

Properties of Definite Bethe–Salpeter Eigenvalue Problems

Meiyue Shao and Chao Yang

Abstract The Bethe–Salpeter eigenvalue problem is solved in condense matter physics to estimate the absorption spectrum of solids. It is a structured eigenvalue problem. Its special structure appears in other approaches for studying electron excitation in molecules or solids also. When the Bethe–Salpeter Hamiltonian matrix is definite, the corresponding eigenvalue problem can be reduced to a symmetric eigenvalue problem. However, its special structure leads to a number of interesting spectral properties. We describe these properties that are crucial for developing efficient and reliable numerical algorithms for solving this class of problems.

1 Introduction

Discretization of the *Bethe–Salpeter equation* (BSE) [15] leads to an eigenvalue problem $H z = \lambda z$, where the coefficient matrix H has the form

$$H = \begin{bmatrix} A & B \\ -\bar{B} & -\bar{A} \end{bmatrix}. \quad (1)$$

The matrix A and B in (1) satisfy

$$A^* = A, \quad \bar{B}^* = B. \quad (2)$$

Here A^* and \bar{A} are the conjugate transpose and complex conjugate of A , respectively. In this paper, we call H a *Bethe–Salpeter Hamiltonian matrix*, or, in short, a BSE Hamiltonian. In condense matter physics, the Bethe–Salpeter eigenvalue problem is derived from a Dyson’s equation for a 2-particle Green’s function used to describe excitation events that involve two particles simultaneously. It is a special case of the J -symmetric eigenvalue problem [3]. This type of eigenvalue problem also appears in linear response (LR) time-dependent density functional theory, and the random

M. Shao • C. Yang (✉)

Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

e-mail: myshao@lbl.gov; cyang@lbl.gov

© Springer International Publishing AG 2017

T. Sakurai et al. (eds.), *Eigenvalue Problems: Algorithms, Software and Applications in Petascale Computing*, Lecture Notes in Computational Science and Engineering 117, https://doi.org/10.1007/978-3-319-62426-6_7

91

phase approximation theory. In these approaches, H is sometimes called a Casida Hamiltonian, a linear response Hamiltonian, or a random phase approximation (RPA) Hamiltonian.

The dimension of A and B can be quite large, because it scales as $\mathcal{O}(N^2)$, where N is number of degrees of freedom required to represent a three-dimensional single particle wavefunction. As a result, efficient numerical algorithms must be developed to solve the Bethe–Salpeter eigenvalue problem. To gain computational efficiency, these methods should take advantage of the special structure of the Hamiltonian in (1).

Let

$$C_n = \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix}, \quad \Omega = \begin{bmatrix} A & B \\ \bar{B} & \bar{A} \end{bmatrix}. \quad (3)$$

Then $H = C_n \Omega$, with both C_n and Ω Hermitian. In most physics problems, the condition

$$\Omega > 0 \quad (4)$$

holds, that is, the matrix Ω is positive definite. We call H a *definite* Bethe–Salpeter Hamiltonian matrix when (4) is satisfied. It has been shown in [16] that, in general, solving a Bethe–Salpeter eigenvalue problem is equivalent to solving a real Hamiltonian eigenvalue problem. However, a definite Bethe–Salpeter eigenvalue problem, which is of most interest in practice, has many additional properties. In this paper we restrict ourselves to this special case, i.e., we assume that the condition (4) holds.

There are several ways to reformulate the definite Bethe–Salpeter eigenvalue problem. One equivalent formulation of $Hz = \lambda z$ yields a generalized eigenvalue problem (GEP) $C_n z = \lambda^{-1} \Omega z$. As Ω is positive definite, $C_n z = \lambda^{-1} \Omega z$ is a Hermitian–definite GEP and hence has real eigenvalues. Another equivalent formulation is $(\Omega - \lambda C_n)z = 0$, where $\Omega - \lambda C_n$ is a *definite pencil* [7, 20] with a *definitizing shift* $\lambda_0 = 0$. In addition, the eigenvalue problem $H^2 z = \lambda^2 z$ can be written as a product eigenvalue problem $(C_n \Omega C_n) \Omega z = \lambda^2 z$ in which both $C_n \Omega C_n$ and Ω are positive definite. These formulations suggest that a definite Bethe–Salpeter eigenvalue problem can be transformed to symmetric eigenvalue problems. As a result, we can analyze various properties of the Bethe–Salpeter eigenvalue problem by combining existing theories of symmetric eigenvalue problems (see, e.g., [14, 20]) with the special structure of H .

In this paper, we describe several spectral properties of a definite BSE Hamiltonian. These properties include the orthogonality of eigenvectors, the Courant–Fischer type of min–max characterization of the eigenvalues, the Cauchy type interlacing properties, and the Weyl type inequalities for establishing bounds on a structurally perturbed definite BSE Hamiltonian. Most properties take into account the special structure of the BSE Hamiltonian. Although the derivations are relatively

straightforward, these properties are important for developing efficient and reliable algorithms for solving the definite Bethe–Salpeter eigenvalue problem.

