```

and $\Omega = \varepsilon \cdot \mathrm{rand}(n)\,, \Omega_\omega = \Omega + \Omega^T$ is a random Hermitian matrix, where $\varepsilon = 10^{-8}$. We are interested in estimating the bound in (2.38) for $k = 1 : 4 : 49$, where k is such that $X_1 = X(:, 1 : k)$ and $X_2 = (:, n - k + 1 : 2n)$. This means that we have to estimate perturbations of an invariant subspace which corresponds to the first k eigenvalues of the



(a) Bound (2.38) and exact value. (b) Effectivity quotient (4.1).

FIG. 4.1. $\varepsilon = 10^{-8}$.

matrix pair (H, M). As an illustration of the performance of our estimator, we study the mapping

$$k \mapsto \frac{\frac{2\|B\|_2^4}{1-4\gamma\|B\|_2^2}\frac{\nu_{\rm fr}\,\|H^{-1}\|_2\|\delta H\|_F}{\mathrm{RelGap}_1} + \frac{\nu_{\rm sec1}\,\mathrm{RelGap}_2 + \nu_{\rm sec2}}{\mathrm{RelGap}_2}\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|_F}{\|\sin\Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F}, \tag{4.1}$$

where X_1 and $\widetilde{X}_1$ span the eigensubspace of the first k eigenvalues (counting from left to right).

The obtained results are presented in Figure 4.1. Figure 4.1(a) displays our bound (2.38) and the exact value. Figure 4.1(b) displays the effectivity quotient, which is the best for $k = 25$ where its value is 1.6916. We see that the bound is dominated by the influence of the relative gaps. Recalling Remark 3.4, we consider this as a further justification of our choice of the parameter α. Also, note that the relative gap does not correctly capture the trend in the perturbation estimate far from the origin. On the other hand, the estimates for a cluster of eigenvalues, which are smallest in the absolute value, are very sharp. This is a known feature of relative estimates which are designed as a sharp tool to study the stability of the inertia of a matrix or matrix pair.

4.2. Analysis of a parameter-dependent family of problems. Let (H, M) be a definite matrix pair. Consider the *quasi-definite* matrix H with

$$H = \begin{bmatrix} 2.0010 & 0.0092 & 0.0071 & 0.4667 & 0.3968 & 0.1007 \\ 0.0092 & 2.0189 & 0.0069 & 0.9746 & 0.4793 & 0.3459 \\ 0.0071 & 0.0069 & 2.0186 & 0.2708 & -0.2937 & -0.5863 \\ 0.4667 & 0.9746 & 0.2708 & -2.0010 & -0.0092 & -0.0071 \\ 0.3968 & 0.4793 & -0.2937 & -0.0092 & -2.0189 & -0.0069 \\ 0.1007 & 0.3459 & -0.5863 & -0.0071 & -0.0069 & 2.0186 \end{bmatrix}.$$

The bounds given here are compared with the bound given in [14, Theorem 5.5], which is of the form

$$\begin{aligned} \|\sin 2\Theta(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F - 2\omega\|X_1\|_2\|W_1\|_2\|\sin\Theta(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F^2 \\ \le \frac{\kappa(X)^3\kappa(\widetilde{X})[1+\kappa(X)^2]\|\widetilde{X}\|_2^2}{\delta}\|\delta H - \lambda\delta M\|_F, \end{aligned} \tag{4.2}$$

ETNA
Kent State University
http://etna.math.kent.edu

where $\omega = \|(W_1^* W_1)^{-1/2} W_1^* W_2 (W_2^* W_2)^{-1/2}\|_2$, $W = X^{-1} = [W_1 \quad W_2]$, $W_1 \in \mathbb{C}^{n \times k}$, $W_2 \in \mathbb{C}^{n \times (n-k)}$, and

$$\delta = \min \left\{ \frac{|\widetilde{\lambda} - \lambda|}{\sqrt{1+\widetilde{\lambda}^2}\sqrt{1+\lambda^2}} ; \lambda \in \sigma(\Lambda_1), \widetilde{\lambda} \in \sigma(\widetilde{\Lambda}_2) \right\}.$$

Note that in (4.2), the sines of the double angle are not bounded directly, and it is necessary to have the information about the sines of the single angles between the subspaces available. Also, for a not so large condition number of the matrix X, this bound can be pessimistic.

