Journal of Mathematical Chemistry 14 (1993) 121–139

The extended negative factor counting method for tridiagonal block matrices with cross links

Yuan-Jie Ye

Department of Chemistry, Peking University, Beijing 100871, PR China and Institute for Theoretical Chemistry, Friedrich-Alexander-University Erlangen-Nürnberg, Egerlandstrasse 3, D-8250 Erlangen, Germany

Received 22 July 1992

The NFC (negative factor counting) method was extended to solve the eigenvalue problem of tridiagonal block matrices with elements corresponding to cross links which may be derived from the quantum-chemical calculation on a native protein molecule. The mathematical proof of the necessary theorem is given in detail.

1. Introduction

With the development of protein engineering, it is necessary that people obtain more knowledge about the relationship between the electronic structures and activities of biological molecules [1–3]. This may contribute significantly to the understanding of electron transfer involved in many biochemical reactions. The NFC (negative factor counting) method [4–6] is a useful tool of dealing with such aperiodic biopolymers to investigate this relationship. Ladik and co- workers [7–15] calculated some model protein molecules at the ab initio level in some simple case including correlation by the NFC method and discussed the level distributions as well as the localization properties of the frontier orbitals of these model protein molecules. It is more difficult to calculate a native protein in such manner, because there are interactions through cross links in the molecule between residues, although they are not near to each other in the primary sequence. Li et al. [16] pointed out that the proof of the NFC theorem by Dean and Martin [4] is not rigorous enough but their theorem is right, and have given a more exact mathematical proof for it. Based on this, the author has developed the extended negative factor counting (ENFC) method to deal with cross link problems. The details of the mathematical proof of the ENFC theorem is given in this article. Pig insulin, one of the smallest native proteins, is taken as the first example of the application of the

ENFC method in the ab initio scheme. The results are reported in a subsequent paper [17].

2. Derivation of the eigenvalue equation

Let us suppose a protein molecule consists of N amino acid residues, and each one of them is considered as a unit. Firstly, let us suppose that all orbitals of the units are known, and the nth unit has m_n orbitals. Thus all the molecular orbitals of the protein can be written as the linear combination of these unit orbitals,

$$\psi = \sum_{n=1}^{N} \sum_{j=1}^{m_n} C_j(n)\phi_j(n) . \tag{1}$$

Here $\phi_j(n)$ is the unit orbital, n denotes the unit and j a given AO in the unit. Using the Ritz's variational method, the following Hartree–Fock–Roothaan equation can be derived:

$$\mathbf{HC} = \mathbf{SCE} , \tag{2}$$

where $\mathbf{H}$ and $\mathbf{S}$ are the Fock matrix and the overlap matrix, respectively. For a protein, eq. (2) is a very large generalized eigenvalue equation. Its order can reach more than ten thousand. At the same time, a large number of elements in the matrices are almost equal to zero.

If the nearest neighbours approximation is applied, that is, only those interactions and overlaps between the units that are first neighbours (with chemical bond or van der Waals forces) are taken into account, the form of eq. (2) remains unchanged but its Fock matrix and overlap matrix will have the forms:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & & \mathbf{0} & & \mathbf{0} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \mathbf{H}_{23} & & & \\ & \ddots & \ddots & \ddots & \mathbf{H}_{n_1 n_1'} & \\ \mathbf{0} & & \ddots & \ddots & \ddots & \mathbf{H}_{n_L n_L'} \\ & \mathbf{H}_{n_1' n_1} & & \ddots & \ddots & \mathbf{0} \\ & & \ddots & & \mathbf{H}_{n_L' n_L} & \\ & & & \ddots & \ddots & \mathbf{H}_{N-1,N} \\ \mathbf{0} & & & \mathbf{0} & \mathbf{H}_{N,N-1} & \mathbf{H}_{NN} \end{pmatrix} , \tag{3}$$

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & & \mathbf{0} & & \mathbf{0} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} & & & \\ & \ddots & \ddots & \ddots & & \\ \mathbf{0} & & \ddots & \ddots & \mathbf{S}_{n_1 n_1'} & \\ & \mathbf{S}_{n_1' n_1} & & \ddots & \ddots & \mathbf{S}_{n_L n_L'} \\ & & \ddots & & \ddots & \mathbf{0} \\ & & & \mathbf{S}_{n_L' n_L} & & \ddots \\ & & & & \ddots & \mathbf{S}_{N-1,N} \\ \mathbf{0} & & & \mathbf{0} & \mathbf{S}_{N,N-1} & \mathbf{S}_{NN} \end{pmatrix}, \tag{4}$$

where, $\mathbf{H}_{nn}$ and $\mathbf{S}_{nn}$ are the matrix blocks of the interactions and overlaps within a unit; $\mathbf{H}_{n,n+1}$ and $\mathbf{S}_{n,n+1}$ are the matrix blocks of the interactions and overlaps between adjacent units in the molecule, and finally $\mathbf{H}_{n_i n_i'}$ and $\mathbf{S}_{n_i n_i'}$ are the matrix blocks of the interactions and overlaps between those units which are close in the three-dimensional conformation but are not adjacent in the primary sequence, such as the interactions and overlaps between two cysteins linked by a disulphur bridge. Let us assume in a macromolecule there are L such cases altogether, n_i is a row index and n_i' a column index, where $i = 1, \ldots, L$; $n_i < n_i'$, and the n_i's are ordered as $n_1' \leqslant n_2' \leqslant \ldots \leqslant n_L'$.

3. Numerical method

3.1. THE BASIC THEOREM

The mathematical proof of the extended negative factor counting method is based on the following theorem [16].

For the generalized eigenvalue equation,

$$\mathbf{H}C = \lambda \mathbf{S}C, \tag{5}$$

introducing a real parameter x, constructing a matrix $\mathbf{H} - x\mathbf{S}$ and partitioning it into a four-block form as

$$\mathbf{H} - x\mathbf{S} = \begin{pmatrix} \mathbf{A}_1 - x\mathbf{R}_1 & \mathbf{B}_2 - x\mathbf{S}_2 \\ \mathbf{B}_2^+ - x\mathbf{S}_2^+ & \mathbf{A}_2 - x\mathbf{R}_2 \end{pmatrix}, \tag{6}$$

the following theorem has been proven to be valid:

THEOREM 1

If a real parameter x is not an eigenvalue of eq. (5), then the number of the eigenvalues of eq. (5) in the interval $(-\infty, x]$ is equal to the numbers of the negative eigen-

values of the matrices $\mathbf{T}_1(x)$ and $\mathbf{T}_2(x)$ which meet the requirement of the following recursion formula:

$$\mathbf{T}_1 = \mathbf{A}_1 - x\mathbf{R}_1\,, \tag{7}$$

$$\mathbf{T}_2 = \mathbf{A}_2 - x\mathbf{R}_2 - (\mathbf{B}_2^+ - x\mathbf{S}_2^+)\mathbf{T}_1^{-1}(\mathbf{B}_2 - x\mathbf{S}_2)\,. \tag{8}$$

The mathematical proof of this theorem was given in ref. [16].