The rest of this paper is organized as follows. In Sect. 2, we analyze the spectral decomposition of H and derive two types of orthogonality conditions on the eigenvectors. Variational properties based on these two types of orthogonality conditions are established in Sect. 3. Finally, we provide several eigenvalue perturbation bounds in Sect. 4.

2 Preliminaries

2.1 Spectral Decomposition

As a highly structured matrix, a definite BSE Hamiltonian admits a structured spectral decomposition as stated in the following theorem.

Theorem 1 ([16, Theorem 3]) *A definite Bethe–Salpeter Hamiltonian matrix is diagonalizable and has real spectrum. Furthermore, it admits a spectral decomposition of the form*

$$H = \begin{bmatrix} X & \bar{Y} \\ Y & \bar{X} \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{bmatrix} \begin{bmatrix} X & -\bar{Y} \\ -Y & \bar{X} \end{bmatrix}^*, \tag{5}$$

where $\Lambda = \text{diag} \{ \lambda_1, \dots, \lambda_n \} \succ 0$, and

$$\begin{bmatrix} X & -\bar{Y} \\ -Y & \bar{X} \end{bmatrix}^* \begin{bmatrix} X & \bar{Y} \\ Y & \bar{X} \end{bmatrix} = I_{2n}. \tag{6}$$

As the eigenvalues of a definite BSE Hamiltonian appear in pairs $\pm\lambda$, we denote by $\lambda_i^+(U)$ ($\lambda_i^-(U)$) the i th smallest *positive* (largest *negative*) eigenvalue of a matrix U with real spectrum. When the matrix is omitted, λ_i^+ (or λ_i^-) represents $\lambda_i^+(H)$ (or $\lambda_i^-(H)$), where H is a definite BSE Hamiltonian. Thus the eigenvalues of H are labeled as

$$\lambda_n^- \leq \dots \leq \lambda_1^- < \lambda_1^+ \leq \dots \leq \lambda_n^+.$$

To represent the structure of the eigenvectors of H , we introduce the notation

$$\phi(U, V) := \begin{bmatrix} U & \bar{V} \\ V & \bar{U} \end{bmatrix},$$

where U and V are matrices of the same size. The structure is preserved under summation, real scaling, complex conjugation, transposition, as well as matrix multiplication

$$\phi(U_1, V_1)\phi(U_2, V_2) = \phi(U_1U_2 + \bar{V}_1V_2, V_1U_2 + \bar{U}_1V_2).$$

The Eqs. (5) and (6) can be rewritten as

$$C_n\Omega = \phi(X, Y)C_n\phi(\Lambda, 0)\phi(X, -Y)^*, \quad \phi(X, -Y)^*\phi(X, Y) = I_{2n}.$$

The converse of Theorem 1 is also true in the following sense: If (5) holds, then H is a definite BSE Hamiltonian because

$$H = C_n[\phi(X, -Y)\phi(\Lambda, 0)\phi(X, -Y)^*].$$

As a result, $H^{-1} = C_n[\phi(X, -Y)\phi(\Lambda^{-1}, 0)\phi(X, -Y)^*]$ is also a definite BSE Hamiltonian.

2.2 Orthogonality on the Eigenvectors

From the spectral decomposition of a definite BSE Hamiltonian H , we immediately obtain two types of orthogonality conditions on the eigenvectors of H .

First, the fact $\Omega = \phi(X, -Y)\phi(\Lambda, 0)\phi(X, -Y)^*$ implies that

$$\phi(X, Y)^*\Omega\phi(X, Y) = \phi(\Lambda, 0).$$

Therefore, the eigenvectors of H are orthogonal with respect to the Ω -inner product defined by $\langle u, v \rangle_\Omega := v^*\Omega u$. The eigenvectors can be normalized as

$$\phi(\tilde{X}, \tilde{Y})^*\Omega\phi(\tilde{X}, \tilde{Y}) = I_{2n}.$$

through a diagonal scaling $\phi(\tilde{X}, \tilde{Y}) = \phi(X, Y)\phi(\Lambda^{-1/2}, 0)$.

Second, it follows directly from (6) that

$$\phi(X, Y)^*C_n\phi(X, Y) = C_n.$$

This indicates that the eigenvectors of H are also orthogonal with respect to the C -inner product defined by $\langle u, v \rangle_C := v^*C_n u$, which is an *indefinite scalar*

product [20]. Furthermore, the positive (negative) eigenvalues of H are also the C -positive (C -negative) eigenvalues of the definite pencil $\Omega - \lambda C_n$.¹

These two types of orthogonal properties can be used to construct structure preserving projections that play a key role in Krylov subspace based eigensolvers. Suppose that $\phi(X_k, Y_k) \in \mathbb{C}^{2n \times 2k}$ is orthonormal with respect to the Ω -inner product. Then projection using $\phi(X_k, Y_k)$ yields a $2k \times 2k$ Hermitian matrix of the form

$$H_k := \phi(X_k, Y_k)^* \Omega H \phi(X_k, Y_k) = \phi(X_k, Y_k)^* \Omega C_n \Omega \phi(X_k, Y_k) =: C_n \phi(A_k, \bar{B}_k). \tag{7}$$

It can be easily shown that the eigenvalues of the projected Hermitian matrix H_k also occur in pairs $\pm\theta$, as H_k admits a structured spectral decomposition $H_k = \phi(U_k, V_k) C_k \phi(\Theta_k, 0) \phi(U_k, V_k)^*$, where $\phi(U_k, V_k)^* \phi(U_k, V_k) = I_{2k}$. Furthermore, the matrix $\phi(X_k, Y_k) \phi(U_k, V_k)$ is again orthonormal with respect to the Ω -inner product. Thus we regard (7) as a structure preserving projection. But we remark that Θ_k is not always positive definite here as H_k can sometimes be singular.

