Convergence of Eigenvectors in the Block Classical Jacobi Method for the Symmetric Eigenvalue Problem

Yusaku Yamamoto Gabriel Okša Marián Vajteršic

Technical Report 2019-04 November 2019

Department of Computer Sciences

Jakob-Haringer-Straße 2
5020 Salzburg
Austria
www.cosy.sbg.ac.at

Technical Report Series

Convergence of Eigenvectors in the Block Classical Jacobi Method for the Symmetric Eigenvalue Problem

Yusaku Yamamoto* Gabriel Okša† Marián Vajteršic‡

November 16, 2019

Abstract. *In the block version of the classical Jacobi method for the symmetric eigenvalue problem, the off-diagonal elements of the iterate matrix $A^{(k)}$ converge to zero. However, this does not necessarily guarantee that $A^{(k)}$ converges to a fixed matrix. In this report, we prove that $A^{(k)}$ actually converges to a fixed diagonal matrix, whose diagonal elements are the eigenvalues of the input matrix $A^{(0)}$. In addition, it is shown that for a simple eigenvalue, the corresponding column of the accumulated orthogonal transformation matrix $Q^{(k)}$ converges to the corresponding eigenvector. For multiple eigenvalues, the orthogonal projector constructed from the corresponding columns of $Q^{(k)}$ converges to an orthogonal projector onto the eigenspace corresponding to those eigenvalues.*

1 Introduction

Consider the block version of the classical Jacobi method for the symmetric eigenvalue problem, in which the off-diagonal block with the largest Frobenius norm is annihilated at each step. In our previous papers [7, 8, 9], we were mainly concerned with the convergence of the off-diagonal norm of the iteration matrix $A^{(k)}$ to zero. However, this does not necessarily guarantee that $A^{(k)}$ converges to a *fixed* diagonal matrix. In this report, we show that $A^{(k)}$ actually converges to a fixed diagonal matrix D and its diagonal elements are the eigenvalues of the initial matrix. This is proved regardless of the multiplicity of the eigenvalues. Furthermore, in the case of simple eigenvalues, we also show that the accumulated orthogonal transformation factor converges to the eigenvector matrix. It should be easy to extend these results to the parallel block Jacobi eigenvalue algorithm with greedy implementation of parallel dynamic ordering (GIPDO) [1, 2].

*Department of Communication Engineering and Informatics, University of Electro-Communications, Tokyo, Japan, email: yusaku.yamamoto@uec.ac.jp

†Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic, email: gabriel.oksa@savba.sk

‡Department of Computer Sciences, University of Salzburg, Austria, and Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic, email: marian@cosy.sbg.ac.at

In the following, we denote the identity matrix of order n by I_n and the $m \times n$ zero matrix by O_{mn}. For a square matrix A, $\mathrm{off}(A)$ denotes its off-diagonal part. For a real symmetric matrix A, we denote its smallest and largest eigenvalues by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$, respectively.

2 Convergence analysis

2.1 Preliminaries

Here we recall several theorems and lemmas to be used in the following subsections. In the following, we consider applying the classical block Jacobi method to a symmetric matrix $A = A^{(0)} \in \mathbb{R}^{n \times n}$ partitioned into a $w \times w$ block structure ($w \geq 2$). For simplicity, we assume that n is divisible by w and consider only equally sized blocks of $\ell \times \ell$, where $\ell = n/w$, but extension to a more general case is obvious. The matrix after the kth step is denoted by $A^{(k)}$. We also assume that the diagonal blocks of A are diagonalized before the first step and the diagonal elements in each diagonal block are ordered non-increasingly. Hence, the diagonal blocks of $A^{(k)}$ remain diagonal throughout the computation.

Each step of the classical block Jacobi method proceeds as follows. We first choose the off-diagonal block of $A^{(k)}$ with the largest Frobenius norm from the upper triangular part. Let the chosen block be $A^{(k)}_{XY}$ ($X < Y$). We then form a 2×2 block matrix

$$\tilde{A}^{(k)} \equiv \begin{pmatrix} A^{(k)}_{XX} & A^{(k)}_{XY} \\ A^{(k)}_{YX} & A^{(k)}_{YY} \end{pmatrix} \in \mathbb{R}^{2\ell \times 2\ell} \tag{1}$$

and compute its orthogonal eigenvector matrix

$$\tilde{P}^{(k)} \equiv \begin{pmatrix} P^{(k)}_{XX} & P^{(k)}_{XY} \\ P^{(k)}_{YX} & P^{(k)}_{YY} \end{pmatrix} \in \mathbb{R}^{2\ell \times 2\ell}. \tag{2}$$

In this paper, we choose the sign of each eigenvector in $\tilde{P}^{(k)}$ so that the diagonal elements of $\tilde{P}^{(k)}$ is nonnegative. Note that $A^{(k)}_{XX}$ and $A^{(k)}_{YY}$ are diagonal as we stated above. Now we embed $\tilde{P}^{(k)}$ into I_n to obtain an $n \times n$ orthogonal matrix $P^{(k)}$ and compute $A^{(k+1)}$ as

$$A^{(k+1)} = (P^{(k)})^\top A^{(k)} P^{(k)}. \tag{3}$$

It is easy to see that $\tilde{A}^{(k)}$ is updated as

$$\tilde{A}^{(k+1)} = (\tilde{P}^{(k)})^\top \tilde{A}^{(k)} \tilde{P}^{(k)} = \begin{pmatrix} A^{(k+1)}_{XX} & O \\ O & A^{(k+1)}_{YY} \end{pmatrix}, \tag{4}$$

where $A^{(k+1)}_{XX}$ and $A^{(k+1)}_{YY}$ are diagonal. At the same time, $P^{(k)}$ is accumulated as

$$Q^{(k+1)} = Q^{(k)} P^{(k)}, \tag{5}$$

with $Q^{(0)} = I_n$.

We first quote a theorem on the global convergence of the classical Jacobi method [12].

Theorem 2.1. *In the classical block Jacobi method, $\|\mathrm{off}(A^{(k)})\|_F$ satisfies the following inequality and therefore converges to zero as $k \to \infty$.*

$$\|\mathrm{off}(A^{(k+1)})\|_F^2 \leq \alpha \|\mathrm{off}(A^{(k)})\|_F^2, \tag{6}$$

where

$$\alpha \equiv 1 - \frac{2}{w(w-1)} < 1. \tag{7}$$

Specifically, letting $S = \|\mathrm{off}(A)\|_F^2$, we have

$$\|\mathrm{off}(A^{(k)})\|_F^2 \leq S\alpha^k. \tag{8}$$

The next theorem and its corollary will be used to prove the convergence of the diagonal elements of $A^{(k)}$.