3.2. THE EXTENDED NEGATIVE FACTOR COUNTING THEOREM

In this subsection, the extended NFC theorem which is used to find the solution of the eqs. (2)–(4) will be proven in detail.

THEOREM 2 (extended negative factor counting theorem)

If a real parameter x is not an eigenvalue of the eqs. (2)–(4), then the number of the eigenvalues of the eqs. (2)– (4) in the interval $(-\infty, x]$ (abbreviated as NE(x)), is equal to the sum of the numbers of the negative eigenvalues of following matrix sequence $\mathbf{U}_n(x), n = 1, \ldots, N$ (abbreviated as NNE[$\mathbf{U}_n(x)$]). That is

$$\mathrm{NE}(x) = \sum_{n=1}^{N} \mathrm{NNE}[\mathbf{U}_n(x)]\,; \tag{9}$$

$$\mathbf{U}_1 = \mathbf{H}_{11} - x\mathbf{S}_{11}\,; \tag{10}$$

$$\mathbf{U}_n = (\mathbf{H}_{nn} - x\mathbf{S}_{nn}) - (\mathbf{H}_{n-1,n}^+ - x\mathbf{S}_{n-1,n}^+)\mathbf{U}_{n-1}^{-1}(\mathbf{H}_{n-1,n} - x\mathbf{S}_{n-1,n}),$$
$$n = 2, \ldots, N, \quad n \neq n_i', \quad i = 1, \ldots, L\,; \tag{11}$$

$$\mathbf{U}_{n_i'} = (\mathbf{H}_{n_i'n_i'} - x\mathbf{S}_{n_i'n_i'}) - [\mathbf{H}_{n_i'-1,n_i'}^+ - x\mathbf{S}_{n_i'-1,n_i'}^+ + (-1)^{K_i}\mathbf{M}_{K_i+1}^{(i)+}]$$
$$\times \mathbf{U}_{n_i'-1}^{-1}[\mathbf{H}_{n_i'-1,n_i'} - x\mathbf{S}_{n_i'-1,n_i'} + (-1)^{K_i}\mathbf{M}_{K_i+1}^{(i)}]$$
$$- \sum_{k=1}^{K_i} \mathbf{M}_k^{(i)+}\mathbf{U}_{n_i+k-1}^{-1}\mathbf{M}_k^{(i)}, \quad i = 1, \ldots, L\,; \tag{12}$$

$$K_i = n_i' - n_i - 1, \quad i = 1, \ldots, L\,; \tag{13}$$

$$\mathbf{M}_1^{(i)} = \mathbf{H}_{n_in_i'} - x\mathbf{S}_{n_in_i'}, \quad i = 1, \ldots, L\,; \tag{14}$$

$$\mathbf{M}_k^{(i)} = (\mathbf{H}_{n_i+k-2,n_i+k-1}^+ - x\mathbf{S}_{n_i+k-2,n_i+k-1}^+)\mathbf{U}_{n_i+k-2}^{-1}\mathbf{M}_{k-1}^{(i)}, \quad i = 1, \ldots, L\,. \tag{15}$$

For arbitrary $i, j = 1, \ldots, L$, if $i > j$ and $n_i < n_j'$, the corresponding $\mathbf{M}^{(i)}n_j' - n_i + 1$ in the matrix sequence $\mathbf{M}_k^{(i)}$ is

$$\mathbf{M}^{(i)}_{n'_j-n_i+1} = [\mathbf{H}^{+}_{n'_j-1,n'_j} - x\mathbf{S}^{+}_{n'_j-1,n'_j} + (-1)^{K_j}\mathbf{M}^{(j)+}_{K_j+1}]\mathbf{U}^{-1}_{n'_j-1}\mathbf{M}^{(i)}_{n'_j-n_i} + (-1)^{n_j+n_i+1}\sum_{k=1}^{n'_j-\max(n_i,n_j)-1}\mathbf{M}^{(j)+}_{k+\max(n_i-n_j,0)}\mathbf{U}^{-1}_{\max(n_i,n_j)+k-1}\mathbf{M}^{(i)}_{k+\max(n_j-n_i,0)}. \quad (16)$$

From now on the following notations will be used:

$$\mathbf{A}_n = \mathbf{H}_{nn} - x\mathbf{S}_{nn}, \quad n = 1, \ldots, N; \quad (17)$$

$$\mathbf{B}_n = \mathbf{H}_{n-1,n} - x\mathbf{S}_{n-1,n}, \quad n = 2, \ldots, N; \quad (18)$$

$$\mathbf{M}^{(i)}_1 = \mathbf{H}_{n_i n'_i} - x\mathbf{S}_{n_i n'_i}, \quad i = 1, \ldots, L. \quad (19)$$

The proof of theorem 2 will be taken in three steps because of the complexity caused by the cross-link blocks in the equation.

Proof

(i) Step 1

There is only one cross-link block in eqs. (2)–(4). They are noted by indices, n_1, n'_1. It is equivalent to the molecular system in fig. 1(a). Out of simplicity it is supposed that there is only one chain in the system. If there are more than one chains, the numbering of the units is arranged as $1, \ldots, N_1, N_1+1, \ldots, N$, and the block between N_1 and N_1+1 is set to zero. The generality of the results is not influenced.