Similarly, if $\phi(X_k, Y_k) \in \mathbb{C}^{2n \times 2k}$ is orthonormal with respect to the C -inner product, that is, $\phi(X_k, Y_k)^* C_n \phi(X_k, Y_k) = C_k$. Then

$$H_k := C_k \phi(X_k, Y_k)^* C_n H \phi(X_k, Y_k) = C_k [\phi(X_k, Y_k)^* \Omega \phi(X_k, Y_k)] \tag{8}$$

is a $2k \times 2k$ definite BSE Hamiltonian. Therefore the projection (8) in C -inner product can also be regarded as structure preserving.

3 Variational Properties

3.1 Min–Max Principles

The i th smallest eigenvalues of the Hermitian–definite pencil $C_n - \mu \Omega$, denoted by μ_i , can be characterized by the Courant–Fischer min–max principle

$$\mu_i = \min_{\dim(\mathcal{Y})=i} \max_{\substack{z \in \mathcal{Y} \\ z \neq 0}} \frac{z^* C_n z}{z^* \Omega z} \tag{9}$$

$$= \max_{\dim(\mathcal{Y})=2n-i+1} \min_{\substack{z \in \mathcal{Y} \\ z \neq 0}} \frac{z^* C_n z}{z^* \Omega z}, \tag{10}$$

¹A vector v is called C -positive, C -negative, C -neutral, respectively, if $v^* C_n v > 0$, $v^* C_n v < 0$, $v^* C_n v = 0$. An eigenvalue of $\Omega - \lambda C_n$ is called C -positive (C -negative) if its associated eigenvector is C -positive (C -negative).

where \mathcal{V} is a linear subspace of \mathbb{C}^{2n} . Notice that

$$\lambda_i^+ = \frac{1}{\mu_{2n-i+1}} > 0, \quad (1 \leq i \leq n).$$

Taking the reciprocal of (9) and (10) yields Theorem 2 below. The theorem is also a direct consequence of the Wielandt min–max principle discussed in [10, Theorem 2.2] for the definite pencil $\Omega - \lambda C_n$.

Theorem 2 *Let $H = C_n \Omega$ be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Then*

$$\lambda_i^+ = \max_{\dim(\mathcal{V})=2n-i+1} \min_{\substack{z \in \mathcal{V} \\ z^* C_n z > 0}} \frac{z^* \Omega z}{z^* C_n z} \tag{11}$$

$$= \min_{\dim(\mathcal{V})=i} \max_{\substack{z \in \mathcal{V} \\ z^* C_n z > 0}} \frac{z^* \Omega z}{z^* C_n z} \tag{12}$$

for $1 \leq i \leq n$.

An important special case is $i = 1$, for which we have the following Corollary 1.

Corollary 1 ([19]) *The smallest positive eigenvalue of a definite Bethe–Salpeter Hamiltonian matrix $H = C_n \Omega$ satisfies*

$$\lambda_1^+ = \min_{x^* x - y^* y \neq 0} \varrho(x, y), \tag{13}$$

where

$$\varrho(x, y) = \frac{\begin{bmatrix} x \\ y \end{bmatrix}^* \begin{bmatrix} A & B \\ \overline{B} & \overline{A} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}}{|x^* x - y^* y|}. \tag{14}$$

is the Thouless functional.

Thanks to this result, the computation of λ_1^+ can be converted to minimizing the Thouless functional (14). Thus optimization based eigensolvers, such as the Davidson algorithm [6] and the LOBPCG algorithm [8], can be adopted to compute λ_1^+ .

Finally, we remark that, from a computational point of view, the use of (12) requires additional care, because for an arbitrarily chosen subspace $\mathcal{V} \subset \mathbb{C}^{2n}$ the quantity

$$\sup_{\substack{z \in \mathcal{V} \\ z^* C_n z > 0}} \frac{z^* \Omega z}{z^* C_n z} = \sup_{\substack{z \in \mathcal{V} \\ z^* C_n z = 1}} z^* \Omega z$$

can easily become $+\infty$ when \mathcal{V} contains C -neutral vectors.

3.2 Trace Minimization Principles

In many applications, only a few smallest positive eigenvalues of H are of practical interest. The computation of these interior eigenvalues requires additional care since interior eigenvalues are in general much more difficult to compute compared to external ones. Recently, (13) has been extended to a trace minimization principle for real BSE Hamiltonians [1], so that several eigenvalues can be computed simultaneously using a blocked algorithm [2, 9]. In the following, we present two trace minimization principles, corresponding to the two types of structured preserving projections discussed in Sect. 2.2.