Case 1: Let M be the positive definite symmetric matrix given in MATLAB notation as

```
M=diag(1:n)+0.1*rand(n);
M=(M+M')/2;
```

We consider random perturbations δH and δM, which satisfy

$$|(\delta H)_{ij}| \le \eta |H_{ij}|, \qquad |(\delta M)_{ij}| \le \eta |(\delta M)_{ij}|,$$

where $\eta = 10^{-8}$. We derive a bound for $\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}\widetilde{X}_1)\|_F$, where X_1 and $\widetilde{X}_1$ contains eigenvectors corresponding to the two smallest eigenvalues of the matrix pair (H, M). The estimate (2.38) from Theorem 2.9 gives

$$\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}\widetilde{X}_1)\|_F \le 1.3296\text{e-}06\,,$$

in comparison with the exact value of $\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}\widetilde{X}_1)\| \approx 3.6229\text{e-}08$. Also, the estimate (4.2) gives $\|\sin 2\Theta_M(\operatorname{Ran}(X_1), \operatorname{Ran}\widetilde{X}_1)\|_F \le 1.1173\text{e-}03$, which is three orders of magnitude larger than our bound.

Case 2: Let M be the real symmetric indefinite matrix given in MATLAB notation as

```
M=-eye(n)+0.1*rand(n);
M=(M+M')/2;
M(3,3)=0.01;
```

The matrix pair (H, M) is definite, which means that there exists an α such that $H - \alpha M = M_\alpha$ is a positive definite matrix. From Theorem 3.1, it follows that the matrix $M_\alpha = H - \alpha M$ is positive definite for $\alpha \in (2.6729, 75.1301)$. From now on, let (H, M_α) be a symmetric definite pair. We consider random perturbations δH and $\delta M_\alpha = \delta H - \alpha \delta M$, which satisfy

$$|(\delta H)_{ij}| \le \eta |H_{ij}|, \qquad |(\delta M_\alpha)_{ij}| \le \eta |\delta (M_\alpha)_{ij}|,$$

where $\eta = 10^{-8}$. In the following experiment, we estimate the perturbation of an invariant subspace which corresponds to the two smallest eigenvalues of the matrix pair (H, M_α). Figure 4.2 and Figure 4.3 display the bounds (3.4) and (3.5), respectively, and also the effectivity quotients for α chosen as (in MATLAB notation)

```
alpha=(2.6729e+000 + 0.00001):0.1:(7.5130e+001-0.00001)
```

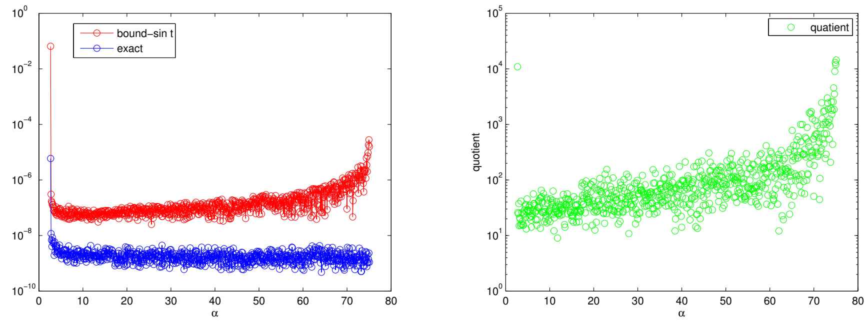
The minimal value of the function (bound (3.4)) shown in Figure 4.2(a) is 2.5213e-08, thus, (3.4) yields

$$\|\sin \Theta_{M_\alpha}(\operatorname{Ran}(X_1), \operatorname{Ran}(\widetilde{X}_1)\|_F \le 2.5213\text{e-}08,$$

in comparison with the exact value $\|\sin \Theta_{M_\alpha}(\operatorname{Ran}(X_1), \operatorname{Ran}\widetilde{X}_1)\|_F \approx 1.5665\text{e-}09$. Also, the minimal value for the effectivity quotient shown in Figure 4.2(b) is 9.0242.

