Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

SYMPLECTIC EIGENVALUES OF POSITIVE-SEMIDEFINITE MATRICES AND THE TRACE MINIMIZATION THEOREM*

NGUYEN THANH SON† AND TATJANA STYKEL‡

Abstract. Symplectic eigenvalues are conventionally defined for symmetric positive-definite matrices via Williamson's diagonal form. Many properties of standard eigenvalues, including the trace minimization theorem, have been extended to the case of symplectic eigenvalues. In this note, we will generalize Williamson's diagonal form for symmetric positive-definite matrices to the case of symmetric positive-semidefinite matrices, which allows us to define symplectic eigenvalues, and prove the trace minimization theorem in the new setting.

Key words. Symplectic eigenvalues, Positive-semidefinite matrices, Trace minimization.

AMS subject classifications. 15A15, 15A18, 70G45.

1. Introduction.

Let us first recall that a matrix $S \in \mathbb{R}^{2n\times 2n}$ is referred to as a *symplectic matrix* if it satisfies the relation:

$$S^T J_{2n} S = J_{2n} \quad \text{with} \quad J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},$$

where I_n denotes the $n \times n$ identity matrix. The set of such matrices forms a multiplicative group [10] and is denoted by $\mathrm{Sp}(2n)$. We also employ the set of rectangular symplectic matrices defined as:

$$\mathrm{Sp}(2k, 2n) = \{S \in \mathbb{R}^{2n\times 2k} \,:\, S^T J_{2n} S = J_{2k}\},$$

for some k with $1 \le k \le n$. In [11, Proposition 3.1], this set was shown to be an embedded Riemannian manifold of $\mathbb{R}^{2n\times 2k}$.

Derived from Williamson's work [28], for a $2n \times 2n$ symmetric positive-definite (SPD) real matrix A, there exists a matrix $S \in \mathrm{Sp}(2n)$ such that

$$S^T A S = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix}, \tag{1.1}$$

where $D = \mathrm{diag}(d_1, \ldots, d_n)$ is a diagonal matrix with positive diagonal entries. The equality (1.1) looks very like the eigenvalue decomposition of SPD matrices except for the fact that the transformation matrix S is symplectic instead of orthogonal. The right-hand side of (1.1) is termed as *Williamson's diagonal form* of A. The positive numbers $d_1, \ldots, d_n$ are referred to as *symplectic eigenvalues*. They form the *symplectic spectrum* of A. The symplectic spectrum of an SPD matrix is unique, while the so-called *diagonalizing*

*Received by the editors on August 16, 2022. Accepted for publication on September 9, 2022. Handling Editor: Heike Fassbender. Corresponding Author: Nguyen Thanh Son

†Institute of Mathematics and Centre for Advanced Analytics and Predictive Sciences (CAAPS), University of Augsburg, Universitätsstraße 12a, 86159 Augsburg, Germany, and Thai Nguyen University of Sciences, 24118 Thai Nguyen, Vietnam (ntson@tnus.edu.vn).

‡Institute of Mathematics and Centre for Advanced Analytics and Predictive Sciences (CAAPS), University of Augsburg, Universitätsstraße 12a, 86159 Augsburg, Germany (tatjana.stykel@math.uni-augsburg.de).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

matrix S in (1.1) is not unique. The set of diagonalizing matrices for a given SPD matrix was characterized via a symplectic version of Autonne's uniqueness theorem in [27, Theorem 3.5].

The symplectic eigenvalues of A are strongly related to the standard eigenvalues of the Hamiltonian matrix $J_{2n}A$ or the symmetric/skew-symmetric matrix pencil $A-\lambda J_{2n}$, and the Hermitian matrix $\mathrm{i}A^{1/2}J_{2n}A^{1/2}$ with $\mathrm{i}=\sqrt{-1}$ and $A^{1/2}$ being the symmetric square root of A or the Hermitian pencil $A-\lambda \mathrm{i}J_{2n}$, which have been intensively investigated in [18, 24, 1, 20, 22, 25, 23, 17]. They find wide applications in quantum mechanics, optics, and structured mechanical systems [14, 2, 19]. Symplectic eigenvalues and eigenvectors, defined below, can be numerically computed using a symplectic Lanczos method via the connection with positive-definite Hamiltonian matrices [1] or by solving a trace minimization problem using a Riemannian optimization method [11, 12, 27].

Many properties of symplectic eigenvalues for SPD matrices have been investigated. Besides trace minimization mentioned above, these are Cauchy's interlacing theorem [4], variational principles and Weyl inequality [6], Schur-Horn theorem [5], Szegő-type theorem [3], and Lidskii's (or the majorization) theorem [17], just to name a few.