Theorem 3 *Let $H = C_n \Omega$ be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Then*

$$-\left(\frac{1}{\lambda_1^+} + \cdots + \frac{1}{\lambda_k^+}\right) = \min_{\phi(X,Y)^* \Omega \phi(X,Y) = I_{2k}} \text{trace}(X^* X - Y^* Y) \tag{15}$$

holds for $1 \leq k \leq n$.

Proof We rewrite the eigenvalue problem $H z = \lambda z$ as $C_n z = \lambda^{-1} \Omega z$. Then by the trace minimization principle for Hermitian–definite GEP, we obtain

$$-\left(\frac{1}{\lambda_1^+} + \cdots + \frac{1}{\lambda_k^+}\right) = \min_{Z^* \Omega Z = I_k} \text{trace}(Z^* C_n Z).$$

Notice that

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathbb{C}^{2n \times k} : \phi(X, Y)^* \Omega \phi(X, Y) = I_{2k} \right\}$$

is a subset of

$$\mathcal{S}_2 := \{Z \in \mathbb{C}^{2n \times k} : Z^* \Omega Z = I_k\}.$$

We have

$$\min_{Z \in \mathcal{S}_2} \text{trace}(Z^* C_n Z) \leq \min_{Z \in \mathcal{S}_1} \text{trace}(Z^* C_n Z).$$

The equality is attainable, since the minimizer in \mathcal{S}_2 can be chosen as the eigenvectors of H , which is also in \mathcal{S}_1 . As a result, (15) follows directly from the fact that $[X^*, Y^*] C_n [X^*, Y^*]^* = X^* X - Y^* Y$. □

Theorem 4 Let $H = C_n \Omega$ be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Then

$$\lambda_1^+ + \cdots + \lambda_k^+ = \min_{\phi(X,Y)^* C_n \phi(X,Y) = C_k} \text{trace}(X^* A X + X^* B Y + Y^* \bar{B} X + Y^* \bar{A} Y) \quad (16)$$

holds for $1 \leq k \leq n$.

Proof As the eigenvalues of H are also the eigenvalues of the definite pencil $\Omega - \lambda C_n$, by the trace minimization property of definite pencils (see, for example, [9, Theorem 2.4]), we obtain

$$\begin{aligned} \lambda_1^+ + \cdots + \lambda_k^+ &= \min_{Z^* C_n Z = I_k} \text{trace}(Z^* \Omega Z) \\ &= \frac{1}{2} \min_{Z^* C_n Z = C_k} \text{trace}(Z^* \Omega Z). \end{aligned}$$

The rest of the proof is nearly identical to that of Theorem 3. Because

$$\mathcal{S}_1 := \left\{ \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathbb{C}^{2n \times k} : \phi(X, Y)^* C_n \phi(X, Y) = C_k \right\}$$

is a subset of

$$\mathcal{S}_2 := \{Z \in \mathbb{C}^{2n \times k} : Z^* C_n Z = C_k\},$$

we have

$$\min_{Z \in \mathcal{S}_2} \text{trace}(Z^* \Omega Z) \leq \min_{Z \in \mathcal{S}_1} \text{trace}(Z^* \Omega Z).$$

The equality is attainable by choosing the corresponding eigenvectors of H , which belong to both \mathcal{S}_1 and \mathcal{S}_2 . \square

Theorems 3 and 4 can both be used to derive structure preserving optimization based eigensolvers. We shall discuss the computation of eigenvalues in separate publications. We also refer the readers to [10] for more general variational principles.

3.3 Interlacing Properties

We have already seen that the two types of orthogonality conditions on the eigenvectors of H can both be used to construct structure preserving projections that can be used for eigenvalue computations. In this subsection we point out some difference on the location of the Ritz values.

When the Ω -inner product is used for projection, we have the following Cauchy type interlacing property.

Theorem 5 *Let $H = C_n\Omega$ be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Suppose that $\phi(X, Y)^*\Omega\phi(X, Y) = I_{2k}$, where $1 \leq k \leq n$. Then the eigenvalues of $\phi(X, Y)^*\Omega H\phi(X, Y)$ are real and appear in pairs $\pm\theta$. Moreover²*

$$\lambda_i^+(\phi(X, Y)^*\Omega H\phi(X, Y)) \leq \lambda_{n+i-k}^+(H), \quad (1 \leq i \leq k). \quad (17)$$

Proof The first half of the theorem follows from the discussions in Sect. 2.2. We only show the interlacing property. Notice that $U := \Omega^{1/2}\phi(X, Y)$ has orthonormal columns in the standard inner product, that is, $U^*U = I_{2k}$. By the Cauchy interlacing theorem, we have

$$\begin{aligned} \lambda_i^+(\phi(X, Y)^*\Omega H\phi(X, Y)) &= \lambda_i^+(U^*\Omega^{1/2}C_n\Omega^{1/2}U) \\ &\leq \lambda_{n+i-k}^+(\Omega^{1/2}C_n\Omega^{1/2}) \\ &= \lambda_{n+i-k}^+(H). \quad \square \end{aligned}$$

In contrast to the standard Cauchy interlacing theorem, there is no nontrivial lower bound on the Ritz value $\lambda_i^+(\phi(X, Y)^*\Omega H\phi(X, Y))$ here. In fact, the projected matrix $\phi(X, Y)^*\Omega H\phi(X, Y)$ can even be zero. For instance,

$$A = I_n, \quad B = 0, \quad X = \frac{1}{\sqrt{2}} \begin{bmatrix} I_k \\ I_k \\ 0 \end{bmatrix}, \quad Y = \frac{1}{\sqrt{2}} \begin{bmatrix} I_k \\ -I_k \\ 0 \end{bmatrix}$$

is an example for such an extreme case (assuming $2k \leq n$).