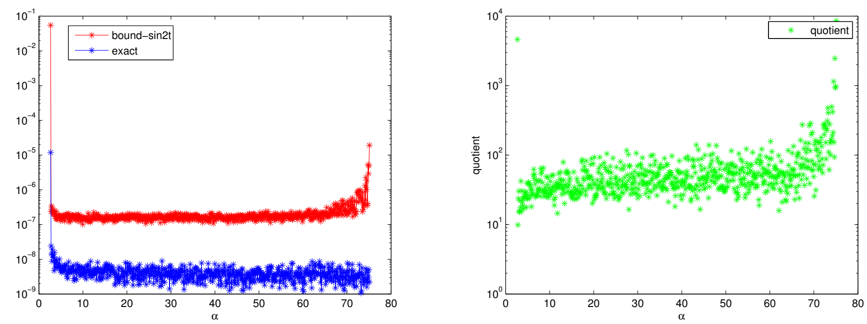
The minimum of the function (the bound in (3.5)) shown in Figure 4.3(a) is 1.0225e-07, hence, the bound (3.5) in Theorem 2.9 gives

$$\|\sin 2\Theta_{M_\alpha}(\operatorname{Ran}(X_1), \operatorname{Ran}(\widetilde{X}_1))\|_F \le 1.0225\text{e-}07\,,$$



(a) Single angle. (b) Effectivity quotient (4.1).

FIG. 4.2. *Bound* (3.4) *for* $\alpha \in (\lambda_1^-, \lambda_m^+)$.



(a) Double angle. (b) Effectivity quotient (4.1).

FIG. 4.3. *Bound* (3.5) *for* $\alpha \in (\lambda_1^-, \lambda_m^+)$.

in comparison with the exact value $\| \sin 2\Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \approx 3.7249\text{e-}09$. The minimal value for the effectivity quotient shown in Figure 4.3(b) is 9.8767. The minimum of the function (the bound (4.2)) is 2.0671e-02, and the maximum is 2.3900e+20, which is a consequence of the dependence of the bounds on the condition numbers of the matrices X and $\widetilde{X}$ from (3.2). This is illustrated in Figure 4.4.

As we have previously pointed out in Remark 3.4, our estimates depend on the parameter α, and from Figure 4.2 and Figure 4.3, it is easy to notice that our bounds are worse for the case when α is close to the edges of interval $(2.6729, 75.1301)$ since the matrix M_α is close to being singular there. We observe that $\alpha = \frac{\lambda_1^- + \lambda_m^+}{2} = 38.9015$ is close to the optimal value of the effectivity quotient. The experiments furthermore suggest that with this choice of α, we improve the sharpness of the estimates (measured by the effectivity quotient of the bounds). To this end we compute

$$\| \sin \Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \leq 6.1705\text{e-}08$$

and

$$\| \sin 2\Theta_{M_\alpha}(\mathrm{Ran}(X_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \leq 1.1696\text{e-}07\,.$$

Note that these bounds are of the same order of magnitude as the minimal values of the functions shown in Figure 4.2 and Figure 4.3.

ETNA
Kent State University
http://etna.math.kent.edu

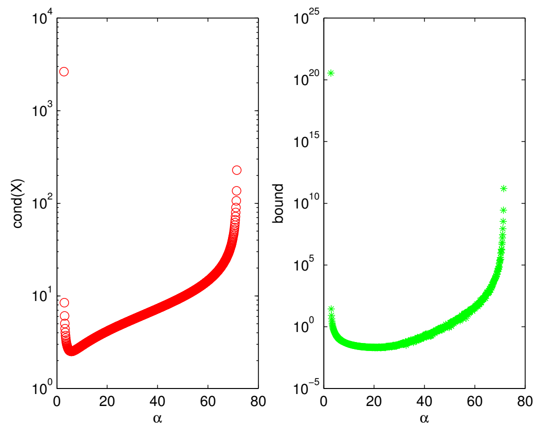


FIG. 4.4. *Condition of the non-singular matrix X from* (3.2) *and bound* (4.2) *for* $\alpha \in (\lambda_1^-, \lambda_m^+)$.