Theorem 2.2 (Hoffman-Wielandt theorem [6])**.** *Let $B, C \in \mathbb{R}^{n\times n}$ be symmetric matrices and $\{\lambda_i(B)\}_{i=1}^n$ and $\{\lambda_i(C)\}_{i=1}^n$ be their respective eigenvalues. Also, let $\mathcal{S}_n$ denote the permutation group of $\{1, 2, \ldots, n\}$. Then,*

$$\min_{\sigma \in \mathcal{S}_n} \sum_{i=1}^n |\lambda_i(B) - \lambda_{\sigma(i)}(C)|^2 \leq \|B - C\|_F^2. \tag{9}$$

Corollary 2.3. *Let $\{\lambda_i(B)\}_{i=1}^n$ and $\{\lambda_i(C)\}_{i=1}^n$ be ordered non-decreasingly (or non-increasingly). Then,*

$$\sum_{i=1}^n |\lambda_i(B) - \lambda_i(C)|^2 \leq \|B - C\|_F^2. \tag{10}$$

Proof. We show that $\sigma(i) = i$ attains the minimum in the left hand side of (9) by induction. Let σ^* be one of the permutations that attains the minimum of the left hand of (9). Assume that $\sigma^*(1) \neq 1$. Then there are indices $2 \leq i, j \leq n$ such that $\sigma^*(1) = i$ and $\sigma^*(j) = 1$. Now we modify σ^* to $\hat{\sigma}^*$ so that $\hat{\sigma}^*(1) = 1$, $\hat{\sigma}^*(j) = i$ and $\hat{\sigma}^*(k) = \sigma^*(k)$ for $k \neq 1, j$. Then, we have

$$\begin{aligned}
& \sum_{i=1}^n |\lambda_i(B) - \lambda_{\sigma^*(i)}(C)|^2 - \sum_{i=1}^n |\lambda_i(B) - \lambda_{\hat{\sigma}^*(i)}(C)|^2 \\
= & \left\{|\lambda_1(B) - \lambda_{\sigma^*(1)}(C)|^2 + |\lambda_j(B) - \lambda_{\sigma^*(j)}(C)|^2\right\} \\
& \quad - \left\{|\lambda_1(B) - \lambda_{\hat{\sigma}^*(1)}(C)|^2 + |\lambda_j(B) - \lambda_{\hat{\sigma}^*(j)}(C)|^2\right\} \\
= & \left\{|\lambda_1(B) - \lambda_i(C)|^2 + |\lambda_j(B) - \lambda_1(C)|^2\right\} - \left\{|\lambda_1(B) - \lambda_1(C)|^2 + |\lambda_j(B) - \lambda_i(C)|^2\right\} \\
= & -2\lambda_1(B)\lambda_i(C) - 2\lambda_j(B)\lambda_1(C) + 2\lambda_1(B)\lambda_1(C) + 2\lambda_j(B)\lambda_i(C) \\
= & \; 2\left\{\lambda_1(B) - \lambda_j(B)\right\}\left\{\lambda_1(C) - \lambda_i(C)\right\} \geq 0,
\end{aligned} \tag{11}$$

which shows that $\hat{\sigma}^*$ also attains the minimum. Hence, by regarding $\hat{\sigma}^*$ as σ^*, we can assume without loss of generality that $\sigma^*(1) = 1$. By repeating this process, we can modify σ^* so that $\sigma^*(i) = i$ $(i = 1, 2, \ldots, n)$ without increasing the value of $\sum_{i=1}^n |\lambda_i(B) - \lambda_{\sigma^*(i)}(C)|^2$. □

The next theorem will be used in subsection 2.3 to prove the convergence of the columns of $Q^{(k)}$ corresponding to simple eigenvalues.

Theorem 2.4 ($\sin\Theta$ theorem [10]). *Let $B \in \mathbb{R}^{m\times m}$ be a symmetric matrix and $\mathbf{y} \in \mathbb{R}^m$ be a vector with $\|\mathbf{y}\| = 1$. Define $\rho = \mathbf{y}^\top B\mathbf{y}$ and $\mathbf{r}(\mathbf{y}) = B\mathbf{y} - \rho\mathbf{y}$. Now, let μ_i be the eigenvalue of B that is closest to ρ, $\mathbf{x}_i$ be the corresponding eigenvector, $\theta = \angle(\mathbf{y}, \mathbf{x}_i)$ and* $\mathrm{gap}(\rho) = \min_{\mu_j\neq\mu_i} |\mu_j - \rho|$. *Then the following inequality holds.*

$$|\sin\theta| \leq \frac{\|\mathbf{r}(\mathbf{y})\|}{\mathrm{gap}(\rho)}. \tag{12}$$

Now we move to results that will be used in subsection 2.4. Let us consider a symmetric 2×2 block matrix

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{13}$$

and denote its eigendecomposition as

$$B = PDP^\top, \tag{14}$$

where

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} D_{11} & O \\ O & D_{22} \end{pmatrix} \tag{15}$$

are orthogonal and diagonal matrices, respectively. The next lemma is used to bound the norm of the off-diagonal blocks of the eigenvector matrix P.

Lemma 2.5. *Assume that $\eta \equiv \lambda_{\min}(B_{11}) - \lambda_{\max}(B_{22}) > 0$ and $\sqrt{2}\|B_{21}\|_F < \eta$. Assume further that the diagonal elements of D are arranged in non-increasing order. Then, $\|P_{12}\|_F = \|P_{21}\|_F \leq \|B_{21}\|_F/(\eta - \sqrt{2}\|B_{21}\|_F)$.*

Proof. Since P is also an eigenvector matrix of the shifted matrix $B' = B + sI$ for any s, we shift B so that B_{22} is positive definite and denote the shifted diagonal blocks by a prime. Then, $\lambda_{\min}(B'_{11}) - \lambda_{\max}(B'_{22}) = \eta$. Now we rewrite the shifted version of Eq. (14) as follows.

$$\begin{pmatrix} B'_{11} & B_{12} \\ B_{21} & B'_{22} \end{pmatrix}\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}\begin{pmatrix} D'_{11} & O \\ O & D'_{22} \end{pmatrix}. \tag{16}$$

By focusing on the $(2,1)$ block, we have the following Sylvester equation for P_{21}:

$$P_{21}D'_{11} - B'_{22}P_{21} = B_{21}P_{11}. \tag{17}$$

We bound $\|P_{21}\|_F$ using the technique in the proof of [3, Theorem 5.1]. Letting $\beta = \lambda_{\max}(B'_{22})$, we have

$$\|B'_{22}P_{21}\|_F \leq \|B'_{22}\|_2\|P_{21}\|_F = \beta\|P_{21}\|_F. \tag{18}$$