It is noteworthy that in the original work [28], A is not necessarily SPD but solely symmetric. It would be a reasonable question to ask whether it is still possible to define symplectic eigenvalues and investigate their properties for a more general case in which A is symmetric positive-semidefinite (SPSD). Such matrices arise, for example, in stability analysis of gyroscopic systems with an SPSD stiffness matrix, see, e.g., [21, Section 2], and in quantization process of superconducting networks modeled by Hamiltonian systems with positive-semidefinite Hamiltonians [7, 9]. It turns out that not every SPSD matrix is appropriate for this purpose and even in the case that symplectic eigenvalues are definable, some properties that are already known for symplectic eigenvalues of SPD matrices are nontrivially extendable to the new setting. In this note, we will set the condition under which an SPSD matrix enjoys Williamson's diagonal form which enables the definition of symplectic eigenvalues. Moreover, we prove the trace minimization theorem for symplectic eigenvalues of such matrices which can be exploited for not only computing symplectic eigenpairs [27] but also deriving the majorization inequality similar to [14, Theorem 1]. Our proof is based on considering positive-semidefinite real pencils studied in [18]. In the case of SPD matrices, a similar result has been independently derived in [23] by exploiting the properties of positive-definite complex pencils discussed in [22].

For ease of presentation, we now collect some necessary materials following [8, 17, 27]. The Kronecker delta is defined as:

$$\delta_{ij}=\begin{cases}1, & \text{if } i=j,\\ 0, & \text{if } i\neq j.\end{cases}$$

The set of all $2n\times 2n$ SPSD real matrices will be denoted by $\mathcal{SPSD}(2n)$. Given a square matrix M, $\mathrm{tr}(M)$ denotes the trace of M; the *kernel* or the *null space* of M is the set $\ker M=\{x : Mx=0\}$; a *square root* of M is a square matrix K of the same size such that $K^2=M$; M^T denotes the transpose of real M and M^* is the complex conjugate and transpose of complex M. Furthermore, $\mathrm{span}\{v_1,\ldots,v_\ell\}$ stands for the subspace spanned by the vectors $v_1,\ldots,v_\ell$; applied to a matrix, it means the subspace spanned by the columns of this matrix. Given $A_j\in\mathbb{R}^{n_j\times n_j}, j=1,\ldots,\ell$, $\mathrm{diag}(A_1,\ldots,A_\ell)$ denotes the (block) diagonal matrix with A_j on the diagonal. A subset $U\subset\mathbb{R}^{2n}$ is said to be *symplectically orthogonal to* a subset $V\subset\mathbb{R}^{2n}$ if $u^TJ_{2n}v=0$ for all $u\in U, v\in V$. Two pairs of vectors (u_1,v_1) and (u_2,v_2) in $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$ are said to be *symplectically normalized* if

$$u_i^TJ_{2n}v_j=\delta_{ij}\quad\text{and}\quad u_i^TJ_{2n}u_j=v_i^TJ_{2n}v_j=0\quad\text{for}\quad i,j=1,2.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

Note that this definition is slightly different from that in [27]. The symplecticity of the matrix $[x_1, \ldots, x_{2k}]$ is equivalent to the fact that each two pairs (x_i, x_{k+i}) and (x_j, x_{k+j}) are symplectically normalized for $i, j = 1, \ldots, k$ and $i \neq j$. It was shown in [8, Theorem 1.15] that a symplectically normalized set of vectors are linear-independent. Therefore, when $k = n$, such vectors form a basis of $\mathbb{R}^{2n}$. In that case, it is called a *symplectic basis*. A subspace $\mathbb{W} \subset \mathbb{R}^{2n}$ is called a *symplectic subspace* if for any $x \in \mathbb{W}\backslash\{0\}$, there exists $y \in \mathbb{W}$ such that $x^T J_{2n} y = 1$. It turns out that we can always construct a symplectic basis for a symplectic subspace [8, Theorem 1.15]. Conversely, a given subspace is symplectic if it owns a symplectic basis as shown in Lemma 2.1 below. Two nonzero vectors written as a matrix $[u, v] \in \mathbb{R}^{2n\times 2}$ form a *symplectic eigenvector pair* corresponding to the *symplectic eigenvalue* d of $A \in \mathcal{SPSD}(2n)$ if it holds

$$Au = dJ_{2n}v, \quad Av = -dJ_{2n}u. \tag{1.2}$$

The set $([u, v], d)$ is then generally referred to as a *symplectic eigenpair*. The existence of Williamson's diagonal form for any given $2n \times 2n$ SPD matrix tells us that there exists a symplectic basis of $\mathbb{R}^{2n}$ consisting of its symplectic eigenvectors.

2. Williamson's diagonal form of symmetric positive-semidefinite matrices. In this section, we will extend the definition of symplectic eigenvalues to the case of SPSD matrices by explicitly constructing Williamson's diagonal forms. For further discussion, we will need the following technical property.