5. Conclusion. In this paper we have presented estimates for the double angle operator associated to the rotation of invariant subspaces of a definite Hermitian pair under the influence of a perturbation of both factors. The angle operator has been defined relative to the scalar product in which the matrix pair is definite. We have obtained, as it is characteristic for double angle theorems, a subspace perturbation estimate in which only the separation of the spectral components of the perturbed matrix pairs appears. As a byproduct, we also obtained bounds for the condition number of the J-unitary matrices B which diagonalize a quasi-definite matrix H. The norm of B is a measure of the reliability of the spectral calculus for H—which yields relative gaps—and convenient for establishing measures of spectral stability, e.g., the relative gap to the unwanted component of the spectrum, for a targeted group of eigenvalues. Numerical experiments confirm that the new bounds are sharper compared to other results in the literature, and they offer the possibility for optimizing the effectivity by the choice of the appropriate scalar product in which to measure the subspace rotation. Investigation of the choice of an appropriate scalar product for a particular application will be the subject of further research.

Acknowledgement. We would like to thank both referees for a careful reading of the manuscript and for their valuable comments and remarks.

This work has been supported in part by the Croatian Science Foundation (HRZZ) under the project "Optimization of parameter dependent mechanical systems" (IP-2014-09-9540), Grant Nr. 9540. Luka Grubišić was supported in part by the Croatian Science Foundation (HRZZ) grant “Mathematical modelling, analysis and computing with applications to complex mechanical systems”, Grant Nr. 9345.

Appendix A. Proof of Theorem 2.7. Before we start with the proof of Theorem 2.7, in the next remark we state some useful facts.

REMARK A.1. From (2.25) it is easy to see that one of the possibilities to choose K_{11} and K_{22} in (2.26) is

$$K_{11} = \sqrt{I - \widetilde{X}_1^* \delta M \widetilde{X}_1} \qquad \text{and}$$

$$K_{22} = \sqrt{I - \widetilde{X}_2^* \delta M \widetilde{X}_2 - K_{21} K_{21}^*}\,. \tag{A.1}$$

ETNA
Kent State University
http://etna.math.kent.edu

Also, $I - \widetilde{X}_2^* \delta M \widetilde{X}_2 = K_{21}K_{21}^* + K_{22}K_{22}^*$. Then by a Cholesky decomposition we find that $I - \widetilde{X}_2^* \delta M \widetilde{X}_2 = K_{21}K_{21}^* + K_{22}K_{22}^* = C_{22}C_{22}^*$, where we can assume that

$$C_{22} = \sqrt{I - \widetilde{X}_2^* \delta M \widetilde{X}_2}, \tag{A.2}$$

and a direct computations proves that the columns of the matrix $\widetilde{Y}_2 C_{22}^{-*}$ are M-orthogonal.

Proof of Theorem 2.7. From (2.30) it follows that

$$\begin{aligned} \| \sin 2\Theta_M(\mathrm{Ran}(\widehat{X}_1), \mathrm{Ran}(\widetilde{X}_1))\|_F \leq & \|K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{Y}_1^* M \widetilde{X}_1 K_{11}^{-*}\|_F \\ & + \|K_{22}^{-1}\widetilde{Y}_2^* M \widetilde{X}_1 K_{11}^{-*}\|_F. \end{aligned} \tag{A.3}$$

As indicated by the above inequality, a bound for $\| \sin 2\Theta_M(\mathrm{Ran}(\widehat{X}_1), \mathrm{Ran}(\widetilde{X}_1))\|_F$ can be obtained by the sum of the bounds for $\|K_{11}^{-1}K_{21}\widetilde{Y}_1^* M \widetilde{X}_1 K_{11}^{-*}\|_F$ and $\|K_{22}^{-1}\widetilde{Y}_2^* M \widetilde{X}_1 K_{11}^{-*}\|_F$, thus the proof of this theorem contains two parts.