Now we bound $\|P_{21}D'_{11}\|_F$ from below. Since all eigenvalues of B'_{11} are larger than those of B'_{22} and the diagonal elements of D' are non-increasingly ordered, we can apply Corollary 2.3 to $\mathrm{diag}(B'_{11}, B'_{22})$ and D' to show that all eigenvalues of D'_{11} are within distance of $\sqrt{\|B_{12}\|_F^2 + \|B_{21}\|_F^2} = \sqrt{2}\|B_{21}\|_F$ from some eigenvalue of B'_{11}. Hence, $\lambda_{\min}(D'_{11}) \geq \beta + \eta - \sqrt{2}\|B_{21}\|_F$ and $\|{D'_{11}}^{-1}\|_2 \leq (\beta + \eta - \sqrt{2}\|B_{21}\|_F)^{-1}$. Thus, it follows that

$$\|P_{21}\|_F \leq \|P_{21}D'_{11}\|_F\|{D'_{11}}^{-1}\|_2 \leq (\beta + \eta - \sqrt{2}\|B_{21}\|_F)^{-1}\|P_{21}D'_{11}\|_F \tag{19}$$

and

$$\|P_{21}D'_{11}\|_F \geq (\beta + \eta - \sqrt{2}\|B_{21}\|_F)\|P_{21}\|_F. \tag{20}$$

From Eqs. (17), (18), (20) and $\|P_{11}\|_2 \leq 1$, we have

$$\|B_{21}\|_F \geq \|B_{21}P_{11}\|_F \geq \|P_{21}D'_{11}\|_F - \|B'_{22}P_{21}\|_F \geq (\eta - \sqrt{2}\|B_{21}\|_F)\|P_{21}\|_F. \tag{21}$$

□

This result is a variant of the famous theorem of Davis and Kahan on the Sylvester equation $AX - XB = C$ [3, Theorem 5.1], in which the 2-norm is used for D'_{11} and B'_{22} and the Frobenius norm is used for B_{21} and P_{21}.

The final Lemma in this subsection concerns the completeness of the set of orthogonal projectors.

Lemma 2.6. *Let $\mathcal{P}_{n,m} \subset \mathbb{R}^{n\times n}$ $(m \leq n)$ be the set of all orthogonal projectors from $\mathbb{R}^n$ to its m-dimensional subspace. Then $\mathcal{P}_{n,m}$ is complete with respect to the metric in $\mathbb{R}^{n\times n}$ induced by the 2-norm or the Frobenius norm.*

Proof. We first prove the lemma for the Frobenius norm. In this case, $\mathbb{R}^{n\times n}$ can be identified with the Euclidean space $\mathbb{R}^{n^2}$. Then, since any closed set of a Euclidean space is complete, all we need to show is that $\mathcal{P}_{n,m}$ is a closed set, that is, if $P^{(1)}, P^{(2)}, \ldots \in \mathcal{P}_{n,m}$ and $\lim_{k\to\infty} P^{(k)} = P^{(\infty)}$ (that is, $\lim_{k\to\infty} \|P^{(k)} - P^{(\infty)}\|_F = 0$), then $P^{(\infty)} \in \mathcal{P}_{n,m}$. First, we note that any $P \in \mathcal{P}_{n,m}$ can be written as $P = YY^\top$, where $Y \in \mathbb{R}^{n\times m}$ and the columns of Y are an orthonormal basis of the m-dimensional subspace onto which P projects a vector in $\mathbb{R}^n$. From this expression, it is clear that P is symmetric and has an eigenvalue 1 with m-fold degeneracy and an eigenvalue 0 with $(n-m)$-fold degeneracy. This is also true of $P^{(1)}, P^{(2)}, \ldots$. Now, since $\lim_{k\to\infty} \|P^{(k)} - P^{(\infty)}\| = 0$, each element of $P^{(k)}$ converges to the corresponding element of $P^{(\infty)}$. Then, since the eigenvalues of a matrix is a continuous function of its elements, $P^{(\infty)}$ also has an eigenvalue 1 with m-fold degeneracy and an eigenvalue 0 with $(n-m)$-fold degeneracy. Also, it is clear that $P^{(\infty)}$ is symmetric. Hence, by letting m orthonormal eigenvectors of $P^{(\infty)}$ corresponding to the eigenvalue 1 be $\mathbf{y}_1^{(\infty)}, \mathbf{y}_2^{(\infty)}, \ldots, \mathbf{y}_m^{(\infty)}$ and $Y^{(\infty)} \equiv [\mathbf{y}_1^{(\infty)}, \mathbf{y}_2^{(\infty)}, \ldots, \mathbf{y}_m^{(\infty)}]$, we can write $P^{(\infty)}$ as $P^{(\infty)} = Y^{(\infty)}(Y^{(\infty)})^\top$. This means that $P^{(\infty)} \in \mathcal{P}_{n,m}$. The lemma for the 2-norm follows easily because for finite-dimensional matrices, the 2-norm and the Frobenius norm is equivalent, i.e. $\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2$ for any $A \in \mathbb{R}^{n\times n}$. □

2.2 Convergence of the diagonal elements of $A^{(k)}$

Regarding the convergence of the diagonal elements of $A^{(k)}$, we have the following theorem. Note that here we do not make any assumptions on the distribution of the eigenvalues of A. Hence, the cases of multiple or clustered eigenvalues are also considered. The theorem will be proved by combining Corollary 2.3 with a technique developed in [11] for the point Jacobi eigenvalue algorithm.

Theorem 2.7. *Assume that $\tilde{P}^{(k)}$ at each step is computed in such a way that the diagonal elements of $\tilde{A}^{(k+1)} = (\tilde{P}^{(k)})^\top \tilde{A}^{(k)} \tilde{P}^{(k)}$ is ordered non-increasingly. Then, as $k \to \infty$, $A^{(k)}$ converges to a diagonal matrix D. The diagonal elements of D are the eigenvalues of A.*

Proof. Consider the transition from $A^{(k)}$ to $A^{(k+1)}$. The only diagonal elements that change are those belonging to $\tilde{A}^{(k)}$. Let the diagonal elements of $\tilde{A}^{(k)}$ and $\tilde{A}^{(k+1)}$ be denoted by $\{\tilde{a}_{qq}^{(k)}\}_{q=1}^{2\ell}$ and $\{\tilde{a}_{qq}^{(k+1)}\}_{q=1}^{2\ell}$, respectively. To bound the change $|\tilde{a}_{qq}^{(k+1)} - \tilde{a}_{qq}^{(k)}|$, we use Corollary 2.3. Let

$$B \equiv \tilde{A}^{(k)}, \quad C \equiv \begin{pmatrix} A_{XX}^{(k)} & O \\ O & A_{YY}^{(k)} \end{pmatrix}. \tag{22}$$

Then, since $\tilde{A}^{(k)}$ and $\tilde{A}^{(k+1)}$ are similar (see Eq. (4)) and $\tilde{A}^{(k+1)}$ is diagonal, the eigenvalues of B are $\{\tilde{a}_{qq}^{(k+1)}\}_{q=1}^{2\ell}$. The eigenvalues of C are $\{\tilde{a}_{qq}^{(k)}\}_{q=1}^{2\ell}$ since $A_{XX}^{(k)}$ and $A_{YY}^{(k)}$ are diagonal. Moreover, both of these diagonal elements are ordered non-increasingly from the assumption. Hence, we can use Corollary 2.3 to obtain

$$\begin{aligned} \sum_{q=1}^{2\ell} |\tilde{a}_{qq}^{(k+1)} - \tilde{a}_{qq}^{(k)}|^2 &\leq \|B - C\|_F^2 \\ &\leq \|A_{XY}^{(k)}\|_F^2 + \|A_{YX}^{(k)}\|_F^2 \\ &\leq \|\mathrm{off}(A^{(k)})\|^2 \leq S\alpha^k, \end{aligned} \tag{23}$$

where we used Eq. (8) in the last inequality. Combining this with the fact that the diagonal elements of $A^{(k)}$ not contained in $\tilde{A}^{(k)}$ are unchanged, we have

$$|a_{ii}^{(k+1)} - a_{ii}^{(k)}| \leq \sqrt{S\alpha^k} \quad (i = 1, 2, \ldots, n). \tag{24}$$