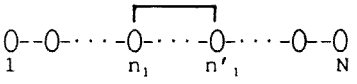
In this case

$$\mathbf{H} - x\mathbf{S} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_2^{+} & \mathbf{A}_2 & \mathbf{B}_3 & \\ & \ddots & \ddots & \mathbf{M}_1^{(1)} \\ \mathbf{0} & & & \mathbf{0} \\ & \mathbf{M}_1^{(1)+} & & \\ & & & \mathbf{B}_N \\ \mathbf{0} & & \mathbf{0} \quad \mathbf{B}_N^{+} & \mathbf{A}_N \end{pmatrix}. \quad (20)$$

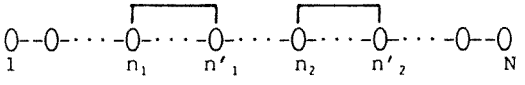
From theorem 1 it is known that the matrix $\mathbf{H} - x\mathbf{S}$ can be partitioned into four blocks as

$$\mathbf{H} - x\mathbf{S} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{G}_2 \\ \mathbf{G}_2^{+} & \mathbf{R}_2 \end{pmatrix}, \quad (21)$$

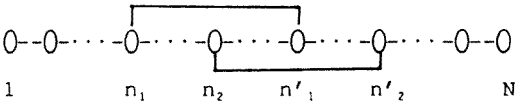
where



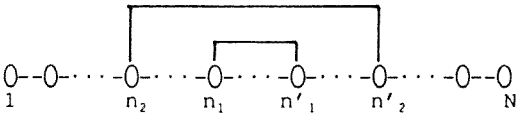
(a)



(b)



(c)



(d)

Fig. 1. The molecular systems have different kinds of cross-links.

$$\mathbf{G}_2 = (\mathbf{B}_2\ \mathbf{0}\ \ldots\ \mathbf{0}), \tag{22}$$

$$\mathbf{R}_2 = \begin{pmatrix} \mathbf{A}_2 & \mathbf{B}_3 & & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_3^+ & \mathbf{A}_3 & \mathbf{B}_4 & & \\ & \ddots & \ddots & \mathbf{M}_1^{(1)} & \\ \mathbf{0} & & \ddots & & \mathbf{0} \\ & \mathbf{M}_1^{(1)+} & & \ddots & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0}\ \ \mathbf{B}_N^+ & \mathbf{A}_N \end{pmatrix}. \tag{23}$$

One obtains formula (10) in theorem 2 at once as

$$\mathbf{U}_1 = \mathbf{H}_{11} - x\mathbf{S}_{11} \tag{24}$$

and the matrix $\mathbf{T}_2$ in theorem 1 will be

$$\begin{aligned} \mathbf{T}_2 &= \mathbf{R}_2 - \mathbf{G}_2^+ \mathbf{U}_1^{-1} \mathbf{G}_2 \\ &= \mathbf{R}_2 - \begin{pmatrix} \mathbf{B}_2^+ \mathbf{U}_1^{-1} \mathbf{B}_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & & & \\ \vdots & & & \\ \mathbf{0} & & & \end{pmatrix}. \end{aligned} \tag{25}$$

Taking the following recursion formula:

$$\mathbf{U}_n = \mathbf{A}_n - \mathbf{B}_n^+ \mathbf{U}_{n-1}^{-1} \mathbf{B}_n, \quad n \neq n_i', \tag{26}$$

the final form of matrix $\mathbf{T}_2$ is obtained as

$$\mathbf{T}_2 = \begin{pmatrix} \mathbf{U}_2 & \mathbf{B}_3 & & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_3^+ & \mathbf{A}_3 & \mathbf{B}_4 & & \\ & \ddots & \ddots & \mathbf{M}_1^{(1)} & \\ \mathbf{0} & & \ddots & & \mathbf{0} \\ & \mathbf{M}_1^{(1)+} & & \ddots & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0}\ \ \mathbf{B}_N^+ & \mathbf{A}_N \end{pmatrix}. \tag{27}$$

From theorem 1 it is known that the number of eigenvalues of eqs. (2)–(4) in the interval $(-\infty, x]$ is equal to the sum of the numbers of the negative eigenvalues of the matrices $\mathbf{U}_1(x)$ and $\mathbf{T}_2(x)$, that is

$$\mathrm{NE}(x) = \mathrm{NNE}(\mathbf{U}_1) + \mathrm{NNE}(\mathbf{T}_2). \tag{28}$$

Theorem 1 can be used to calculate the number of negative eigenvalues of the matrix $\mathbf{T}_2$. The eigenvalue equation which has to be solved as

$$\mathbf{T}_2 C = \lambda C \tag{29}$$

(that is, $\mathbf{S} = \mathbf{I}$ in eq. (5)), and one has to take into account only the negative eigenvalues. This can be done by introducing another real parameter x' and set $x' = 0$. In other words, one has to calculate the number of eigenvalues of eq. (29) in the interval $(-\infty, 0]$, so that only the matrix $\mathbf{T}_2$ will be partitioned as

$$\mathbf{T}_2 = \begin{pmatrix} \mathbf{U}_2 & \mathbf{G}_3 \\ \mathbf{G}_3^+ & \mathbf{R}_3 \end{pmatrix}, \tag{30}$$

where

$$\mathbf{G}_3 = (\mathbf{B}_3\ \mathbf{0}\ \ldots\ \mathbf{0}), \tag{31}$$

$$\mathbf{R}_3 = \begin{pmatrix} \mathbf{A}_3 & \mathbf{B}_4 & & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_4^+ & \mathbf{A}_4 & \mathbf{B}_5 & & \\ & \ddots & \ddots & \mathbf{M}_1^{(1)} & \\ \mathbf{0} & & \ddots & & \mathbf{0} \\ & \mathbf{M}_1^{(1)+} & & \ddots & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0}\ \ \mathbf{B}_N^+ & \mathbf{A}_N \end{pmatrix} \tag{32}$$

and the matrix $\mathbf{T}_3$ is obtained as

$$\mathbf{T}_3 = \mathbf{R}_3 - \mathbf{G}_3^+ \mathbf{U}_2^{-1} \mathbf{G}_3 \tag{33}$$

and again from theorem 1 it is known that

$$\mathrm{NNE}(\mathbf{T}_2) = \mathrm{NNE}(\mathbf{U}_2) + \mathrm{NNE}(\mathbf{T}_3)\,. \tag{34}$$

From recursion formula (26) it follows that

$$\mathbf{U}_3 = \mathbf{A}_3 - \mathbf{B}_3^+ \mathbf{U}_2^{-1} \mathbf{B}_3 \tag{35}$$

and the final form of the matrix $\mathbf{T}_3$ is

$$\mathbf{T}_3 = \begin{pmatrix} \mathbf{U}_3 & \mathbf{B}_4 & & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_4^+ & \mathbf{A}_4 & \mathbf{B}_5 & & \\ & \ddots & \ddots & \mathbf{M}_1^{(1)} & \\ \mathbf{0} & & \ddots & & \mathbf{0} \\ & \mathbf{M}_1^{(1)+} & & \ddots & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0}\ \ \mathbf{B}_N^+ & \mathbf{A}_N \end{pmatrix}. \tag{36}$$