First let us derive a bound for $\|K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{Y}_1^* M \widetilde{X}_1 K_{11}^{-*}\|_F$. Using (2.31) it follows that

$$\|K_{22}^{-1}K_{21}K_{11}^{-1}\widetilde{Y}_1^* M \widetilde{X}_1 K_{11}^{-*}\|_F = \|K_{22}^{-1}K_{21}CC + K_{22}^{-1}K_{21}SS\|_F = \|K_{22}^{-1}K_{21}\|_F\,,$$

where $C = K_{11}^{-1}\widetilde{X}_1^* M \widehat{X}_1$ and $S = -K_{11}^{-1}\widetilde{X}_1^* M \widehat{X}_2$, which is easy to conclude from (2.28) and (2.29).

It remains to derive a bound for $\|K_{22}^{-1}K_{21}\|_F$. Note that

$$\|K_{22}^{-1}K_{21}\|_F \leq \|K_{22}^{-1}\|_2 \|K_{21}K_{11}^*\|_F \|K_{11}^{-1}\|_2. \tag{A.4}$$

In (A.4), an estimate for $\|K_{11}^{-1}\|_2$ and also for $\|K_{22}^{-1}\|_2$ is provided in the proof of [8, Theorem 3.4]:

$$\|K_{11}^{-1}\|_2 \leq \frac{\sqrt{1-\eta_M}}{\sqrt{1-2\eta_M}} \qquad \text{and} \qquad \|K_{22}^{-1}\|_2 \leq \frac{\sqrt{1-\eta_M}}{\sqrt{1-2\eta_M}}\,, \tag{A.5}$$

where η_M is defined in (2.33).

From (2.25) it is easy to see that $K_{21}K_{11}^* = -\widetilde{X}_2 \delta M \widetilde{X}_1$. Using this and the fact that $\widetilde{X}$ is an $\widetilde{M}$-orthogonal matrix, one can write

$$\|K_{21}K_{11}\|_F = \|\widetilde{X}_2^* \widetilde{M}^{1/2} \widetilde{M}^{-1/2} \delta M \widetilde{M}^{-1/2} \widetilde{M}^{1/2} \widetilde{X}_1\|_F.$$

Since the columns of the matrices $\widetilde{X}_2^* \widetilde{M}^{1/2}$ and $\widetilde{M}^{1/2}\widetilde{X}_1$ are orthogonal, we conclude that

$$\|K_{21}K_{11}\|_F = \|\widetilde{M}^{-1/2} \delta M \widetilde{M}^{-1/2}\|_F. \tag{A.6}$$

Now, inserting (A.5) and (A.6) into (A.4), it follows that

$$\|K_{22}^{-1}K_{21}\|_F \leq \frac{1-\eta_M}{1-2\eta_M} \|\widetilde{M}^{-1/2} \delta M \widetilde{M}^{-1/2}\|_F. \tag{A.7}$$

In the second part of the proof, we derive an upper bound for $\|K_{22}^{-1}\widetilde{Y}_2^* M \widetilde{X}_1 K_{11}^{-*}\|_F$. Note that

$$\|K_{22}^{-1}\widetilde{Y}_2^* M \widetilde{X}_1 K_{11}^{-*}\|_F \leq \|K_{22}^{-1}\|_2 \|\widetilde{Y}_2^* M \widetilde{X}_1\|_F \|K_{11}^{-1}\|_2\,.$$

Bounds for $\|K_{11}^{-1}\|_2$ and $\|K_{22}^{-1}\|_2$ are given in (A.5). It remains to estimate $\|\widetilde{Y}_2^* M \widetilde{X}_1\|_F$. From (2.23), it holds that

$$\widetilde{H}\widetilde{X}_1 = \widetilde{M}\widetilde{X}_1\widetilde{\Lambda}_1. \tag{A.8}$$

Multiplying (A.8) from the left with $\widetilde{Y}_2^*$ and using that $\widetilde{Y}_2^*\widetilde{H} = \widetilde{\Lambda}_2\widetilde{Y}_2^*\widehat{M}$, we have