For projections based on the C -inner product, we establish Theorem 6 below. Similar to Theorem 5, Ritz values are only bounded in one direction. However, in this case, it is possible to provide a meaningful (though complicated) upper bound for the Ritz value. We refer the readers to [1, Theorem 4.1] for the case of real BSE. Further investigation in this direction is beyond the scope of this paper.

Theorem 6 *Let $H = C_n\Omega$ be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Suppose that $\phi(X, Y)^*C_n\phi(X, Y) = C_k$, where $1 \leq k \leq n$. Then the eigenvalues of $C_k\phi(X, Y)^*C_nH\phi(X, Y)$ appear in pairs $\pm\theta$. Moreover*

$$\lambda_i^+(C_k\phi(X, Y)^*C_nH\phi(X, Y)) \geq \lambda_i^+(H), \quad (1 \leq i \leq k). \quad (18)$$

Proof Notice that the eigenvalues of $C_k(\phi(X, Y)^*\Omega\phi(X, Y))$ can also be regarded as the eigenvalues of the definite pencil $\phi(X, Y)^*(\Omega - \lambda C_n)\phi(X, Y)$. Then the

²In the case when $\phi(X, Y)^*\Omega H\phi(X, Y)$ is singular, we assign half of the zero eigenvalues with the positive sign in the notation $\lambda_i^+(\phi(X, Y)^*\Omega H\phi(X, Y))$.

conclusion follows from the Cauchy interlacing property of definite pencils [9, Theorem 2.3]. \square

From a computational perspective, (18) provides more useful information than (17), because the Ritz value $\lambda_i^+(C_k\phi(X, Y)^*C_nH\phi(X, Y))$ is bounded in terms of the corresponding eigenvalue $\lambda_i^+(H)$ to be approximated. The inequality (17) gives an upper bound of the Ritz value. But we have less control over the location of $\lambda_i^+(\phi(X, Y)^*\Omega H\phi(X, Y))$.

Finally, we remark that the trace minimization principle (16) can also be derived by the interlacing property (18).

4 Eigenvalue Perturbation Bounds

4.1 Weyl Type Inequalities

In the perturbation theory of symmetric eigenvalue problems, Weyl's inequality implies that the eigenvalues of a Hermitian matrix are well conditioned when a Hermitian perturbation is introduced. In the following we establish similar results for definite Bethe–Salpeter eigenvalue problems.

Theorem 7 *Let H and $H + \Delta H$ be definite Bethe–Salpeter Hamiltonian matrices. Then*

$$\left| \frac{\lambda_i^+(H + \Delta H) - \lambda_i^+(H)}{\lambda_i^+(H)} \right| \leq \kappa_2(H) \frac{\|\Delta H\|_2}{\|H\|_2}, \quad (1 \leq i \leq n),$$

where $\kappa_2(H) = \|H\|_2\|H^{-1}\|_2$.

Proof Let $\Delta\Omega = C_n\Delta H$. Then $\Omega + \Delta\Omega$ is positive definite. We rewrite $Hz = \lambda z$ as the GEP $C_n z = \lambda^{-1}\Omega z$. It follows from the Weyl inequality on Hermitian–definite GEP [12, Theorem 2.1] that

$$\left| \frac{1}{\lambda_i^+(H)} - \frac{1}{\lambda_i^+(H + \Delta H)} \right| \leq \frac{\|\Omega^{-1}\|_2\|\Delta\Omega\|_2}{\lambda_i^+(H + \Delta H)}.$$

By simple arithmetic manipulations, we arrive at

$$\left| \frac{\lambda_i^+(H + \Delta H) - \lambda_i^+(H)}{\lambda_i^+(H)} \right| \leq \kappa_2(\Omega) \frac{\|\Delta\Omega\|_2}{\|\Omega\|_2} = \kappa_2(H) \frac{\|\Delta H\|_2}{\|H\|_2}. \quad \square$$

Theorem 7 characterizes the sensitivity of the eigenvalues of H when a structured perturbation is introduced—the relative condition number of $\lambda_i^+(H)$ is bounded by $\kappa_2(H)$. When the perturbation is also a definite BSE Hamiltonian, the eigenvalues are perturbed monotonically. We have the following result.