LEMMA 2.1. *If a subspace $\mathbb{W} \subset \mathbb{R}^{2n}$ admits a symplectic basis, then it is a symplectic subspace.*

Proof. Assume that the columns of $[w_1, \ldots, w_m, w_{m+1}, \ldots, w_{2m}]$ with $1 \leq m \leq n$ form a symplectic basis of $\mathbb{W}$. Then, any $x \in \mathbb{W}\backslash\{0\}$ can be represented via this basis as $x = \sum_{j=1}^{m}(\lambda_j w_j + \lambda_{m+j} w_{m+j})$ with $\sum_{j=1}^{2m} \lambda_j^2 \neq 0$. Setting

$$y = \frac{\sum\limits_{j=1}^{m}(\lambda_j w_{m+j} - \lambda_{m+j} w_j)}{\sum\limits_{j=1}^{2m} \lambda_j^2},$$

we can verify that $x^T J_{2n} y = 1$. This implies that the subspace $\mathbb{W}$ is symplectic. ∎

For a moment, we consider an SPD matrix A which has Williamson's diagonal form (1.1). Multiplying (1.1) with $S^{-T} = J_{2n} S J_{2n}^T$ from the left and with $S^{-1} = J_{2n}^T S^T J_{2n}$ from the right, we derive

$$A = J_{2n} S \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} (J_{2n} S)^T.$$

Then, for $0 \leq m < n$, we set

$$\hat{A} = J_{2n} S \begin{bmatrix} \hat{D}_{m+1:n} & 0 \\ 0 & \hat{D}_{m+1:n} \end{bmatrix} (J_{2n} S)^T,$$

where $\hat{D}_{m+1:n} = \mathrm{diag}(0, \ldots, 0, d_{m+1}, \ldots, d_n) \in \mathbb{R}^{n\times n}$. We reverse the above multiplications to obtain

$$\hat{A} S = J_{2n} S \begin{bmatrix} 0 & -\hat{D}_{m+1:n} \\ \hat{D}_{m+1:n} & 0 \end{bmatrix},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

and

$$S^T \hat{A} S = \begin{bmatrix} \hat{D}_{m+1:n} & 0 \\ 0 & \hat{D}_{m+1:n} \end{bmatrix}. \tag{2.3}$$

Several facts can be drawn from the above construction. First, $\hat{A}$ is an SPSD matrix of rank $2n-2m$. Second, if we write S by its columns as $S = [s_1, \ldots, s_{2n}]$, then the kernel of $\hat{A}$ is $\operatorname{span}\{s_1, \ldots, s_m, s_{n+1}, \ldots, s_{n+m}\}$. Moreover, it is a symplectic subspace of dimension $2m$ due to Lemma 2.1. And third, it is symplectically orthogonal to the subspace $\operatorname{span}\{s_{m+1}, \ldots, s_n, s_{n+m+1}, \ldots, s_{2n}\}$ which is nothing else but the range of $\hat{A}$.

Obviously, the decomposition (2.3) is like (1.1), and therefore it can be thought of as Williamson's diagonal form of the SPSD matrix $\hat{A}$. This gives us a clear vision of how to construct Williamson's diagonal form of SPSD matrices and under which condition this is possible. In practice, we will proceed somewhat in the reverse direction. Indeed, given $A \in \mathcal{SPSD}(2n)$ whose kernel is a symplectic space. As any symplectic subspace must have even dimension [8, Sect. 1.1], we assume that $\dim(\ker A) = 2m$ with $0 \le m < n$. Let $W = [W_1\ W_2]$ with $W_1,\ W_2 \in \mathbb{R}^{2n\times m}$ be a symplectic matrix whose columns form a symplectic basis of $\ker A$. Note that such a basis can be computed using the symplectic Gram-Schmidt process [8, Theorem 1.15]. We adapt the procedure proposed in [25] for constructing Williamson's diagonal form of SPD matrices to the more general case where the matrix A is SPSD as follows.

First, it is worth to recall that any SPSD matrix has a unique SPSD square root, denoted by $A^{1/2}$, and that $\ker A^{1/2} = \ker A$ [15, Theorem 7.2.6]. Next, we will show that $\ker(A^{1/2}J_{2n}A^{1/2}) = \ker A^{1/2}$. It is trivial that $\ker(A^{1/2}J_{2n}A^{1/2}) \supseteq \ker A^{1/2}$. Now, consider any $y \in \mathbb{R}^{2n}$ such that $A^{1/2}J_{2n}A^{1/2}y = 0$. Assume that $J_{2n}A^{1/2}y \neq 0$. Because $\ker A^{1/2}$ is a symplectic subspace, there exists $z \in \ker A^{1/2}$ such that $z^T J_{2n}J_{2n}A^{1/2}y = 1$ which leads to $y^T A^{1/2}z = -1$. This contradicts to the fact that $z \in \ker A^{1/2}$ and hence $J_{2n}A^{1/2}y = 0$. As J_{2n} is nonsingular, we deduce that $A^{1/2}y = 0$. This implies the desired equality. Observe moreover that the matrix $A^{1/2}J_{2n}A^{1/2}$ is skew-symmetric. Then, this matrix can be transformed into the real Schur form:

$$Q^T A^{1/2} J_{2n} A^{1/2} Q = \operatorname{diag}\left(0_{2\times 2}, \ldots, 0_{2\times 2}, \begin{bmatrix} 0 & d_{m+1} \\ -d_{m+1} & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & d_n \\ -d_n & 0 \end{bmatrix}\right), \tag{2.4}$$

where $0 < d_{m+1} \le \cdots \le d_n$ and $Q \in \mathbb{R}^{2n\times 2n}$ is an orthogonal matrix, see, e.g., [13, Theorem 7.4.1]. After that, we multiply both sides of (2.4) with the permutation matrix $P = [\mathrm{e}_1, \mathrm{e}_3, \ldots, \mathrm{e}_{2n-1}, \mathrm{e}_2, \mathrm{e}_4, \ldots, \mathrm{e}_{2n}]$ from the right and with P^T from the left, where e_j denotes the j-th canonical unit vector in $\mathbb{R}^{2n}$. As a result, we obtain

$$P^T Q^T A^{1/2} J_{2n} A^{1/2} Q P = \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix},$$

where $D = \operatorname{diag}(0, \ldots, 0, d_{m+1}, \ldots, d_n)$. Finally, setting

$$\tilde{S} = J_{2n}A^{1/2}QP \begin{bmatrix} 0_{m\times(n-m)} & 0_{m\times(n-m)} \\ 0_{(n-m)\times(n-m)} & -\tilde{D}_{m+1:n}^{-1/2} \\ 0_{m\times(n-m)} & 0_{m\times(n-m)} \\ \tilde{D}_{m+1:n}^{-1/2} & 0_{(n-m)\times(n-m)} \end{bmatrix} =: [\tilde{S}_1\ \tilde{S}_2] \in \mathbb{R}^{2n\times(2n-2m)}, \tag{2.5}$$

where $\tilde{D}_{m+1:n}^{-1/2} = \operatorname{diag}(1/\sqrt{d_{m+1}}, \ldots, 1/\sqrt{d_n})$ and $\tilde{S}_1, \tilde{S}_2 \in \mathbb{R}^{2n\times(n-m)}$, we define $S = [W_1\ \tilde{S}_1\ W_2\ \tilde{S}_2]$. The following theorem shows that S is a symplectic matrix which diagonalizes A.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

THEOREM 2.2. *Given a matrix $A \in \mathcal{SPSD}(2n)$ whose kernel is a symplectic subspace of $\mathbb{R}^{2n}$ of dimension $2m$. Then the matrix S constructed as above is symplectic and it holds*

$$S^T A S = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix}, \tag{2.6}$$

where $D = \mathrm{diag}(d_1, \ldots, d_n)$ with $0 = d_1 = \cdots = d_m < d_{m+1} \leq \cdots \leq d_n$.

Proof. It follows from $\ker A^{1/2} = \ker A = \mathrm{span}(W)$ that $A^{1/2}W = 0$. Therefore, in view of (2.5), $W^T J_{2n} \tilde{S} = 0$. Furthermore, taking into account the fact that the matrices W and $\tilde{S}$ are both symplectic, we deduce that S is symplectic.

The diagonal form (2.6) follows directly from the equalities $\tilde{S}_i^T A \tilde{S}_j = \delta_{ij}\mathrm{diag}(d_{m+1}, \ldots, d_n)$, $AW_j = 0$, and $W_j^T A = 0$ for $i, j = 1, 2$. ∎

In view of the connection of $d_j, j = 1, \ldots, n$, determined in (2.6) with the standard eigenvalues of Hamiltonian matrix $J_{2n}A$, see [27, Proposition 3.2] and references therein, we can conclude that $d_j, j = 1, \ldots, n$, are unique. This fact gives rise to the following definition.

DEFINITION 2.3. *For $A \in \mathcal{SPSD}(2n)$ with a symplectic kernel, the right-hand side of (2.6) is called* Williamson's diagonal form *of A. The nonnegative diagonal elements in this form $0 \leq d_1 \leq \cdots \leq d_n$ are called the* symplectic eigenvalues *of A.*

It immediately follows from (2.6) that the symplecticity of the kernel of A is also a necessary condition for the existence of Williamson's diagonal form of A. The SPSD matrices that do not satisfy this condition are not difficult to find. Consider, for instance, $A = \mathrm{diag}(I_n, 0_{n\times n})$. Then $\ker A = \mathrm{span}\{\mathrm{e}_{n+1}, \ldots, \mathrm{e}_{2n}\}$. One can directly verify that $x^T J_{2n} y = 0$ for all $x, y \in \ker A$, which means that $\ker A$ is an isotropic subspace [8, Section 1.2]. In fact, all SPSD matrices whose kernel is an isotropic subspace do not have Williamson's diagonal form.