$$\widetilde{\Lambda}_2\widetilde{Y}_2^*\widehat{M}\widetilde{X}_1 = \widetilde{Y}_2^*\widetilde{M}\widetilde{X}_1\widetilde{\Lambda}_1.$$

Since $\widehat{M} = M + T_r^*\delta M T_r$ and $\widetilde{M} = M + \delta M$, the previous equation can be rewritten in the form

$$\widetilde{\Lambda}_2\widetilde{Y}_2^* M\widetilde{X}_1 - \widetilde{Y}_2^* M\widetilde{X}_1\widetilde{\Lambda}_1 = -\widetilde{\Lambda}_2\widetilde{Y}_2^* T_r^*\delta M T_r\widetilde{X}_1 + \widetilde{Y}_2^*\delta M\widetilde{X}_1\widetilde{\Lambda}_1.$$

Also, we can use the fact that $T_r\widetilde{Y}_1 = \widetilde{X}_1$, $\widetilde{Y}_2 = T_r\widetilde{X}_2$, and then obtain the structured Sylvester equation

$$\widetilde{\Lambda}_2\widetilde{Y}_2^* M\widetilde{X}_1 - \widetilde{Y}_2^* M\widetilde{X}_1\widetilde{\Lambda}_1 = -\widetilde{\Lambda}_2\widetilde{X}_2^*\delta M\widetilde{Y}_1 + \widetilde{Y}_2^*\delta M\widetilde{X}_1\widetilde{\Lambda}_1. \tag{A.9}$$

In particular, the first part of the right-hand side of (A.9) can be rewritten as

$$-\widetilde{\Lambda}_2\widetilde{X}_2^*\delta M\widetilde{Y}_1 = -\widetilde{\Lambda}_2\widetilde{X}_2^*\widetilde{M}^{1/2}\widetilde{M}^{-1/2}\delta M M^{-1/2} M^{1/2}\widetilde{Y}_1 K_{11}^{-*}K_{11}^*.$$

Note that the matrices $\widetilde{X}_2^*\widetilde{M}^{1/2} =: Q_2^*$ and $M^{1/2}\widetilde{Y}_1 K_{11}^{-*} =: Z_1$ have orthogonal columns. The second part of the left-hand side in (A.9) can be expressed as

$$\widetilde{Y}_2^*\delta M\widetilde{X}_1\widetilde{\Lambda}_1 = C_{22}C_{22}^{-1}\widetilde{Y}_2^* M^{1/2} M^{-1/2}\delta M\widetilde{M}^{-1/2}\widetilde{M}^{1/2}\widetilde{X}_1\widetilde{\Lambda}_1.$$

Also, the matrices $\widetilde{M}^{1/2}\widetilde{X}_1 =: Q_1$ and $M^{1/2}\widetilde{Y}_2 C_{22}^{-*} =: Z_2$ have orthogonal columns. Then (A.9) reads as

$$\begin{aligned}\widetilde{\Lambda}_2\widetilde{Y}_2^* M\widetilde{X}_1 - \widetilde{Y}_2^* M\widetilde{X}_1\widetilde{\Lambda}_1 &= -\widetilde{\Lambda}_2 Q_2^*\widetilde{M}^{-1/2}\delta M M^{-1/2} Z_1 K_{11}^* \\ &\quad + C_{22} Z_2^* M^{-1/2}\delta M\widetilde{M}^{-1/2} Q_1\widetilde{\Lambda}_1,\end{aligned} \tag{A.10}$$

which is also a Sylvester equation with a structured right-hand side. Applying [16, Lemma 2.2] to (A.10), we obtain

$$\|\widetilde{X}_2^* M\widetilde{Y}_1\|_F \le \frac{\sqrt{\|K_{11}^*\|_2^2 + \|C_{22}\|_2^2}\,\|\widetilde{M}^{-1/2}\delta M M^{-1/2}\|_F}{\mathrm{RelGap}_2}, \tag{A.11}$$

where RelGap_2 is defined in (2.35).