The above steps can be repeated until the matrix $\mathbf{T}_{n_1}$ is obtained,

$$\mathbf{T}_{n_1} = \begin{pmatrix} \mathbf{U}_{n_1} & \mathbf{B}_{n_1+1} & & \mathbf{0} & \mathbf{M}_1^{(1)} & & \mathbf{0} \\ \mathbf{B}_{n_1+1}^{+} & \mathbf{A}_{n_1+1} & \mathbf{B}_{n_1+2} & & & & \\ & \ddots & \ddots & \ddots & & & \\ \mathbf{0} & & & & & & \\ & & & & & & \mathbf{0} \\ \mathbf{M}_1^{(1)+} & & & & & & \\ & & & & & & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0} & \mathbf{B}_N^{+} & & \mathbf{A}_N \end{pmatrix}. \tag{37}$$

One can write

$$\mathrm{NE}(x) = \mathrm{NNE}(\mathbf{U}_1) + \mathrm{NNE}(\mathbf{U}_2) + \ldots + \mathrm{NNE}(\mathbf{U}_{n_1-1}) + \mathrm{NNE}(\mathbf{T}_{n_1}). \tag{38}$$

When the matrix $\mathbf{T}_{n_1}$ is partitioned into four blocks according to $\mathbf{U}_{n_1}$,

$$\mathbf{T}_{n_1} = \begin{pmatrix} \mathbf{U}_{n_1} & \mathbf{G} \\ \mathbf{G}^{+} & \mathbf{R} \end{pmatrix}, \tag{39}$$

the following formulae will be obtained:

$$\mathbf{G} = (\mathbf{B}_{n_1+1}\ \mathbf{0}\ \ldots\ \mathbf{0}\ \mathbf{M}_1^{(1)}\ \mathbf{0}\ \ldots\ \mathbf{0}), \tag{40}$$

$$\mathbf{R} = \begin{pmatrix} \mathbf{A}_{n_1+1} & \mathbf{B}_{n_1+2} & & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{n_1+2}^{+} & \mathbf{A}_{n_1+2} & \mathbf{B}_{n_1+3} & & \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & & \mathbf{0} \\ & & & & \mathbf{B}_N \\ \mathbf{0} & & & \mathbf{0} & \mathbf{B}_N^{+} \quad \mathbf{A}_N \end{pmatrix}. \tag{41}$$

and

$$\mathbf{G}^{+}\mathbf{U}_{n_1}^{-1}\mathbf{G} = \begin{pmatrix} \mathbf{B}_{n_1+1}^{+}\mathbf{U}_{n_1}^{-1}\mathbf{B}_{n_1+1} & \mathbf{0} & \ldots & \mathbf{B}_{n_1+1}^{+}\mathbf{U}_{n_1}^{-1}\mathbf{M}_1^{(1)} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0} & & & \\ \vdots & & & \vdots & & & \\ \mathbf{M}_1^{(1)+}\mathbf{U}_{n_1}^{-1}\mathbf{B}_{n_1+1} & \mathbf{0} & \ldots & \mathbf{M}_1^{(1)+}\mathbf{U}_{n_1}^{-1}\mathbf{M}_1^{(1)} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0} & & & \\ \vdots & & & \vdots & & & \\ \mathbf{0} & & & \mathbf{0} & & & \end{pmatrix}. \tag{42}$$

Introducing the following recursion formula:

$$\mathbf{M}_{k+1}^{(1)} = \mathbf{B}_{n_1+k}^{+}\mathbf{U}_{n_1+k-1}^{-1}\mathbf{M}_{k}^{(1)} \tag{43}$$

and from the second formula in theorem 1,

$$\mathbf{T}_{n_1+1} = \mathbf{R} - \mathbf{G}^{+}\mathbf{U}_{n_1}^{-1}\mathbf{G}, \tag{44}$$

the final form of the matrix $\mathbf{T}_{n_1+1}$ then can be obtained as follows:

$$\mathbf{T}_{n_1+1} = \begin{pmatrix} \mathbf{U}_{n_1+1} & \mathbf{B}_{n_1+2} & \mathbf{0} & -\mathbf{M}_2^{(1)} & \mathbf{0} \\ \mathbf{B}_{n_1+2}^{+} & \mathbf{A}_{n_1+2} & \mathbf{B}_{n_1+3} & & \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & & \\ -\mathbf{M}_2^{(1)+} & & \mathbf{A}_{n_1'} - \boldsymbol{\Sigma}_1^{(1)} & & \mathbf{0} \\ & & & & \mathbf{B}_N \\ \mathbf{0} & & \mathbf{0} & \mathbf{B}_N^{+} & \mathbf{A}_N \end{pmatrix}, \tag{45}$$

where

$$\boldsymbol{\Sigma}_m^{(1)} = \sum_{k=1}^{m} \mathbf{M}_k^{(1)+}\mathbf{U}_{n_1+k-1}^{-1}\mathbf{M}_k^{(1)} \tag{46}$$

and the following equality is valid:

$$\mathrm{NE}(x) = \mathrm{NNE}(\mathbf{U}_1) + \mathrm{NNE}(\mathbf{U}_2) + \ldots + \mathrm{NNE}(\mathbf{U}_{n_1}) + \mathrm{NNE}(\mathbf{T}_{n_1+1}). \tag{47}$$

Comparing (45) with (37), it can be found that the differences between $\mathbf{T}_{n_1+1}$ and $\mathbf{T}_{n_1}$ are that a new block of $\mathbf{M}_2^{(1)}$ is introduced in the matrix $\mathbf{T}_{n_1+1}$, its sign is changed once, and the block $\mathbf{A}_{n_1'}$ in the matrix $\mathbf{T}_{n_1}$ is changed into $\mathbf{A}_{n_1'} - \mathbf{M}_1^{(1)+}\mathbf{U}_{n_1}^{-1}\mathbf{M}_1^{(1)}$ in the matrix $\mathbf{T}_{n_1+1}$.