The observation that owning a symplectic kernel is a necessary and sufficient condition for the existence of Williamson's diagonal form of an SPSD matrix was previously made in [17, Remark 2.6]. A justification for this observation can be drawn from [17, Proposition 5.2] where one of the symplectic projections is uniquely determined by the columns of W. Our proof here is however more constructive.

For the sake of convenience, in the rest of this paper, we restrict ourselves to the case of SPSD matrices that satisfy the symplectic kernel condition mentioned in Theorem 2.2.

3. Trace minimization theorem. The trace minimization theorem for the symplectic eigenvalues of SPD matrices was established in [14, 4]. This theorem was then exploited in [27] to compute symplectic eigenpairs of SPD matrices. A popular way of justifying such a statement is to use minimax principles; see [26, 24] for standard eigenvalues of Hermitian matrix pencils and [4] for symplectic eigenvalues of SPD matrices. In [4], the derived minimax principles, however, invoke the reciprocals of symplectic eigenvalues. This approach is certainly inapplicable to the case of SPSD matrices. In this section, we will prove the trace minimization theorem for symplectic eigenvalues in a more general setting where A can be SPSD. Note that restricted to the case of SPD matrices, our proof differs from the ones presented in [14, 4] and rather relates to the results of [23, Appendix A2]. For completeness, we restate the trace minimization theorem in a more general form in which the symplectic eigenvalues can be zero.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

THEOREM 3.1. *Let $A \in \mathcal{SPSD}(2n)$ have symplectic eigenvalues $0 \le d_1 \le \cdots \le d_n$. Then for any k with $1 \le k \le n$, it holds that*

$$\min_{X \in \mathrm{Sp}(2k,2n)} \operatorname{tr}(X^T A X) = 2\sum_{j=1}^{k} d_j. \tag{3.7}$$

For a proof of this result, we need some preparation. We start with reformulating the definition of symplectic eigenvalues and eigenvectors. Recalling (1.2), $[u, v] \in \mathbb{R}^{2n\times 2}$ is a symplectic eigenvector pair corresponding to the symplectic eigenvalue d of A if and only if one of the following four equivalent conditions holds

$$\begin{aligned}
\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} &= d \begin{bmatrix} 0 & J_{2n} \\ -J_{2n} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \\
\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} &= -d \begin{bmatrix} 0 & J_{2n} \\ -J_{2n} & 0 \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix}, \\
\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} v \\ -u \end{bmatrix} &= d \begin{bmatrix} 0 & J_{2n} \\ -J_{2n} & 0 \end{bmatrix} \begin{bmatrix} v \\ -u \end{bmatrix}, \\
\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} -u \\ v \end{bmatrix} &= -d \begin{bmatrix} 0 & J_{2n} \\ -J_{2n} & 0 \end{bmatrix} \begin{bmatrix} -u \\ v \end{bmatrix}.
\end{aligned} \tag{3.8}$$

This fact suggests constructing the matrix pencil:

$$\mathcal{A} - \lambda \mathcal{J} := \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} - \lambda \begin{bmatrix} 0 & J_{2n} \\ -J_{2n} & 0 \end{bmatrix}, \tag{3.9}$$

with the following properties.

LEMMA 3.2. *Let $A \in \mathcal{SPSD}(2n)$ and $\mathcal{A} - \lambda\mathcal{J}$ be the matrix pencil as in (3.9). Then the following statements hold:*

1. *The real number d is a symplectic eigenvalue of A if and only if d and $-d$ are eigenvalues of multiplicities two of the matrix pencil $\mathcal{A} - \lambda\mathcal{J}$.*
2. *The matrix $\mathcal{A}$ is SPSD and the matrix $\mathcal{J}$ is symmetric and nonsingular.*
3. *The matrix $\mathcal{J}$ possesses only 1 and -1 as its eigenvalues. Moreover, both of them are of multiplicity $2n$.*

Proof. The first statement follows immediately from (3.8), while the second one is straightforward. To justify the last one, let X_1 and X_2 be any two basis matrices of $\mathbb{R}^n$. Then it is direct to show that the matrices

$$\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \\ 0 & -X_2 \\ X_1 & 0 \end{bmatrix}, \quad \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \\ 0 & X_2 \\ -X_1 & 0 \end{bmatrix},$$

consist of eigenvectors of $\mathcal{J}$ associated with eigenvalues 1 and -1, respectively. ∎

The relations in (3.8) together with Lemma 3.2 establish a connection between symplectic eigenpairs of A and eigenpairs of the matrix pencil (3.9). Hence, the result on trace minimization of positive-semidefinite

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

matrix pencils derived in [18] is here helpful. Let us first collect the necessary facts from there. Given two $\ell \times \ell$ symmetric real matrices A and B, the matrix pencil $A - \lambda B$ is said to be *positive-semidefinite* if there is a number $\lambda_0 \in \mathbb{R}$ such that $A - \lambda_0 B$ is positive-semidefinite. Such constant λ_0 is called a *definitizing shift*. Obviously, the matrix pencil $\mathcal{A} - \lambda\mathcal{J}$ in (3.9) is positive-semidefinite with a shift $\lambda_0 = 0$.