It remains to determine a bound for $\sqrt{\|K_{11}\|_2^2 + \|C_{22}\|_2^2}$, where K_{11} and C_{22} are defined by (A.1) and (A.2), respectively. Let us just estimate $\|K_{11}\|_2^2$; in a similar way, the results for $\|C_{22}\|_2^2$ are obtained.

We assume that $\|\widetilde{X}_1^*\delta M\widetilde{X}_1\|_2 < 1$, which ensures the existence of $(I - \widetilde{X}_1^*\delta M\widetilde{X}_1)^{1/2}$ defined by the following series from [11, Theorem 6.2.8]:

$$K_{11} = (I - \widetilde{X}_1^*\delta M\widetilde{X}_1)^{1/2} = I - \sum_{i=1}^{\infty}\frac{(2n-1)!!}{2^n\cdot n!}(\widetilde{X}_1^*\delta M\widetilde{X}_1)^n, \tag{A.12}$$

where $(2n-1)!! = 1 \cdot 3 \cdot 5 \cdots (2n-1)$. Write $K_{11} = I + \Gamma$. From (A.12), we have

$$\|\Gamma\|_2 \le \sum_{i=1^\infty} \frac{(2n-1)!!}{2^n \cdot n!} \|\widetilde{X}_1^* \delta M \widetilde{X}_1\|_2^n$$

$$\le \frac{1}{2} \frac{\|\widetilde{X}_1^* \delta M \widetilde{X}_1\|_2}{1 - \|\widetilde{X}_1^* \delta M \widetilde{X}_1\|_2} \le \frac{1}{2} \frac{\|\widetilde{X}^* \delta M \widetilde{X}\|_2}{1 - \|\widetilde{X}^* \delta M \widetilde{X}\|_2}. \tag{A.13}$$

Using the $\widetilde{M}$-orthogonality of the matrix $\widetilde{X}$, it can easily be seen that the matrices $\widetilde{X}$ and $M^{-1/2}(I + M^{-1/2}\delta M M^{-1/2})^{-1/2}$ are unitary similar, that is, there exist a unitary matrix Q such that

$$\widetilde{X} = M^{-1/2}(I + M^{-1/2}\delta M M^{-1/2})^{-1/2} Q. \tag{A.14}$$

Set $V = M^{-1/2}\delta M M^{-1/2}$, then from (A.14) it follows that

$$\|\widetilde{X}^* \delta M \widetilde{X}\|_2 = \|(I+V)^{-1/2} V (I+V)^{-1/2}\|_2 \le \frac{\eta_M}{1-\eta_M}. \tag{A.15}$$

Now, inserting (A.15) into (A.13), we have

$$\|\Gamma\|_2 \le \frac{1}{2} \frac{\eta_M}{1 - 2\eta_M}. \tag{A.16}$$

Using (A.16) it follows that

$$\|K_{11}\|_2 \le \frac{2 - 3\eta_M}{2 - 4\eta_M} \tag{A.17}$$

and similarly for $\|C_{22}\|_2$. Using (A.17) we establish

$$\sqrt{\|K_{11}\|_2^2 + \|C_{22}\|_2^2} \le \sqrt{2} \frac{2 - 3\eta_M}{2 - 4\eta_M}. \tag{A.18}$$

Now insert (A.18) into (A.11) to obtain a bound for $\|\widetilde{Y}_2^* M \widetilde{X}_1\|_F$. Finally, the proof simply follows by inserting (A.7), (A.11), (A.5) into (A.3). ◻

REFERENCES

[1] S. ALBEVERIO AND A. K. MOTOVILOV, *Sharpening the norm bound in the subspace perturbation theory*, Complex Anal. Oper. Theory, 7 (2013), pp. 1389–1416.