Theorem 8 *Let $H, \Delta H \in \mathbb{C}^{2n \times 2n}$ be definite Bethe–Salpeter Hamiltonian matrices. Then*

$$\lambda_i^+(H + \Delta H) \geq \lambda_i^+(H) + \lambda_1^+(\Delta H), \quad (1 \leq i \leq n).$$

Proof Let $\Delta\Omega = C_n\Delta H$. Then by Theorem 2 we have

$$\begin{aligned} \lambda_i^+(H + \Delta H) &= \max_{\dim(\mathcal{Y})=2n-i+1} \min_{\substack{z \in \mathcal{Y} \\ z^* C_n z > 0}} \left(\frac{z^* \Omega z}{z^* C_n z} + \frac{z^* \Delta\Omega z}{z^* C_n z} \right) \\ &\geq \max_{\dim(\mathcal{Y})=2n-i+1} \min_{\substack{z \in \mathcal{Y} \\ z^* C_n z > 0}} \left(\frac{z^* \Omega z}{z^* C_n z} + \lambda_1^+(\Delta H) \right) \\ &= \lambda_i^+(H) + \lambda_1^+(\Delta H). \quad \square \end{aligned}$$

A special perturbation in the context of Bethe–Salpeter eigenvalue problems is to drop the off-diagonal blocks in H . Such a perturbation is known as the *Tamm–Dancoff approximation* (TDA) [5, 18]. Similar to the monotonic perturbation behavior above, it has been shown in [16] that TDA overestimates all positive eigenvalues of H . In the following, we present a simpler proof of this property than the one given in [16].

Theorem 9 ([16, Theorem 4]) *Let H be a definite Bethe–Salpeter Hamiltonian matrix as defined in (1). Then*

$$\lambda_i^+(H) \leq \lambda_i^+(A), \quad (1 \leq i \leq n).$$

Proof Notice that $H^2 = (C_n\Omega C_n)\Omega$ with both $C_n\Omega C_n$ and Ω positive definite. By the arithmetic–geometric inequality on positive definite matrices [4, Sect. 3.4], we obtain

$$\lambda_i^+(H) = \sqrt{\lambda_{2i}^+((C_n\Omega C_n)\Omega)} \leq \lambda_{2i}^+\left(\frac{C_n\Omega C_n + \Omega}{2}\right) = \lambda_{2i}^+\left(\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}\right) = \lambda_i^+(A). \quad \square$$

Combining Theorems 7 and 9, we obtain the following corollary. It characterizes to what extent existing results in the literature obtained from TDA are reliable.

Corollary 2 *If H is a Bethe–Salpeter Hamiltonian matrix as defined in (1), then*

$$0 \leq \frac{\lambda_i^+(A) - \lambda_i^+(H)}{\lambda_i^+(H)} \leq \kappa_2(H) \frac{\|B\|_2}{\|H\|_2}, \quad (1 \leq i \leq n).$$

4.2 Residual Bounds

Another type of perturbation bounds on eigenvalues measures the accuracy of approximate eigenvalues in terms of the residual norm. These bounds are of interest in eigenvalue computations. In the following we discuss several residual bounds for the definite Bethe–Salpeter eigenvalue problem.

Theorem 10 *Let $H = C_n \Omega$ be a definite Bethe–Salpeter Hamiltonian matrix. Suppose that $X, Y \in \mathbb{C}^{n \times k}$ satisfy*

$$\phi(X, Y)^* \Omega \phi(X, Y) = I_{2k}, \quad \phi(X, Y)^* C_n \phi(X, Y) = \begin{bmatrix} \Theta & 0 \\ 0 & -\Theta \end{bmatrix}^{-1},$$

for some k between 1 and n , where $\Theta = \text{diag}\{\theta_1, \dots, \theta_k\} > 0$. Then there exists a BSE Hamiltonian $\Delta H = C_n \Delta \Omega = C_n \phi(\Delta A, \overline{\Delta B})$ such that

$$(H + \Delta H)\phi(X, Y) = \phi(X, Y) \begin{bmatrix} \Theta & 0 \\ 0 & -\Theta \end{bmatrix}. \quad (19)$$

and

$$\|\Delta H\|_2 \leq 2\|H\|_2^{1/2}\|R\|_2, \quad (20)$$

where

$$R = H\phi(X, Y) - \phi(X, Y) \begin{bmatrix} \Theta & 0 \\ 0 & -\Theta \end{bmatrix}.$$

Proof It follows from the definition of R that

$$\phi(X, Y)^* C_n R = I_{2k} - \phi(X, Y)^* C_n \phi(X, Y) \begin{bmatrix} \Theta & 0 \\ 0 & -\Theta \end{bmatrix} = 0.$$