Furthermore, it was shown in [18, Proposition 4.1] that the eigenvalues of the positive-semidefinite matrix pencil $A - \lambda B$ with nonsingular B are real, and the number of non-positive (resp. non-negative) eigenvalues of $A - \lambda B$ is the same as the number of negative (resp. positive) eigenvalues of B. In other words, if q and p with $q + p = \ell$ is the number of negative and positive eigenvalues of B, respectively, then there are q non-positive eigenvalues of $A - \lambda B$, namely, $\alpha_q^- \leq \cdots \leq \alpha_1^- \leq 0$, and p non-negative eigenvalues of $A - \lambda B$, namely, $0 \leq \alpha_1^+ \leq \cdots \leq \alpha_p^+$. Thus, they can overall be arranged as:

$$\alpha_q^- \leq \cdots \leq \alpha_1^- \leq \alpha_1^+ \leq \cdots \leq \alpha_p^+.$$

The following theorem provides an important result on trace minimization for positive-semidefinite matrix pencils.

THEOREM 3.3. [18, Theorem 3.1] *Let A, $B \in \mathbb{R}^{\ell\times\ell}$ be symmetric and let B be nonsingular with p positive and q negative eigenvalues. Assume moreover that the matrix pencil $A - \lambda B$ is positive-semidefinite. Let p_1 and q_1 be two integers such that $0 \leq p_1 \leq p$, $0 \leq q_1 \leq q$, and $0 < p_1 + q_1$. Then, the function*

$$\mathbb{R}^{\ell\times(p_1+q_1)} \ni \ X \mapsto \operatorname{tr}(X^T A X) \in \mathbb{R}, \tag{3.10}$$

subjected to the condition:

$$X^T B X = \operatorname{diag}(I_{p_1}, -I_{q_1}), \tag{3.11}$$

is bounded from below by:

$$\sum_{j=1}^{p_1} \alpha_j^+ - \sum_{j=1}^{q_1} \alpha_j^-. \tag{3.12}$$

If, additionally, there exists a matrix $X_0 \in \mathbb{R}^{\ell\times(p_1+q_1)}$ satisfying (3.11) and consisting of eigenvectors of the matrix pencil $A - \lambda B$ associated with the eigenvalues $\alpha_1^+, \ldots, \alpha_{p_1}^+, \alpha_1^-, \ldots, \alpha_{q_1}^-$, then the lower bound (3.12) becomes the minimal value of the function (3.10) which is reached at X_0.

Now, we are ready to prove the trace minimization Theorem 3.1 for symplectic eigenvalues of SPSD matrices.

Proof of Theorem 3.1. Consider the matrix pencil $\mathcal{A} - \lambda\mathcal{J}$ as in (3.9). By Theorem 2.2 and Lemma 3.2, the eigenvalues of $\mathcal{A} - \lambda\mathcal{J}$, sorted in the nondecreasing order, are as follows:

$$-d_n = -d_n \leq \cdots \leq -d_1 = -d_1 \leq d_1 = d_1 \leq \cdots \leq d_n = d_n,$$

where $d_1, \ldots, d_n$ are the symplectic eigenvalues of A. Thus, an application of Theorem 3.3 to the matrix pencil $\mathcal{A} - \lambda\mathcal{J}$ for $p_1 = q_1 = 2k$ yields

$$\operatorname{tr}(\mathcal{X}^T \mathcal{A} \mathcal{X}) \geq 4 \sum_{j=1}^{k} d_j, \tag{3.13}$$

for all $\mathcal{X} \in \mathbb{R}^{4n\times 4k}$ satisfying the constraint:

$$\mathcal{X}^T \mathcal{J} \mathcal{X} = \operatorname{diag}(I_{2k}, -I_{2k}). \tag{3.14}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

Next, for any $X \in \mathbb{R}^{2n\times 2k}$, we construct the matrix:

$$\mathcal{X} = \frac{1}{\sqrt{2}} \begin{bmatrix} X & XJ_{2k}^T \\ XJ_{2k}^T & X \end{bmatrix} \in \mathbb{R}^{4n\times 4k}. \tag{3.15}$$

This matrix fulfills

$$\begin{aligned} \operatorname{tr}(\mathcal{X}^T\mathcal{A}\mathcal{X}) &= \operatorname{tr}(X^TAX + J_{2k}X^TAXJ_{2k}^T) \\ &= \operatorname{tr}(X^TAX + X^TAXJ_{2k}^TJ_{2k}) \\ &= 2\operatorname{tr}(X^TAX). \end{aligned} \tag{3.16}$$