Now, let m be any nonnegative integer. Then,

$$|a_{ii}^{(k+m)} - a_{ii}^{(k)}| \leq \sum_{j=k}^{k+m-1} |a_{ii}^{(j+1)} - a_{ii}^{(j)}| \leq \sum_{j=k}^{k+m-1} \sqrt{S\alpha^j} = \sqrt{S\alpha^k} \cdot \frac{1 - (\sqrt{\alpha})^m}{1 - \sqrt{\alpha}} < \frac{\sqrt{S\alpha^k}}{1 - \sqrt{\alpha}}, \tag{25}$$

showing that $\{a_{ii}^{(k)}\}_{k=1}^{\infty}$ is a Cauchy sequence for $1 \leq i \leq n$. Hence, it converges to a constant d_i.

The next task is to show that d_i is an eigenvalue of A. Let $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. Then,

$$\lim_{k\to\infty} A^{(k)} = \lim_{k\to\infty} (Q^{(k)})^\top A Q^{(k)} = D. \tag{26}$$

Here, $\{Q^{(k)}\}_{k=0}^{\infty}$ is generally not convergent. However, since $Q^{(k)}$ is an orthogonal matrix for any k and the set of orthogonal matrices is a compact set, we can choose a subsequence $\{Q^{(k_p)}\}_{p=1}^{\infty}$ that converges to some orthogonal matrix Q. For this subsequence, it also holds that

$$\lim_{p\to\infty} (Q^{(k_p)})^\top A Q^{(k_p)} = D. \tag{27}$$

Since the left-hand side is equal to $Q^\top A Q$, we finally obtain

$$Q^\top A Q = D \tag{28}$$

for a diagonal matrix D and an orthogonal matrix Q. This shows that the diagonal elements $d_1, d_2, \ldots, d_n$ are the eigenvalues of A. □

2.3 Convergence of the columns of $Q^{(k)}$ corresponding to simple eigenvalues

Let $\mathbf{q}_i^{(k)}$ be the ith column of $Q^{(k)}$. Also, as in Theorem 2.7, let d_i denote the eigenvalue of A to which $a_{ii}^{(k)}$ converges. In this subsection, we consider the convergence of $\mathbf{q}_i^{(k)}$ in the case where d_i is a simple eigenvalue. Note that other eigenvalues of A can be multiple or clustered. We have the following theorem.

Theorem 2.8. *Under the same assumption as in Theorem 2.7, if d_i is a simple eigenvalue of A, then $\mathbf{q}_i^{(k)}$ converges to the eigenvector of A corresponding to d_i as $k \to \infty$.*

Proof. Let δ_i be the smallest distance from d_i to other eigenvalues of A and k_0 be an integer such that $\|\mathrm{off}(A^{(k)})\|_F \leq \delta_i/4$ if $k \geq k_0$. In the following, we assume that $k \geq k_0$ and consider the transition from $\mathbf{q}_i^{(k)}$ to $\mathbf{q}_i^{(k+1)}$.

From Eq. (5), $\mathbf{q}_i^{(k+1)}$ is the ith column of $Q^{(k)}P^{(k)}$. We consider the case where the column i belongs to either column block X or Y; otherwise, $\mathbf{q}_i^{(k+1)} = \mathbf{q}_i^{(k)}$ since $P^{(k)}$ is identical to I_n except for the Xth and Yth column blocks. Let the Xth and Yth column blocks of $Q^{(k)}$ be denoted by $Q_X^{(k)}$ and $Q_Y^{(k)}$, respectively, and the local column index of $\mathbf{q}_i^{(k)}$ within the $n \times (2\ell)$ matrix $\left(Q_X^{(k)}\ Q_Y^{(k)}\right)$ be q. Then we have

$$\mathbf{q}_i^{(k)} = \left(Q_X^{(k)}\ Q_Y^{(k)}\right)\tilde{\mathbf{e}}_q, \quad \mathbf{q}_i^{(k+1)} = \left(Q_X^{(k)}\ Q_Y^{(k)}\right)\tilde{\mathbf{p}}_q^{(k)}, \tag{29}$$

where $\tilde{\mathbf{e}}_q$ and $\tilde{\mathbf{p}}_q^{(k)}$ are the qth columns of $I_{2\ell}$ and $\tilde{P}^{(k)}$, respectively. By noting that the matrix $\left(Q_X^{(k)}\ Q_Y^{(k)}\right)$ has orthonormal columns, we obtain

$$\|\mathbf{q}_i^{(k+1)} - \mathbf{q}_i^{(k)}\| = \left\|\left(Q_X^{(k)}\ Q_Y^{(k)}\right)(\tilde{\mathbf{p}}_q^{(k)} - \tilde{\mathbf{e}}_q)\right\| = \|\tilde{\mathbf{p}}_q^{(k)} - \tilde{\mathbf{e}}_q\|. \tag{30}$$

To bound the right-hand side, we use Theorem 2.4. By putting $B = \tilde{A}^{(k)}$ and $\mathbf{y} = \tilde{\mathbf{e}}_q$ in Theorem 2.4, we have

$$\begin{aligned} \rho &= \mathbf{y}^\top B\mathbf{y} = \tilde{a}_{qq}^{(k)} = a_{ii}^{(k)}, \\ \mathbf{r}(\mathbf{y}) &= B\mathbf{y} - \rho\mathbf{y} = \tilde{\mathbf{a}}_q^{(k)} - \tilde{a}_{qq}^{(k)}\tilde{\mathbf{e}}_q, \end{aligned} \tag{31}$$

where $\tilde{\mathbf{a}}_q^{(k)}$ is the qth column of $\tilde{A}^{(k)}$. Hence,

$$\|\mathbf{r}(\mathbf{y})\| = \left(\sum_{\substack{j=1\\ j\neq q}}^{2\ell}\left(\tilde{a}_{jq}^{(k)}\right)^2\right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{2}}\|\mathrm{off}(\tilde{A}^{(k)})\|_F \leq \frac{1}{\sqrt{2}}\|\mathrm{off}(A^{(k)})\|_F \leq \sqrt{\frac{1}{2}S\alpha^k}. \tag{32}$$

On the other hand, from Eq. (23), we obtain

$$|\tilde{a}_{qq}^{(k+1)} - \tilde{a}_{qq}^{(k)}| = |a_{ii}^{(k+1)} - a_{ii}^{(k)}| \leq \|\mathrm{off}(A^{(k)})\|_F \leq \frac{\delta_i}{4}. \tag{33}$$