When the matrix $\mathbf{T}_{n_1+1}$ is partitioned into four blocks according to $\mathbf{U}_{n_1+1}$ and theorem 1 is used again, the matrix $\mathbf{T}_{n_1+2}$ will be obtained and will have a similar form to matrix $\mathbf{T}_{n_1+1}$:

$$\mathbf{T}_{n_1+2} = \begin{pmatrix} \mathbf{U}_{n_1+2} & \mathbf{B}_{n_1+3} & & \mathbf{0} & \mathbf{M}_3^{(1)} & & \mathbf{0} \\ \mathbf{B}_{n_1+3}^{+} & \mathbf{A}_{n_1+3} & \mathbf{B}_{n_1+4} & & & & \\ & \ddots & \ddots & \ddots & & & \\ \mathbf{0} & & & & & & \\ & & & & & & \\ \mathbf{M}_3^{(1)+} & & & & \mathbf{A}_{n_1'} - \boldsymbol{\Sigma}_2^{(1)} & & \mathbf{0} \\ & & & & & \ddots & \mathbf{B}_N \\ \mathbf{0} & & & & \mathbf{0} & \mathbf{B}_N^{+} & \mathbf{A}_N \end{pmatrix}, \tag{48}$$

It is found that a new block $\mathbf{M}_3^{(1)}$ occurs, its sign has been changed once again, and the block in the diagonal position of n_1' is changed into

$$\mathbf{A}_{n_1'} - \sum_{k=1}^{2} \mathbf{M}_k^{(1)+} \mathbf{U}_{n_1+k-1}^{-1} \mathbf{M}_k^{(1)} . \tag{49}$$

Substituting for K_1

$$K_1 = n_1' - n_1 - 1 \tag{50}$$

it can be worked out that the above partitioning has to be carried out K_1 times with the help of the recursion formula (44) to obtain $\mathbf{T}_{n_1'-1}$. In each step a new block $\mathbf{M}_{k+1}^{(1)}$ is introduced and its sign is opposite to that of $\mathbf{M}_k^{(1)}$. The final form of the matrix $\mathbf{T}_{n_1'-1}$ will become

$$\mathbf{T}_{n_1'-1} = \begin{pmatrix} \mathbf{U}_{n_1'-1} & \mathbf{B}_{n_1'} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)} & & \mathbf{0} & & \mathbf{0} \\ \mathbf{B}_{n_1'}^{+} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)+} & \mathbf{A}_{n_1'} - \boldsymbol{\Sigma}_{K_1}^{(1)} & \mathbf{B}_{n_1'+1} & & & \\ & \mathbf{B}_{n_1'+1}^{+} & \mathbf{A}_{n_1'+1} & \ddots & & \\ \mathbf{0} & & \ddots & \ddots & & \mathbf{0} \\ & & & & & \mathbf{B}_N \\ \mathbf{0} & & & & \mathbf{B}_N^{+} & \mathbf{A}_N \end{pmatrix}. \tag{51}$$

Partitioning the matrix $\mathbf{T}_{n_1'-1}$ into four blocks according to $\mathbf{U}_{n_1'-1}$, one obtains

$$\mathbf{T}_{n_1'} = \begin{pmatrix} \mathbf{U}_{n_1'} & \mathbf{B}_{n_1'+1} & & \mathbf{0} \\ \mathbf{B}_{n_1'+1}^{+} & \mathbf{A}_{n_1'+1} & \ddots & \\ & \ddots & \ddots & \mathbf{B}_N \\ \mathbf{0} & & \mathbf{B}_N^{+} & \mathbf{A}_N \end{pmatrix}, \tag{52}$$

where

$$\mathbf{U}_{n_1'} = \mathbf{A}_{n_1'} - [\mathbf{B}_{n_1'}^{+} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)+}]\mathbf{U}_{n_1'-1}^{-1}[\mathbf{B}_{n_1'} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)}] - \sum_{k=1}^{K_1} \mathbf{M}_k^{(1)+}\mathbf{U}_{n_1+k-1}^{-1}\mathbf{M}_k^{(1)}. \tag{53}$$

This is just the formula (12) in theorem 2 in this case, and formula (47) now is extended to

$$\mathrm{NE}(x) = \mathrm{NNE}(\mathbf{U}_1) + \mathrm{NNE}(\mathbf{U}_2) + \ldots + \mathrm{NNE}(\mathbf{U}_{n_1'-1}) + \mathrm{NNE}(\mathbf{T}_{n_1'}). \tag{54}$$

Continuing the four-part partitioning as shown above, the matrices $\mathbf{U}_n$ $(n = n_1' + 1, \ldots, N)$ can be obtained after the following relationship will be valid:

$$\mathrm{NNE}(\mathbf{T}_{n_1'}) = \mathrm{NNE}(\mathbf{U}_{n_1'}) + \mathrm{NNE}(\mathbf{U}_{n_1'+1}) + \ldots + \mathrm{NNE}(\mathbf{U}_{N-1}) + \mathrm{NNE}(\mathbf{U}_N). \tag{55}$$

Putting this into eq. (54) the final result is obtained as

$$\mathrm{NE}(x) = \sum_{n=1}^{N} \mathrm{NNE}[\mathbf{U}_n(x)] \tag{56}$$

and thus the proof of theorem 2 in this case is completed.

(ii) Step 2

There are two cross-link blocks in eqs. (2)–(4). Their positions are marked as n_1, n_1', n_2, n_2'. There are altogether three different distributions in this case.

(2.a) $n_1 < n_1' < n_2 < n_2'$. This case is equivalent to the molecular system in fig. 1(b).

The proof of the theorem in this case is quite similar to step 1. The only difference is that the different sequences of the matrices $\mathbf{M}_k$ should be labelled as $\mathbf{M}_k^{(i)}$ $(i = 1, 2, k = 1, \ldots, K_i + 1)$.