Let

$$\Delta \Omega = C_n R \phi(X, Y)^* \Omega + \Omega \phi(X, Y) R^* C_n.$$

Then $\Delta\Omega$ is Hermitian. Since Θ is real, we have

$$R = C_n\phi(A, \bar{B})\phi(X, Y) - \phi(X, Y)C_n\phi(\Theta, 0) = C_n\phi(AX + BY - X\Theta, A\bar{Y} + B\bar{X} + Y\Theta),$$

indicating that

$$\Omega\phi(X, Y)R^*C_n = \phi(A, \bar{B})\phi(X, Y)\phi(AX + BY - X\Theta, A\bar{Y} + B\bar{X} + Y\Theta)^*$$

has the block structure $\phi(\cdot, \cdot)$. Thus $\Delta H := C_n\Delta\Omega$ is a BSE Hamiltonian. It can be easily verified that (19) is satisfied. Finally,

$$\|\Delta H\|_2 = \|\Delta\Omega\|_2 \leq 2\|\Omega^{1/2}\|_2\|\Omega^{1/2}\phi(X, Y)\|_2\|R^*\|_2 = 2\|H\|_2^{1/2}\|R\|_2. \quad \square$$

Roughly speaking, (20) implies that for definite BSE, Rayleigh–Ritz based algorithms that produce small residual norms are backward stable. When $\kappa_2(H)$ is of modest size, backward stability implies forward stability according to Theorem 7. The following theorem provides a slightly better estimate compared to simply combining Theorems 7 and 10.

Theorem 11 *Under the same assumption of Theorem 10, there exist k positive eigenvalues of H , $\lambda_{j_1} \leq \dots \leq \lambda_{j_k}$, such that*

$$|\theta_i - \lambda_{j_i}| \leq \|H\|_2^{1/2}\|R\|_2, \quad (1 \leq i \leq k).$$

Proof Notice that $U := \Omega^{1/2}\phi(X, Y)$ has orthonormal columns (in the standard inner product), and

$$\Omega^{1/2}R = \Omega^{1/2}C_n\Omega^{1/2}U - U \begin{bmatrix} \Theta & 0 \\ 0 & -\Theta \end{bmatrix}.$$

By the residual bound for standard Hermitian eigenvalue problems (see [14, Theorem 11.5.1] or [17, Sect. IV.4.4]), we obtain that there are $2k$ eigenvalues of $\Omega^{1/2}C_n\Omega^{1/2}$, $\tilde{\lambda}_{-j_k} \leq \dots \leq \tilde{\lambda}_{-j_1} \leq \tilde{\lambda}_{j_1} \leq \dots \leq \tilde{\lambda}_{j_k}$, such that

$$\max \left\{ |\theta_i + \tilde{\lambda}_{-j_i}|, |\theta_i - \tilde{\lambda}_{j_i}| \right\} \leq \|\Omega\|_2^{1/2}\|R\|_2 = \|H\|_2^{1/2}\|R\|_2, \quad (1 \leq i \leq k).$$

Note that at least one of the inequalities $\tilde{\lambda}_{j_i} > 0$ and $\tilde{\lambda}_{-j_i} < 0$ holds. As the eigenvalues of $\Omega^{1/2}C_n\Omega^{1/2}$ are identical to those of H , the conclusion follows immediately by choosing

$$\lambda_{j_i} = \begin{cases} \tilde{\lambda}_{j_i}, & \text{if } \tilde{\lambda}_{j_i} > 0, \\ -\tilde{\lambda}_{-j_i}, & \text{otherwise,} \end{cases} \quad (1 \leq i \leq k). \quad \square$$

Finally, we end this section by a Temple–Kato type quadratic residual bound as stated in Theorem 12. The quadratic residual bound explains the fact that the accuracy of a computed eigenvalues is in general much higher compared to that of the corresponding eigenpair. Such a behavior has been reported for real Bethe–Salpeter eigenvalue problem in [9]. It is certainly possible to extend Theorem 12 to a subspace manner using techniques in [11, 13].

Theorem 12 *Let (θ, \hat{z}) be an approximate eigenpair of a definite BSE Hamiltonian $H = C_n \Omega$ satisfying*

$$\frac{\hat{z}^* \Omega \hat{z}}{\hat{z}^* C_n \hat{z}} = \theta.$$

Then the eigenvalue of H closest to θ , denoted by λ , satisfies

$$|\theta^{-1} - \lambda^{-1}| \leq \frac{\|H^{-1}\|_2 \|H\hat{z} - \theta\hat{z}\|_2^2}{\text{gap}(\theta) \hat{z}^* \Omega \hat{z}},$$

where

$$\text{gap}(\theta) := \min_{\lambda_i(H) \neq \theta} |\theta^{-1} - \lambda_i(H)^{-1}|.$$

Proof The theorem is a direct consequence of [14, Theorem 11.7.1] on the equivalent Hermitian eigenvalue problem $(\Omega^{-1/2} C_n \Omega^{-1/2})(\Omega^{1/2} z) = \lambda(\Omega^{1/2} z)$. \square

5 Summary

The Bethe–Salpeter eigenvalue problem is an important class of structured eigenvalue problems arising from several physics and chemistry applications. The most important case, the definite Bethe–Salpeter eigenvalue problem, has a number of interesting properties. We identified two types of orthogonality conditions on the eigenvectors, and discussed several properties of the corresponding structure preserving projections. Although most of our theoretical results can be derived by extending similar results for general symmetric eigenvalue problems to this class of problems, they play an important role in developing and analyzing structure preserving algorithms for solving this type of problems. Numerical algorithms will be discussed in a separate publication.

Acknowledgements The authors thank Ren-Cang Li for helpful discussions. Support for this work was provided through Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research.