Moreover, taking $X \in \operatorname{Sp}(2k, 2n)$, one can verify by a direct calculation that the matrix $\mathcal{X}$ in (3.15) satisfies (3.14). Then, using (3.16) and (3.13), we obtain that

$$\min_{X\in\operatorname{Sp}(2k,2n)} \operatorname{tr}(X^TAX) = \frac{1}{2} \min_{\substack{\mathcal{X} \text{ as in (3.15)},\\ X \in \operatorname{Sp}(2k,2n)}} \operatorname{tr}(\mathcal{X}^T\mathcal{A}\mathcal{X}) \tag{3.17}$$

$$\geq \frac{1}{2} \min_{\tilde{\mathcal{X}}^T\mathcal{J}\tilde{\mathcal{X}}=\operatorname{diag}(I_{2k},-I_{2k})} \operatorname{tr}(\tilde{\mathcal{X}}^T\mathcal{A}\tilde{\mathcal{X}}) \tag{3.18}$$

$$\geq 2\sum_{j=1}^{k} d_j,$$

in which (3.18) is due to the fact that the constraint set in the right-hand side of (3.17) is a subset of that in (3.18). Finally, we choose $X = [s_1, \ldots, s_k, s_{n+1}, \ldots, s_{n+k}]$ consisting of (symplectically) normalized symplectic eigenvector pairs $[s_j, s_{n+j}]$ associated with the symplectic eigenvalue d_j of A for $j = 1, \ldots, k$. In this case, we have $\operatorname{tr}(X^TAX) = 2\sum_{j=1}^{k} d_j$ which completes the proof. ∎

Note that since $\{0\}$ is itself a (trivial) symplectic subspace, Theorem 3.1 holds for SPD matrices too. For such matrices, our proof is actually related to that in [23] via the embedding of $2n \times 2n$ complex matrices into $4n \times 4n$ real matrices. An extension of the trace minimization theorem to the SPSD case has been obtained in [23] by considering an SPD perturbation $A+\epsilon I$ and taking the limit as $\epsilon \to 0$. However, utilizing Williamson's normal form (2.6), we can directly verify (3.7) for SPSD A by exploiting the spectral properties of the Hermitian pencil $A - \lambda \mathrm{i} J_{2n}$ similar to [23]. Indeed, consider the Kronecker product $\tilde{Q} = Q_2 \otimes I_k$ with a unitary matrix:

$$Q_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{bmatrix}.$$

Clearly, $\tilde{Q}$ is also unitary, and for any $X \in \operatorname{Sp}(2k, 2n)$, we have

$$(X\tilde{Q})^* \,\mathrm{i}J_{2n}X\tilde{Q} = \tilde{Q}^* \,\mathrm{i}J_{2k}\tilde{Q} = \operatorname{diag}(I_k, -I_k).$$

In particular, setting $Y = [s_1, \ldots, s_k, s_{n+1}, \ldots, s_{n+k}]\tilde{Q}$, where s_j, $j = 1, \ldots, k, n+1, \ldots, n+k$, are the columns of the diagonalizing symplectic matrix S in (2.6), one can show that the columns of Y are eigenvectors associated with the eigenvalues $d_1, \ldots, d_k, -d_1, \ldots, -d_k$ of the Hermitian pencil $A - \lambda \mathrm{i} J_{2n}$ and that Y satisfies the constraint $Y^*\mathrm{i}J_{2n}Y = \operatorname{diag}(I_k, -I_k)$. Then, we obtain

$$\min_{X\in\operatorname{Sp}(2k,2n)} \operatorname{tr}(X^TAX) = \min_{\tilde{X}^*\mathrm{i}J_{2n}\tilde{X}=\operatorname{diag}(I_k,-I_k)} \operatorname{tr}(\tilde{X}^*A\tilde{X}) = 2\sum_{j=1}^{k} d_j,$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

where the first equality holds by taking $\tilde{X} = X\tilde{Q}$ and the second minimum is attained at Y due to the complex extension of Theorem 3.3, see [22, Theorem 1.1]. Obviously, the real matrix $X = [s_1, \ldots, s_k, s_{n+1}, \ldots, s_{n+k}]$ solves the first trace minimization problem with the real constraint.

Remark 3.4. For the computation of the smallest symplectic eigenvalues of the SPSD matrix A, a direct use of the symplectic Lanczos procedure [1] will not work because it involves the inverse of A which does not exist. On the other side, as the theoretical results in [27] still hold for SPSD matrices, Theorem 3.1 plays a key role in enabling the use of the numerical method proposed there for computing the required symplectic eigenpairs.

Acknowledgment. The authors would like to thank the anonymous referee for the constructive comments and the editor for making possible to access [23].