It also holds that $|\tilde{a}_{qq}^{(k)} - d_i| = |a_{ii}^{(k)} - d_i| \leq \delta_i/4$. To see this, consider the Hoffman-Wielandt theorem with $B = A^{(k)}$ and $C = \mathrm{diag}(a_{11}^{(k)}, a_{22}^{(k)}, \ldots, a_{nn}^{(k)})$. $a_{ii}^{(k)}$ is an eigenvalue of C and d_i is an eigenvalue of B. If $|a_{ii}^{(k)} - d_i| > \delta_i/4$, there must be another eigenvalue d_j of B, which is distant from d_i at least by δ_i, such that $|a_{ii}^{(k)} - d_j| \leq \delta_i/4$. But, then, due to the inequality (33),

$$|a_{ii}^{(k+1)} - d_i| \geq |d_j - d_i| - |a_{ii}^{(k+1)} - a_{ii}^{(k)}| - |a_{ii}^{(k)} - d_j| \geq \frac{\delta_i}{2} > \frac{\delta_i}{4}. \tag{34}$$

Hence, it follows that $|a_{ii}^{(k')} - d_i| > \delta_i/4$ for any $k' \geq k$, but this contradicts to the assumption that $a_{ii}^{(k)}$ converges to d_i.

Now we consider another diagonal element $\tilde{a}_{rr}^{(k+1)}$ of $\tilde{A}^{(k+1)}$, where $r \neq q$. Then, again from the Hoffman-Wielandt theorem, with $B = A^{(k+1)}$ and $C = \mathrm{diag}(a_{11}^{(k+1)}, a_{22}^{(k+1)}, \ldots, a_{nn}^{(k+1)})$, there must be another eigenvalue d_s of A (and therefore of $A^{(k+1)}$) such that $|\tilde{a}_{rr}^{(k+1)} - d_s| \leq \delta_i/4$, because d_i must be associated with $\tilde{a}_{qq}^{(k+1)}$. Hence,

$$|\tilde{a}_{rr}^{(k+1)} - \tilde{a}_{qq}^{(k)}| \geq |d_i - d_s| - |\tilde{a}_{rr}^{(k+1)} - d_s| - |\tilde{a}_{qq}^{(k)} - d_i| \geq \delta_i/2. \tag{35}$$

Eqs. (33) and (35) show that $\tilde{a}_{qq}^{(k+1)} = a_{ii}^{(k+1)}$ is the eigenvalue of $\tilde{A}^{(k)}$ closest to $\tilde{a}_{qq}^{(k)}$ and all other eigenvalues are distant from $\tilde{a}_{qq}^{(k)}$ by at least $\delta_i/2$. Thus, $\mathrm{gap}(\rho) \geq \delta_i/2$. Noting that the eigenvector of $\tilde{A}^{(k)}$ corresponding to the eigenvalue $\tilde{a}_{qq}^{(k+1)}$ is $\tilde{\mathbf{p}}_q^{(k)}$ and defining $\theta = \angle(\tilde{\mathbf{e}}_q, \tilde{\mathbf{p}}_q^{(k)})$, we obtain from Theorem 2.4,

$$|\sin\theta| \leq \frac{\|\mathbf{r}(\mathbf{y})\|}{\mathrm{gap}(\rho)} \leq \frac{\sqrt{2S\alpha^k}}{\delta_i}. \tag{36}$$

Note that, since we have chosen the sign of the eigenvector $\tilde{\mathbf{p}}_q^{(k)}$ so that its qth element is nonnegative (see subsection 2.1), the inner product $(\tilde{\mathbf{p}}_q^{(k)})^\top \tilde{\mathbf{e}}_q$ is nonnegative and hence $\cos\theta \geq 0$. Inserting these results into (30) gives

$$\begin{aligned} \|\mathbf{q}_i^{(k+1)} - \mathbf{q}_i^{(k)}\|^2 &= \|\tilde{\mathbf{p}}_q^{(k)} - \tilde{\mathbf{e}}_q\|^2 \\ &= \|\tilde{\mathbf{p}}_q^{(k)}\|^2 - 2\left(\tilde{\mathbf{p}}_q^{(k)}\right)^\top \tilde{\mathbf{e}}_q + \|\tilde{\mathbf{e}}_q\|^2 \\ &= 2(1 - \cos\theta) \\ &\leq 2(1 - \cos^2\theta) = 2\sin^2\theta \leq \frac{4S\alpha^k}{\delta_i^2}, \end{aligned} \tag{37}$$

where we used $1 \leq 1 + \cos\theta$ in the first inequality. Now, let m be any nonnegative integer. Then,

$$\|\mathbf{q}_i^{(k+m)} - \mathbf{q}_i^{(k)}\| \leq \sum_{j=k}^{k+m-1} \|\mathbf{q}_i^{(j+1)} - \mathbf{q}_i^{(j)}\| \leq \sum_{j=k}^{k+m-1} \frac{2\sqrt{S\alpha^j}}{\delta_i} < \frac{2\sqrt{S\alpha^k}}{\delta_i(1 - \sqrt{\alpha})}. \tag{38}$$

This shows that $\left\{\mathbf{q}_i^{(k')}\right\}_{k'=k}^{\infty}$ is a Cauchy sequence and therefore converges to a constant vector $\mathbf{q}_i$ as $k \to \infty$.

The remaining task is to show that $\mathbf{q}_i$ is the eigenvector of A belonging to the eigenvalue d_i. Let us recall the subsequence $\{Q^{(k_p)}\}_{p=1}^{\infty}$ defined in the proof of Theorem 2.7. It converges to an orthogonal matrix Q satisfying Eq. (28). Moreover, since it is a subsequence of $\{Q^{(k)}\}_{k=0}^{\infty}$, its ith column converges to $\mathbf{q}_i$. Hence, by premultiplying both sides of Eq. (28) by Q and extracting the ith column, we have

$$A\mathbf{q}_i = d_i\mathbf{q}_i, \tag{39}$$

which shows that $\mathbf{q}_i$ is the eigenvector of A corresponding to d_i. □

2.4 The case of multiple and clustered eigenvalues

Let d_i be the eigenvalue of A to which $a_{ii}^{(k)}$ converges. If d_i is not a simple eigenvalue, the ith column of $Q^{(k)}$, which we denote by $\mathbf{q}_i^{(k)}$, does not converge to a fixed vector in general. Even if d_i is simple, if the distance δ_i to the nearest eigenvalue is very small, the convergence of $\mathbf{q}^{(k)}$ to the eigenvector can be very slow, because the bound in Eq. (38) becomes very large. To deal with these cases, in this subsection, we consider the convergence of a subspace spanned by multiple columns of $Q^{(k)}$. In the case of multiple eigenvalues, it is natural to consider the corresponding eigenspace rather than individual eigenvectors. Also, in the case of clustered eigenvalues, the convergence of the corresponding eigenspace can be much faster than that of individual eigenvectors if the cluster is separated from the rest of the eigenvalues.