[2] S. ALBEVERIO, A. K. MOTOVILOV, AND A. V. SELIN, *The a priori tan Θ theorem for eigenvectors*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 685–697.

[3] J. BARLOW AND J. DEMMEL, *Computing accurate eigensystems of scaled diagonally dominant matrices*, SIAM J. Numer. Anal., 27 (1990), pp. 762–791.

[4] J. L. BARLOW AND I. SLAPNIČAR, *Optimal perturbation bounds for the Hermitian eigenvalue problem*, Linear Algebra Appl., 309 (2000), pp. 19–43.

[5] R. BHATIA AND R. C. LI, *On perturbations of matrix pencils with real spectra. II*, Math. Comp., 65 (1996), pp. 637–645.

[6] C. DAVIS AND W. M. KAHAN, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.

[7] S. C. EISENSTAT AND I. C. F. IPSEN, *Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices*, BIT, 38 (1998), pp. 502–509.

[8] L. GRUBIŠIĆ, N. TRUHAR, AND S. MIODRAGOVIĆ, *The rotation of eigenspaces of perturbed matrix pairs II*, Linear Multilinear Algebra, 62 (2014), pp. 1010–1031.

[9] L. GRUBIŠIĆ, N. TRUHAR, AND K. VESELIĆ, *The rotation of eigenspaces of perturbed matrix pairs*, Linear Algebra Appl., 436 (2012), pp. 4161–4178.

[10] L. GRUBIŠIĆ AND K. VESELIĆ, *On weakly formulated Sylvester equations and applications*, Integral Equations Operator Theory, 58 (2007), pp. 175–204.

[11] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[12] W. KIRSCH, B. METZGER, AND P. MÜLLER, *Random block operators*, J. Stat. Phys., 143 (2011), pp. 1035–1054.

[13] A. V. KNYAZEV AND M. E. ARGENTATI, *Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates*, SIAM J. Sci. Comput., 23 (2002), pp. 2008–2040

[14] R.-C. LI, *On perturbations of matrix pencils with real spectra*, Math. Comp., 62 (1994), pp. 231–265.

[15] ———, *Relative perturbation theory. III. More bounds on eigenvalue variation*, Linear Algebra Appl., 266 (1997), pp. 337–345.

[16] ———, *Relative perturbation theory. II. Eigenspace and singular subspace variations*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 471–492.

[17] ———, *Relative perturbation theory. IV.* $\sin 2\theta$ *theorems*, Linear Algebra Appl., 311 (2000), pp. 45–60.

[18] A. PUSHNITSKI, *The scattering matrix and the differences of spectral projections*, Bull. Lond. Math. Soc., 40 (2008), pp. 227–238.

[19] ———, *Operator theoretic methods for the eigenvalue counting function in spectral gaps*, Ann. Henri Poincaré, 10 (2009), pp. 793–822.

[20] A. SEELMANN, *Notes on the* $\sin 2\Theta$ *theorem*, Integral Equations Operator Theory, 79 (2014), pp. 579–597.

[21] N. TRUHAR AND R.-C. LI, *A sin* 2Θ *theorem for graded indefinite Hermitian matrices*, Linear Algebra Appl., 359 (2003), pp. 263–276.

[22] N. TRUHAR AND S. MIODRAGOVIĆ, *Relative perturbation theory for definite matrix pairs and hyperbolic eigenvalue problem*, Appl. Numer. Math., 98 (2015), pp. 106–121.

[23] N. TRUHAR AND I. SLAPNIČAR, *Relative perturbation bound for invariant subspaces of graded indefinite Hermitian matrices*, Linear Algebra Appl., 301 (1999), pp. 171–185.

[24] K. VESELIĆ, *A Jacobi eigenreduction algorithm for definite matrix pairs*, Numer. Math., 64 (1993), pp. 241–269.

[25] ———, *Perturbation theory for the eigenvalues of factorised symmetric matrices*, Linear Algebra Appl., 309 (2000), pp. 85–102.