(2.b) $n_1 < n_2 < n_1' < n_2'$. This is equivalent to the molecular system in fig. 1(c). In this case the four-block partitioning is the same as above before $\mathbf{T}_{n_2}$ is obtained. The matrix $\mathbf{T}_{n_2}$ is

$$\mathbf{T}_{n_2} = \begin{pmatrix} \mathbf{U}_{n_2} & \mathbf{Q} & \mathbf{0} & \mathbf{P}_1 & \mathbf{P}_2 & & \mathbf{0} \\ \mathbf{Q}^+ & \mathbf{A}_{n_2+1} & \mathbf{B}_{n_2+2} & & & & \\ \mathbf{0} & & \cdots & \cdots & & & \\ & & \cdots & \cdots & & & \\ \mathbf{P}_1^+ & & & \mathbf{A}_{n_1'} - \boldsymbol{\Sigma}^{(1)} & & & \\ & & & \cdots & \cdots & & \\ & & & \cdots & \cdots & & \\ \mathbf{P}_2^+ & & & & \mathbf{A}_{n_2'} & & \mathbf{0} \\ & & & & \cdots & \cdots & \\ & & & & & \cdots & \mathbf{B}_N \\ \mathbf{0} & & & & \mathbf{0} & \mathbf{B}_N^+ & \mathbf{A}_N \end{pmatrix}, \tag{57}$$

where

$$\mathbf{Q} = \mathbf{B}_{n_2+1}, \tag{58}$$

$$\mathbf{P}_1 = (-1)^{n_2-n_1}\mathbf{M}^{(1)}_{n_2-n_1+1}, \tag{59}$$

$$\mathbf{P}_2 = \mathbf{M}^{(2)}_1 \tag{60}$$

and

$$\boldsymbol{\Sigma}^{(1)} = \sum_{k=1}^{n_2-n_1} \mathbf{M}^{(1)+}_k \mathbf{U}^{-1}_{n_1+k-1} \mathbf{M}^{(1)}_k . \tag{61}$$

Partitioning $\mathbf{T}_{n_2}$ again into four blocks according to $\mathbf{U}_{n_2}$, one obtains

$$\mathbf{T}_{n_2} = \begin{pmatrix} \mathbf{U}_{n_2} & \mathbf{G} \\ \mathbf{G}^+ & \mathbf{R} \end{pmatrix}, \tag{62}$$

$$\mathbf{G} = (\mathbf{Q}\,\mathbf{0} \ldots \mathbf{P}_1\,\mathbf{0} \ldots \mathbf{P}_2\,\mathbf{0} \ldots \mathbf{0}), \tag{63}$$

$$\mathbf{R} = \begin{pmatrix} \mathbf{A}_{n_2+1} & \mathbf{B}_{n_2+2} & & \mathbf{0} & & & \mathbf{0} & \\ \mathbf{B}^{+}_{n_2+2} & \mathbf{A}_{n_2+2} & \mathbf{B}_{n_2+3} & & & & & \\ & \cdots & \cdots & & & & & \\ & & \cdots & \cdots & & & & \\ \mathbf{0} & & & \mathbf{A}_{n_1'} - \mathbf{\Sigma}^{(1)} & & & & \\ & & & \cdots & & & & \\ & & & & \cdots & & & \\ & & & & \mathbf{A}_{n_2'} & & \mathbf{0} & \\ & & & & & \cdots & & \\ & & & & & & \cdots & \mathbf{B}_N \\ \mathbf{0} & & & & \mathbf{0} & \mathbf{B}^{+}_N & \mathbf{A}_N & \end{pmatrix} \tag{64}$$

and

$$\mathbf{G}^{+}\mathbf{U}^{-1}_{n_2}\mathbf{G} = \begin{pmatrix} \mathbf{Q}^{+}\mathbf{U}^{-1}_{n_2}\mathbf{Q} & \mathbf{0} & \cdots & \mathbf{Q}^{+}\mathbf{U}^{-1}_{n_2}\mathbf{P}_1 & \mathbf{0} & \cdots & \mathbf{Q}^{+}\mathbf{U}^{-1}_{n_2}\mathbf{P}_2 & \mathbf{0} & \cdots \\ \mathbf{0} & & & \mathbf{0} & & & \mathbf{0} & & \\ \vdots & & & \vdots & & & \vdots & & \\ \mathbf{P}^{+}_1\mathbf{U}^{-1}_{n_2}\mathbf{Q} & \mathbf{0} & \cdots & \mathbf{P}^{+}_1\mathbf{U}^{-1}_{n_2}\mathbf{P}_1 & \mathbf{0} & \cdots & \mathbf{P}^{+}_1\mathbf{U}^{-1}_{n_2}\mathbf{P}_2 & \mathbf{0} & \cdots \\ \mathbf{0} & & & \mathbf{0} & & & \mathbf{0} & & \\ \vdots & & & \vdots & & & \vdots & & \\ \mathbf{P}^{+}_2\mathbf{U}^{-1}_{n_2}\mathbf{Q} & \mathbf{0} & \cdots & \mathbf{P}^{+}_2\mathbf{U}^{-1}_{n_2}\mathbf{P}_1 & \mathbf{0} & \cdots & \mathbf{P}^{+}_2\mathbf{U}^{-1}_{n_2}\mathbf{P}_2 & \mathbf{0} & \cdots \\ \mathbf{0} & & & \mathbf{0} & & & \mathbf{0} & & \\ \vdots & & & \vdots & & & \vdots & & \end{pmatrix}. \tag{65}$$

From $\mathbf{T}_{n_2+1} = \mathbf{R} - \mathbf{G}^{+}\mathbf{U}^{-1}_{n_2}\mathbf{G}$, together with the formulae above, it can be shown that

$\mathbf{T}_{n_2+1} =$

$$\begin{pmatrix} \mathbf{U}_{n_2+1} & \mathbf{B}_{n_2+2} & \mathbf{0} & \cdot & (-1)^{n_2-n_1+1}\mathbf{M}^{(1)}_{k_1} & \mathbf{0} & \cdot & -\mathbf{M}^{(2)}_2 & \mathbf{0} & \cdot & \cdot \\ \mathbf{B}^{+}_{n_2+2} & \mathbf{A}_{n_2+2} & \mathbf{B}_{n_2+3} & & \mathbf{0} & & & \mathbf{0} & & & \\ \mathbf{0} & \cdots & \cdots & \cdots & & & & & & & \\ \cdot & & \cdots & \cdots & \cdots & & & & & & \\ (-1)^{n_2-n_1+1}\mathbf{M}^{(1)+}_{k_1} & \mathbf{0} & & & \mathbf{A}_{n_1'} - \mathbf{\Sigma}^{(1)}_{n_2-n_1+1} & & & -\mathbf{\Sigma}^{(12)}_1 & & & \\ \mathbf{0} & & & & \cdots & \cdots & & & & & \\ \cdot & & & & & \cdots & & & & & \\ -\mathbf{M}^{(2)+}_2 & \mathbf{0} & & & -\mathbf{\Sigma}^{(12)+}_1 & & & \mathbf{A}_{n_2'} - \mathbf{\Sigma}^{(2)}_1 & & & \mathbf{0} \\ \mathbf{0} & & & & & & & \cdots & \cdots & & \\ \cdot & & & & & & & & \cdots & \cdots & \mathbf{B}_N \\ \cdot & & & & & & & \mathbf{0} & & \mathbf{B}^{+}_N & \mathbf{A}_N \end{pmatrix}, \tag{66}$$