To be specific, let us define the set of s consecutive column indices, $\mathcal{I} = \{i, i+1, \ldots, i+s-1\}$, which is confined within one column block, and consider the convergence of the subspace spanned by $Q_{\mathcal{I}}^{(k)} = \{\mathbf{q}_i^{(k)}, \mathbf{q}_{i+1}^{(k)}, \ldots, \mathbf{q}_{i+s-1}^{(k)}\}$. To this end, we consider the convergence of the orthogonal projector $\mathcal{P}_{\mathcal{I}}^{(k)} = Q_{\mathcal{I}}^{(k)}(Q_{\mathcal{I}}^{(k)})^\top$ onto this subspace. This is because unlike the orthonormal basis, which has a freedom of rotation within the subspace, the orthogonal projector is determined uniquely from the subspace and there is one-to-one correspondence between them. Now we prove the following theorem.

Theorem 2.9. *Under the same assumption as in Theorem 2.7, assume further that there exists $\delta_{\mathcal{I}} > 0$ and for any indices $i \in \mathcal{I}$ and $j \notin \mathcal{I}$, $|d_i - d_j| \geq \delta_{\mathcal{I}}$. Then, as $k \to \infty$, the orthogonal projector $\mathcal{P}_{\mathcal{I}}^{(k)} = Q_{\mathcal{I}}^{(k)}(Q_{\mathcal{I}}^{(k)})^\top$ converges to an orthogonal projector onto the subspace $\mathcal{S}_{\mathcal{I}}$ spanned by the eigenvectors corresponding to $d_i, d_{i+1}, \ldots, d_{i+s-1}$.*

Proof. Let k_0 be an integer such that $\|\mathrm{off}(A^{(k)})\|_F \leq \delta_{\mathcal{I}}/5$ if $k \geq k_0$. In the following, we assume that $k \geq k_0$ and consider the transition from $\mathcal{P}_{\mathcal{I}}^{(k)}$ to $\mathcal{P}_{\mathcal{I}}^{(k+1)}$.

From Eq. (5), $Q_{\mathcal{I}}^{(k+1)}$ consists of the ith through $(i+s-1)$th columns of $Q^{(k)}P^{(k)}$. We consider the case where these columns are contained in the column block X. The case where these columns are contained in the column block Y can be treated similarly. When they are contained neither in column block X nor Y, $Q_{\mathcal{I}}^{(k+1)} = Q_{\mathcal{I}}^{(k)}$ since $P^{(k)}$ is identical to I_n except for the Xth and Yth column blocks. Let the Xth and Yth column blocks of $Q^{(k)}$ be denoted by $Q_X^{(k)}$ and $Q_Y^{(k)}$, respectively, and the local column index of $\mathbf{q}_i^{(k)}, \mathbf{q}_{i+1}^{(k)}, \ldots, \mathbf{q}_{i+s-1}^{(k)}$ within the $n \times (2\ell)$ matrix $\begin{pmatrix} Q_X^{(k)} & Q_Y^{(k)} \end{pmatrix}$ be $q, q+1, \ldots, q+s-1$, respectively. Moreover, let $\begin{pmatrix} Q_X^{(k)} & Q_Y^{(k)} \end{pmatrix}$ and $\tilde{P}^{(k)}$ be partitioned as

$$\begin{pmatrix} Q_X^{(k)} & Q_Y^{(k)} \end{pmatrix} = \begin{pmatrix} \tilde{Q}_{\mathcal{J}_1}^{(k)} & \tilde{Q}_{\mathcal{I}}^{(k)} & \tilde{Q}_{\mathcal{J}_2}^{(k)} \end{pmatrix}, \quad \tilde{P}^{(k)} = \begin{pmatrix} \tilde{P}_{\mathcal{J}_1\mathcal{J}_1}^{(k)} & \tilde{P}_{\mathcal{J}_1\mathcal{I}}^{(k)} & \tilde{P}_{\mathcal{J}_1\mathcal{J}_2}^{(k)} \\ \tilde{P}_{\mathcal{I}\mathcal{J}_1}^{(k)} & \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} & \tilde{P}_{\mathcal{I}\mathcal{J}_2}^{(k)} \\ \tilde{P}_{\mathcal{J}_2\mathcal{J}_1}^{(k)} & \tilde{P}_{\mathcal{J}_2\mathcal{I}}^{(k)} & \tilde{P}_{\mathcal{J}_2\mathcal{J}_2}^{(k)} \end{pmatrix}, \tag{40}$$

where the first, second and third column blocks in the right-hand side (and also row blocks for $\tilde{P}^{(k)}$) consist of $q-1$, s and $2\ell-(q-1)-s$ columns (rows), respectively. Then we have

$$Q_{\mathcal{I}}^{(k+1)} = \begin{pmatrix} \tilde{Q}_{\mathcal{J}_1}^{(k)} & \tilde{Q}_{\mathcal{I}}^{(k)} & \tilde{Q}_{\mathcal{J}_2}^{(k)} \end{pmatrix} \begin{pmatrix} \tilde{P}_{\mathcal{J}_1\mathcal{I}}^{(k)} \\ \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \\ \tilde{P}_{\mathcal{J}_2\mathcal{I}}^{(k)} \end{pmatrix} = \tilde{Q}_{\mathcal{I}}^{(k)} \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} + \tilde{Q}_{\mathcal{J}}^{(k)} \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)}, \tag{41}$$

where

$$\tilde{Q}_{\mathcal{J}}^{(k)} \equiv \begin{pmatrix} \tilde{Q}_{\mathcal{J}_1}^{(k)} & \tilde{Q}_{\mathcal{J}_2}^{(k)} \end{pmatrix} \in \mathbb{R}^{n\times(2\ell-s)}, \quad \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \equiv \begin{pmatrix} \tilde{P}_{\mathcal{J}_1\mathcal{I}}^{(k)} \\ \tilde{P}_{\mathcal{J}_2\mathcal{I}}^{(k)} \end{pmatrix} \in \mathbb{R}^{(2\ell-s)\times s}. \tag{42}$$