REFERENCES

[1] P. Amodio. On the computation of few eigenvalues of positive definite Hamiltonian matrices. *Future Gener. Comput. Syst.*, 22:403–411, 2006.
[2] P. Benner, H. Fassbender, and M. Stoll. Solving large-scale quadratic eigenvalue problems with Hamiltonian eigenstructure using a structure-preserving Krylov subspace method. *Electron. Trans. Numer. Anal.*, 29:212–229, 2008.
[3] R. Bhatia and T. Jain. A Szegő type theorem and distribution of symplectic eigenvalues. *J. Spectr. Theory*, 11:1369–1389, 2011.
[4] R. Bhatia and T. Jain. On the symplectic eigenvalues of positive definite matrices. *J. Math. Phys.*, 56:112201, 2015.
[5] R. Bhatia and T. Jain. A Schur-Horn theorem for symplectic eigenvalues. *Linear Algebra Appl.*, 599:133–139, 2020.
[6] R. Bhatia and T. Jain. Variational principles for symplectic eigenvalues. *Canad. Math. Bull.*, 64:553–559, 2020.
[7] F.J. Buijnsters, A. Fasolino, and M.I. Katsnelson. Zero modes in magnetic systems: General theory and an efficient computational scheme. *Phys. Rev. B*, 89:174433, 2014.
[8] M. de Gosson. *Symplectic Geometry and Quantum Mechanics.* Advances in Partial Differential Equations. Birkhäuser, Basel, 2006.
[9] I.L. Egusquiza and A. Parra-Rodziguez. Algebraic canonical quantization of lumped superconducting networks. *Phys. Rev. B*, 106:024510, 2022.
[10] S. Fiori. Solving minimal-distance problems over the manifold of real-symplectic matrices. *SIAM J. Matrix Anal. Appl.*, 32:938–968, 2011.
[11] B. Gao, N.T. Son, P.-A. Absil, and T. Stykel. Riemannian optimization on the symplectic Stiefel manifold. *SIAM J. Optim.*, 31:1546–1575, 2021.
[12] B. Gao, N.T. Son, P.-A. Absil, and T. Stykel. Geometry of the symplectic Stiefel manifold endowed with the Euclidean Metric. In: F. Nielsen and F. Barbaresco (editors), *Geometric Science of Information*, GSI 2021. Lecture Notes in Computer Science, vol. 12829. Springer, Cham, 789–798, 2021.
[13] G. Golub and C.V. Loan. *Matrix Computations, 4th ed.* The Johns Hopkins University Press, Baltimore, London, 2013.
[14] T. Hiroshima. Additivity and multiplicativity properties of some Gaussian channels for Gaussian inputs. *Phys. Rev. A*, 73:012330, 2006.
[15] R.A. Horn and C.R. Johnson. *Matrix Analysis, 2nd ed.* Cambridge University Press, New York, NY, 2013.
[16] M. Idel, S.S. Gaona, and M.M. Wolf. Perturbation bounds for Williamson's symplectic normal form. *Linear Algebra Appl.*, 525:45–58, 2017.
[17] T. Jain and H.K. Mishra. Derivatives of symplectic eigenvalues and a Lidskii type theorem. *Canad. J. Math.*, 74:457–485, 2022.
[18] J. Kovač-Striko and K. Veselić. Trace minimization and definiteness of symmetric pencils. *Linear Algebra Appl.*, 216:139–158, 1995.
[19] M. Krbek, T. Tyc, and J. Vlach. Inequalities for quantum marginal problems with continuous variables. *J. Math. Phys.*, 55:062201, 2014.
[20] P. Lancaster and R. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Rev.*, 47:407–443, 2006.
[21] P. Lancaster. Stability of linear gyroscopic systems: A review. *Linear Algebra Appl.*, 439:686–706, 2013.
[22] X. Liang, R.-C. Li, and Z. Bai. Trace minimization principles for positive semi-definite pencils. *Linear Algebra Appl.*, 438:3085–3106, 2013.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 607-616, September 2022.

[23] X. Liang, L. Wang, L.-H. Zhang, and R.-C. Li. On generalizing trace minimization. Submitted to *Linear Algebra Appl.*, 2021.

[24] I. Nakić and K. Veselić. Wielandt and Ky-Fan theorem for matrix pairs. *Linear Algebra Appl.*, 369:77–93, 2003.

[25] K. Parthasarathy. The symplectry group of Gaussian states in $L^2(\mathbb{R}^n)$. In: A. Shiryaev, S. Varadhan, and E. Presman (editors), *Prokhorov and Contemporary Probability Theory*, vol. 33. Springer Proceedings in Mathematics and Statistics, Berlin Heidelberg, 349–369, 2013.

[26] A.H. Sameh, and J.A. Wisniewski. A trace minimization algorithm for generalized eigenvalue problem. *SIAM J. Numer. Anal.*, 19:1243–1259, 1982.

[27] N.T. Son, P.-A. Absil, B. Gao, and T. Stykel. Computing symplectic eigenpairs of symmetric positive-definite matrices via trace minimization and Riemannian optimization. *SIAM J. Matrix Anal. Appl.*, 42:1732–1757, 2021.

[28] J. Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *Amer. J. Math.*, 58:141–163, 1936.