in which $k_1 = n_2 - n_1 + 2$, and

$$\mathbf{\Sigma}_m^{(i)} = \sum_{k=1}^{m} \mathbf{M}_k^{(i)+} \mathbf{U}_{n_i+k-1}^{-1} \mathbf{M}_k^{(i)}, \quad i = 1,2; \tag{67}$$

$$\mathbf{\Sigma}_m^{(12)} = (-1)^{n_2-n_1} \sum_{k=1}^{m} \mathbf{M}_{k+n_2-n_1}^{(1)+} \mathbf{U}_{n_2+k-1}^{-1} \mathbf{M}_k^{(2)} \tag{68}$$

and the matrix $\mathbf{\Sigma}_1^{(12)}$ is just at the position where the matrix $\mathbf{M}_{n_1'-n_2+1}^{(2)}$ will be obtained later.

In the next four-block partitioning the matrix $\mathbf{T}_{n_2+2}$ is obtained as follows:

$\mathbf{T}_{n_2+2} =$

$$\begin{pmatrix}
\mathbf{U}_{n_2+2} & \mathbf{B}_{n_2+3} & \mathbf{0} & \cdots & (-1)^{n_2-n_1+2}\mathbf{M}_{k_1+1}^{(1)} & \mathbf{0} & \cdots & \mathbf{M}_3^{(2)} & \mathbf{0} & \cdots & \cdots \\
\mathbf{B}_{n_2+3}^{+} & \mathbf{A}_{n_2+3} & \mathbf{B}_{n_2+4} & & \mathbf{0} & & & \mathbf{0} & & & \\
\mathbf{0} & \cdots & \cdots & \cdots & & & & & & & \\
\vdots & & \cdots & \cdots & \cdots & & & & & & \\
(-1)^{n_2-n_1+2}\mathbf{M}_{k_1+1}^{(1)+} & \mathbf{0} & & & \mathbf{A}_{n_1'} - \mathbf{\Sigma}_{n_2-n_1+2}^{(1)} & & & -\mathbf{\Sigma}_2^{(12)} & & & \\
\mathbf{0} & & & & \cdots & \cdots & & & & & \\
\vdots & & & & & \cdots & & & & & \\
\mathbf{M}_3^{(2)+} & & \mathbf{0} & & -\mathbf{\Sigma}_2^{(12)+} & & & \mathbf{A}_{n_2'} - \mathbf{\Sigma}_2^{(2)} & & & \mathbf{0} \\
\mathbf{0} & & & & & & & \cdots & \cdots & & \\
\vdots & & & & & & & & \cdots & \cdots & \mathbf{B}_N \\
\vdots & & & & & & & \mathbf{0} & & \mathbf{B}_N^{+} & \mathbf{A}_N
\end{pmatrix}. \tag{69}$$

This procedure will be continued until the matrix $\mathbf{T}_{n_1'-1}$ is obtained,

$\mathbf{T}_{n_1'-1} =$

$$\begin{pmatrix}
\mathbf{U}_{n_1'-1} & \mathbf{B}_{n_1'} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)} & \mathbf{0} & \cdots & (-1)^{n_1'-n_2-1}\mathbf{M}_{n_1'-n_2}^{(2)} & \mathbf{0} & \cdots & \cdots \\
\mathbf{B}_{n_1'}^{+} + (-1)^{K_1}\mathbf{M}_{K_1+1}^{(1)+} & \mathbf{A}_{n_1'} - \mathbf{\Sigma}_{K_1}^{(1)} & \mathbf{B}_{n_1'+1} & & -\mathbf{\Sigma}_{n_1'-n_2-1}^{(12)} & & & \\
\mathbf{0} & & & & & & & \\
\vdots & & & & & & & \\
(-1)^{n_1'-n_2-1}\mathbf{M}_{n_1'-n_2}^{(2)+} & -\mathbf{\Sigma}_{n_1'-n_2-1}^{(12)+} & & & \mathbf{A}_{n_2'} - \mathbf{\Sigma}_{n_1'-n_2-1}^{(2)} & & & \mathbf{0} \\
\mathbf{0} & & & & & & & \mathbf{B}_N \\
\vdots & & & & & & & \\
\vdots & & & & \mathbf{0} & & \mathbf{B}_N^{+} & \mathbf{A}_N
\end{pmatrix}. \tag{70}$$

Finally partitioning again this matrix into the four-block form one obtains for the matrix $\mathbf{T}_{n_1'}$,

$$\mathbf{T}_{n_1'} =$$

$$\begin{pmatrix}
\mathbf{U}_{n_1'} & \mathbf{B}_{n_1'+1} & \mathbf{0} & \cdots & (-1)^{n_1'-n_2}\mathbf{M}^{(2)}_{n_1'-n_2+1} & \mathbf{0} & \cdots & \cdots \\
\mathbf{B}^{+}_{n_1'+1} & \mathbf{A}_{n_1'+1} & \mathbf{B}_{n_1'+2} & \ddots & \mathbf{0} & & & \\
\mathbf{0} & \ddots & \ddots & \ddots & & & & \\
\vdots & & & & & & & \\
(-1)^{n_1'-n_2}\mathbf{M}^{(2)+}_{n_1'-n_2+1} & \mathbf{0} & & \mathbf{A}_{n_2'} - \boldsymbol{\Sigma}^{(2)}_{n_1'-n_2} & & \mathbf{0} & & \\
\mathbf{0} & & & & \ddots & & \mathbf{B}_N & \\
\vdots & & & & & \ddots & \ddots & \\
\vdots & & & \mathbf{0} & & \mathbf{B}^{+}_N & \mathbf{A}_N &
\end{pmatrix} . \tag{71}$$

and the matrix $\mathbf{M}^{(2)}_{n_1'-n_2+1}$, which is different to the others, will become

$$\mathbf{M}^{(2)}_{n_1'-n_2+1} = [\mathbf{B}_{n_1'} + (-1)^{K_1}\mathbf{M}^{(1)+}_{K_1+1}]\mathbf{U}^{-1}_{n_1'-1}\mathbf{M}^{(2)}_{n_1'-n_2} + (-1)^{n_1+n_2+1}\sum_{k=1}^{n_1'-n_2-1}\mathbf{M}^{(1)+}_{k+n_2-n_1}\mathbf{U}^{-1}_{n_2+k-1}\mathbf{M}^{(2)}_{k} . \tag{72}$$

By analogy, it is valid that

$$\mathrm{NE}(x) = \mathrm{NNE}(\mathbf{U}_1) + \mathrm{NNE}(\mathbf{U}_2) + \ldots + \mathrm{NNE}(\mathbf{U}_{n_1'-1}) + \mathrm{NNE}(\mathbf{T}_{n_1'}) . \tag{73}$$

The next four-block partitionings are the same as those in step 1 and thus theorem 2 can be proved to be valid in this case also.