From Eq. (41), we obtain

$$\begin{aligned}
\left\| \mathcal{P}_{\mathcal{I}}^{(k+1)} - \mathcal{P}_{\mathcal{I}}^{(k)} \right\|_2 &= \left\| Q_{\mathcal{I}}^{(k+1)} (Q_{\mathcal{I}}^{(k+1)})^\top - Q_{\mathcal{I}}^{(k)} (Q_{\mathcal{I}}^{(k)})^\top \right\|_2 \\
&= \left\| \left(\tilde{Q}_{\mathcal{I}}^{(k)} \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} + \tilde{Q}_{\mathcal{J}}^{(k)} \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right) \left(\tilde{Q}_{\mathcal{I}}^{(k)} \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} + \tilde{Q}_{\mathcal{J}}^{(k)} \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top - \tilde{Q}_{\mathcal{I}}^{(k)} \left(\tilde{Q}_{\mathcal{I}}^{(k)} \right)^\top \right\|_2 \\
&\le \left\| \tilde{Q}_{\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top - I_s \right) \left(\tilde{Q}_{\mathcal{I}}^{(k)} \right)^\top \right\|_2 + \left\| \tilde{Q}_{\mathcal{I}}^{(k)} \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \left(\tilde{Q}_{\mathcal{J}}^{(k)} \right)^\top \right\|_2 \\
&\quad + \left\| \tilde{Q}_{\mathcal{J}}^{(k)} \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top \left(\tilde{Q}_{\mathcal{I}}^{(k)} \right)^\top \right\|_2 + \left\| \tilde{Q}_{\mathcal{J}}^{(k)} \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \left(\tilde{Q}_{\mathcal{J}}^{(k)} \right)^\top \right\|_2 \\
&\le \left\| \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top - I_s \right\|_2 + \left\| \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \right\|_2 \\
&\quad + \left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top \right\|_2 + \left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \right\|_2,
\end{aligned} \tag{43}$$

where we used the fact that both $\tilde{Q}_{\mathcal{I}}^{(k)}$ and $\tilde{Q}_{\mathcal{J}}^{(k)}$ has orthonormal columns and therefore $\left\| \tilde{Q}_{\mathcal{I}}^{(k)} \right\|_2 = \left\| \tilde{Q}_{\mathcal{J}}^{(k)} \right\|_2 = 1$. To further simplify the last expression, we use the singular value decomposition $\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} = U\Sigma V^\top$ and the fact that $\left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} + \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} = I_s$. Then,

$$\begin{aligned}
\left\| \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top - I_s \right\|_2 &= \left\| U(\Sigma^2 - I_s)U^\top \right\|_2 \\
&= \left\| \Sigma^2 - I_s \right\|_2 \\
&= \left\| V(\Sigma^2 - I_s)V^\top \right\|_2 \\
&= \left\| \left(\tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right)^\top \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} - I_s \right\|_2 = \left\| \left(\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right)^\top \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right\|_2.
\end{aligned} \tag{44}$$

Inserting this into (43) and using $\left\| \tilde{P}_{\mathcal{I}\mathcal{I}}^{(k)} \right\|_2 \le 1$ and $\left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right\|_2 \le 1$ gives

$$\left\| \mathcal{P}_{\mathcal{I}}^{(k+1)} - \mathcal{P}_{\mathcal{I}}^{(k)} \right\|_2 \le 2 \left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right\|_2^2 + 2 \left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right\|_2 \le 4 \left\| \tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)} \right\|_2. \tag{45}$$

To bound $\left\|\tilde{P}^{(k)}_{\mathcal{J}\mathcal{I}}\right\|_2$, we use Lemma 2.5. Let us partition $\tilde{A}^{(k)}$ in the same way as $\tilde{P}^{(k)}$ is partitioned in Eq. (40):

$$\tilde{A}^{(k)} = \begin{pmatrix} \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_1} & \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{I}} & \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_2} \\ \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_1} & \tilde{A}^{(k)}_{\mathcal{I}\mathcal{I}} & \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_2} \\ \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_1} & \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{I}} & \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_2} \end{pmatrix}. \tag{46}$$

We first apply Lemma 2.5 with

$$B_{11} = \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_1}, \quad B_{12} = \left(\tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{I}} \ \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_2}\right), \quad B_{21} = \begin{pmatrix} \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_1} \\ \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_1} \end{pmatrix}, \quad B_{22} = \begin{pmatrix} \tilde{A}^{(k)}_{\mathcal{I}\mathcal{I}} & \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_2} \\ \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{I}} & \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_2} \end{pmatrix}. \tag{47}$$

Then, since we have assumed that the columns belonging to the index set $\mathcal{I}$ are contained in the column block X, the columns belonging to the index set $\mathcal{J}_1$ are also contained in the column block X, and therefore $B_{11} = \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_1}$ is diagonal. Moreover, from Corollary 2.3, we have

$$\lambda_{\min}(B_{11}) = \tilde{a}^{(k)}_{q-1,q-1} = a^{(k)}_{i-1,i-1} \geq d_{i-1} - \left\|\mathrm{off}(A^{(k)})\right\|_F \geq d_{i-1} - \delta_{\mathcal{I}}/5. \tag{48}$$

On the other hand, we have from Corollary 2.3,

$$\tilde{a}^{(k)}_{q,q} = a^{(k)}_{i,i} \leq d_i + \left\|\mathrm{off}(A^{(k)})\right\|_F \leq d_i + \frac{\delta_{\mathcal{I}}}{5}. \tag{49}$$

Hence, again from Corollary 2.3

$$\lambda_{\max}(B_{22}) \leq \tilde{a}^{(k)}_{q,q} + \|\mathrm{off}(B_{22})\|_F \leq d_i + \frac{2}{5}\delta_{\mathcal{I}}. \tag{50}$$

Since $d_{i-1} - d_i \geq \delta_{\mathcal{I}}$ from the assumption, Eqs. (48) and (50) show that η in Lemma 2.5 can be chosen as

$$\eta = \lambda_{\min}(B_{11}) - \lambda_{\max}(B_{22}) \geq d_{i-1} - d_i - \frac{3}{5}\delta_{\mathcal{I}} \geq \frac{2}{5}\delta_{\mathcal{I}}. \tag{51}$$

Also,

$$\sqrt{2}\|B_{21}\|_F \leq \left\|\mathrm{off}(A^{(k)})\right\|_F \leq \frac{\delta_{\mathcal{I}}}{5} < \eta \tag{52}$$

Hence, the assumptions of Lemma 2.5 are satisfied and we have from the lemma,

$$\left\|(\tilde{P}^{(k)}_{\mathcal{J}_1\mathcal{I}} \ \tilde{P}^{(k)}_{\mathcal{J}_1\mathcal{J}_2})\right\|_F = \|P_{12}\|_F \leq \frac{\|B_{21}\|_F}{\eta - \sqrt{2}\|B_{21}\|_F} \leq \frac{5\left\|\mathrm{off}(A^{(k)})\right\|_F}{\sqrt{2}\delta_{\mathcal{I}}} \leq \frac{5\sqrt{S}\alpha^k}{\sqrt{2}\delta_{\mathcal{I}}}. \tag{53}$$

Next, we apply Lemma 2.5 with

$$B_{11} = \begin{pmatrix} \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_1} & \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{I}} \\ \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_1} & \tilde{A}^{(k)}_{\mathcal{I}\mathcal{I}} \end{pmatrix}, \quad B_{12} = \begin{pmatrix} \tilde{A}^{(k)}_{\mathcal{J}_1\mathcal{J}_2} \\ \tilde{A}^{(k)}_{\mathcal{I}\mathcal{J}_2} \end{pmatrix}, \quad B_{21} = \left(\tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_1} \ \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{I}}\right), \quad B_{22} = \tilde{A}^{(k)}_{\mathcal{J}_2\mathcal{J}_2}. \tag{54}$$