References

1. Bai, Z., Li, R.C.: Minimization principles for the linear response eigenvalue problem I: theory. *SIAM J. Matrix Anal. Appl.* **33**(4), 1075–1100 (2012). doi:[10.1137/110838960](https://doi.org/10.1137/110838960)
2. Bai, Z., Li, R.C.: Minimization principles for the linear response eigenvalue problem II: computation. *SIAM J. Matrix Anal. Appl.* **34**(2), 392–416 (2013). doi:[10.1137/110838972](https://doi.org/10.1137/110838972)
3. Benner, P., Fassbender, H., Yang, C.: Some remarks on the complex J -symmetric eigenproblem. Preprint MPIMD/15–12, Max Planck Institute Magdeburg. Available from <http://www.mpi-magdeburg.mpg.de/preprints/> (2015)
4. Bhatia, R., Kittaneh, F.: Notes on matrix arithmetic–geometric mean inequalities. *Linear Algebra Appl.* **308**, 203–211 (2000). doi:[10.1016/S0024-3795\(00\)00048-3](https://doi.org/10.1016/S0024-3795(00)00048-3)
5. Dancoff, S.M.: Non-adiabatic meson theory of nuclear forces. *Phys. Rev.* **78**(4), 382–385 (1950). doi:[10.1103/PhysRev.78.382](https://doi.org/10.1103/PhysRev.78.382)
6. Davidson, E.R.: The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. *J. Comput. Phys.* **17**(1), 87–94 (1975). doi:[10.1016/0021-9991\(75\)90065-0](https://doi.org/10.1016/0021-9991(75)90065-0)
7. Gohberg, I., Lancaster, P., Rodman, L.: *Matrices and Indefinite Scalar Products*. Operator Theory: Advances and Applications, vol. 8. Birkhäuser, Basel (1983)
8. Knyazev, A.V.: Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method. *SIAM J. Sci. Comput.* **23**(2), 517–541 (2001). doi:[10.1137/S1064827500366124](https://doi.org/10.1137/S1064827500366124)
9. Kressner, D., Miloloža Pandur, M., Shao, M.: An indefinite variant of LOBPCG for definite matrix pencils. *Numer. Algorithms* **66**(4), 681–703 (2014). doi:[10.1007/s11075-013-9754-3](https://doi.org/10.1007/s11075-013-9754-3)
10. Liang, X., Li, R.C.: Extensions for Wielandt’s min–max principles for positive semi-definite pencils. *Linear Multilinear Algebra* **62**(8), 1032–1048 (2014). doi:[10.1080/03081087.2013.803242](https://doi.org/10.1080/03081087.2013.803242)
11. Mathias, R.: Quadratic residual bounds for the Hermitian eigenvalue problem. *SIAM J. Matrix Anal. Appl.* **19**(2), 541–550 (1998). doi:[10.1137/S0895479896310536](https://doi.org/10.1137/S0895479896310536)
12. Nakatsukasa, Y.: Perturbation behavior of a multiple eigenvalue in generalized Hermitian eigenvalue problems. *BIT Numer. Math.* **50**, 109–121 (2010). doi:[10.1007/s10543-010-0254-8](https://doi.org/10.1007/s10543-010-0254-8)
13. Nakatsukasa, Y.: Eigenvalue perturbation bounds for Hermitian block tridiagonal matrices. *Appl. Numer. Math.* **62**, 67–78 (2012). doi:[10.1016/j.apnum.2011.09.010](https://doi.org/10.1016/j.apnum.2011.09.010)
14. Parlett, B.N.: *The Symmetric Eigenvalue Problem*. Classics in Applied Mathematics, vol. 20. SIAM, Philadelphia, PA (1998). Corrected reprint of the 1980 original
15. Salpeter, E.E., Bethe, H.A.: A relativistic equation for bounded-state problems. *Phys. Rev.* **84**(6), 1232–1242 (1951). doi:[10.1103/PhysRev.84.1232](https://doi.org/10.1103/PhysRev.84.1232)
16. Shao, M., H. da Jornada, F., Yang, C., Deslippe, J., Louie, S.G.: Structure preserving parallel algorithms for solving the Bethe–Salpeter eigenvalue problem. *Linear Algebra Appl.* **488**, 148–167 (2016). doi:[10.1016/j.laa.2015.09.036](https://doi.org/10.1016/j.laa.2015.09.036)
17. Stewart, G.W., Sun, J.: *Matrix Perturbation Theory*. Academic, Boston, MA (1990)
18. Tamm, I.Y.: Relativistic interaction of elementary particles. *J. Phys. (USSR)* **9**, 449–460 (1945)
19. Thouless, D.J.: Vibrational states of nuclei in the random phase approximation. *Nucl. Phys.* **22**, 78–95 (1961). doi:[10.1016/0029-5582\(61\)90364-9](https://doi.org/10.1016/0029-5582(61)90364-9)
20. Veselić, K.: *Damped Oscillations of Linear Systems—A Mathematical Introduction*. Lecture Notes in Mathematics, vol. 2023. Springer, Heidelberg (2011)