Then, by noting that B_{11} is again diagonal and repeating the same argument as led to (53), we obtain

$$\left\|(\tilde{P}^{(k)}_{\mathcal{J}_2\mathcal{J}_1} \ \tilde{P}^{(k)}_{\mathcal{J}_2\mathcal{I}})\right\|_F \leq \frac{5\sqrt{S}\alpha^k}{\sqrt{2}\delta_{\mathcal{I}}}. \tag{55}$$

From Eqs. (45), (53) and (55), we have

$$
\begin{aligned}
\left\|\mathcal{P}_{\mathcal{I}}^{(k+1)}-\mathcal{P}_{\mathcal{I}}^{(k)}\right\|_2 \le 4\left\|\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)}\right\|_2 &\le 4\left\|\tilde{P}_{\mathcal{J}\mathcal{I}}^{(k)}\right\|_F \\
&\le 4\sqrt{\left\|(\tilde{P}_{\mathcal{J}_1\mathcal{I}}^{(k)}\ \ \tilde{P}_{\mathcal{J}_1\mathcal{J}_2}^{(k)})\right\|_F^2+\left\|(\tilde{P}_{\mathcal{J}_2\mathcal{J}_1}^{(k)}\ \ \tilde{P}_{\mathcal{J}_2\mathcal{I}}^{(k)})\right\|_F^2} \\
&\le \frac{20\sqrt{S\alpha^k}}{\delta_{\mathcal{I}}}.
\end{aligned} \tag{56}
$$

Thus, for any nonnegative integer m,

$$
\begin{aligned}
\left\|\mathcal{P}_{\mathcal{I}}^{(k+m)}-\mathcal{P}_{\mathcal{I}}^{(k)}\right\|_2 &\le \sum_{j=k}^{k+m-1}\left\|\mathcal{P}_{\mathcal{I}}^{(j+1)}-\mathcal{P}_{\mathcal{I}}^{(j)}\right\|_2 \\
&\le \sum_{j=k}^{k+m-1}\frac{20\sqrt{S\alpha^j}}{\delta_{\mathcal{I}}}=\frac{20\sqrt{S\alpha^k}}{\delta_{\mathcal{I}}}\sum_{j=0}^{m-1}(\sqrt{\alpha})^j \le \frac{20\sqrt{S\alpha^k}}{\delta_{\mathcal{I}}(1-\sqrt{\alpha})},
\end{aligned} \tag{57}
$$

showing that $\left\{\mathcal{P}_{\mathcal{I}}^{(k')}\right\}_{k'=k}^{\infty}$ is a Cauchy sequence. Since the set of orthogonal projectors from $\mathbb{R}^n$ to its s-dimensional subspace is complete, as shown in Lemma 2.6, this sequence converges to an orthogonal projector $\mathcal{P}_{\mathcal{I}}$ from $\mathbb{R}^n$ to its s-dimensional subspace.

The final task is to show that $\mathcal{P}_{\mathcal{I}}$ is the orthogonal projector onto the subspace $\mathcal{S}_{\mathcal{I}}$ spanned by the eigenvectors corresponding to $d_i, d_{i+1}, \ldots, d_{i+s-1}$. To this end, we recall the subsequence $\{Q^{(k_p)}\}_{p=1}^{\infty}$ defined in the proof of Theorem 2.7. Since it converges to an orthogonal matrix Q satisfying Eq. (28), its column vectors $\mathbf{q}_i^{(k_p)}, \mathbf{q}_{i+1}^{(k_p)}, \ldots, \mathbf{q}_{i+s-1}^{(k_p)}$ converge to the eigenvectors corresponding to $d_i, d_{i+1}, \ldots, d_{i+s-1}$, respectively, and therefore $\mathcal{P}_{\mathcal{I}}^{(k_p)} = Q_{\mathcal{I}}^{(k_p)}\left(Q_{\mathcal{I}}^{(k_p)}\right)^{\top}$, where $Q_{\mathcal{I}}^{(k_p)} = \{\mathbf{q}_i^{(k_p)}, \mathbf{q}_{i+1}^{(k_p)}, \ldots, \mathbf{q}_{i+s-1}^{(k_p)}\}$, converges to the orthogonal projector onto $\mathcal{S}_{\mathcal{I}}$. Noting that $\left\{\mathcal{P}_{\mathcal{I}}^{(k_p)}\right\}_{p=1}^{\infty}$ is a subsequence of the convergent sequence $\left\{\mathcal{P}_{\mathcal{I}}^{(k)}\right\}_{k=1}^{\infty}$, we can conclude that $\mathcal{P}_{\mathcal{I}} = \lim_{k\to\infty}\mathcal{P}_{\mathcal{I}}^{(k)}$ is the orthogonal projector onto $\mathcal{S}_{\mathcal{I}}$. □

References

[1] M. Bečka, G. Okša and M. Vajteršic, Dynamic ordering for a parallel block-Jacobi SVD algorithm. *Parallel Comput.*, Vol. 28, pp. 243–262 (2002).

[2] M. Bečka, G. Okša and M. Vajteršic, New dynamic orderings for the parallel one-sided block-Jacobi SVD algorithm. *Parallel Proc. Lett.*, Vol. 25, No. 2, 1550003 (2015).

[3] C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation: III*, *SIAM J. Num. Anal.*, Vol. 7, pp. 1–46 (1970).

[4] Z. Drmač, Global convergence proof for cyclic Jacobi methods with block rotations, *SIAM J. Matrix Anal. Appl.*, Vol. 31, pp. 1329–1350 (2010).

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins UP, Baltimore, 2013.

[6] A. J. Hoffman and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, *Duke Math. J.*, Vol. 20, pp. 37–39 (1953).

[7] G. Okša, Y. Yamamoto, and M. Vajteršic, Asymptotic quadratic convergence of the serial block-Jacobi EVD algorithm for Hermitian matrices, *Numerische Mathematik*, Vol. 136, No. 4, pp. 1071–1095 (2017).

[8] G. Okša, Y. Yamamoto, M. Bečka and M. Vajteršic, Asymptotic quadratic convergence of the parallel block-Jacobi EVD algorithm with dynamic ordering for Hermitian matrices. *BIT Numerical Mathematics*, Vol. 58, No. 4, pp. 1099–1123 (2018).

[9] G. Okša, Y. Yamamoto, M. Bečka and M. Vajteršic, Asymptotic quadratic convergence of the two-sided serial and parallel block-Jacobi SVD algorithm. *SIAM J. Matrix Anal. Appl.*, Vol. 40, No. 2, pp. 639–671 (2019).

[10] B. N. Parlett, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1987.

[11] T. Tsuchiya and K. Aishima, A note on convergence and a posteriori error estimates of the classical Jacobi method, *Nonlinear Theory and Its Applications, IEICE*, Vol. 6, pp. 391–396 (2015).

[12] Y. Yamamoto, Z. Lan, S. Kudo, Convergence analysis of the parallel classical block Jacobi method for the symmetric eigenvalue problem, *JSIAM Letters*, Vol. 6, pp. 57–60 (2014).