It should be noted that only the representation of the $\mathbf{M}^{(2)}_{n_1'-n_2+1}$ is changed as compared to all the matrix sequences discussed before.

(2.c) $n_2 < n_1 < n_1' < n_2'$. This case is corresponding to the molecular system in fig. 1(d).

It is found that the four-block partitions are nearly the same as those in the case of (2.b). The only difference is the form of the special matrix $\mathbf{M}^{(2)}_{n_1'-n_2+1}$ in the matrix sequence of $\mathbf{M}^{(2)}_{k}$:

$$\mathbf{M}^{(2)}_{n_1'-n_2+1} = [\mathbf{B}_{n_1'} + (-1)^{K_1}\mathbf{M}^{(1)+}_{K_1+1}]\mathbf{U}^{-1}_{n_1'-1}\mathbf{M}^{(2)}_{n_1'-n_2} + (-1)^{n_1+n_2+1}\sum_{k=1}^{n_1'-n_1-1}\mathbf{M}^{(1)+}_{k}\mathbf{U}^{-1}_{n_1+k-1}\mathbf{M}^{(2)}_{k+n_1-n_2} . \tag{74}$$

Formulae (72) and (74) can be combined in a unified form because the only difference between the cases (2.b) and (2.c) is that the orders of n_1 and n_2 are different. The upper limit of the summation can be replaced by $n_1' - \max(n_1, n_2) - 1$ for both cases. This is valid also for the subscript of $\mathbf{U}^{-1}$ in the summation. The subscript of the matrix $\mathbf{M}^{(1)+}$ is replaced by $k + \max(n_2 - n_1, 0)$. It is $k + n_2 - n_1$ when $n_2 > n_1$ and k when $n_2 < n_1$. The subscript of $\mathbf{M}^{(2)}$ is replaced by $k + \max(n_1 - n_2,$

0). It is k when $n_2 > n_1$ and $k + n_1 - n_2$ when $n_2 < n_1$. The unified expression of (72) and (74) then becomes

$$\mathbf{M}^{(2)}_{n'_1-n_2+1} = [\mathbf{B}_{n'_1} + (-1)^{K_1}\mathbf{M}^{(1)+}_{K_1+1}]\mathbf{U}^{-1}_{n'_1-1}\mathbf{M}^{(2)}_{n'_1-n_2} + (-1)^{n_1+n_2+1} \times \sum_{k=1}^{n'_1-\max(n_1,n_2)-1} \mathbf{M}^{(1)+}_{k+\max(n_2-n_1,0)}\mathbf{U}^{-1}_{\max(n_1,n_2)+k-1}\mathbf{M}^{(2)}_{k+\max(n_1-n_2,0)} . \tag{75}$$

To summarize, the method used to prove theorem 2 in the above case is to continue the four-block partitioning and apply theorem 1 in every step of partitioning. If there are any cross-links between different positions in a chain or between chains, a matrix sequence of $\mathbf{M}^{(i)}_k$ $(i = 1, 2, k = 1, \ldots, K_i)$ should be calculated. If and only if $n_2 < n'_1$, the cross product between the different sequences of $\mathbf{M}^{(i)}_k$ occurs and the form of $\mathbf{M}^{(2)}_{n'_1-n_2+1}$ in the matrix sequence $\mathbf{M}^{(2)}_k$ is changed while the others remain the same. The same method can be used to prove the theorem in the general case of $i = 1, \ldots, L$.

(iii) Step 3

There are L cross link blocks in eqs. (2)–(4), and their indices are $n_1, n'_1, \ldots, n_i, n'_i, \ldots, n_L, n'_L$. This is the most general case of the eqs. (2)–(4). From the above given proof it is known that there are altogether L matrix sequences of $\mathbf{M}^{(i)}_k$, $i = 1, \ldots, L$, which satisfy the recursion formula (15) and are derived with the aid of the four block partitions. For every $n_i, i = 1, \ldots, L$, take an n'_j from $j = 1, \ldots, L$. If $i > j$ and $n_i < n'_j$, then the cross product between the sequences $M^{(i)}$ and $\mathbf{M}^{(j)}$ occurs and produces the sum in the form of the matrix $\mathbf{M}^{(i)}_{n'_j-n_i+1}$ to change into

$$\mathbf{M}^{(i)}_{n'_j-n_i+1} = [\mathbf{H}^{+}_{n'_j-1,n'_j} - x\mathbf{S}^{+}_{n'_j-1,n'_j} + (-1)^{K_j}\mathbf{M}^{(j)+}_{K_j+1}]\mathbf{U}^{-1}_{n'_j-1}\mathbf{M}^{(i)}_{n'_j-n_i} + (-1)^{n_j+n_i+1} \times \sum_{k=1}^{n'_j-\max(n_i,n_j)-1} \mathbf{M}^{(j)+}_{k+\max(n_i-n_j,0)}\mathbf{U}^{-1}_{\max(n_i,n_j)+k-1}\mathbf{M}^{(i)}_{k+\max(n_j-n_i,0)} . \tag{76}$$

In all cases the forms of the matrix sequence $\mathbf{U}_n$ including $\mathbf{U}_{n'_i}$ are unchanged. In every four-block partition theorem 1 is used and it is valid that

$$\mathrm{NE}(x) = \sum_{n=1}^{N} \mathrm{NNE}[\mathbf{U}_n(x)] . \tag{77}$$

The proof of theorem 2 is completed after i, j are taken over all cross-link blocks.

Using theorem 2, eqs. (2)–(4) can be solved by the bisection method to obtain the eigenvalues. The eigenvectors can be obtained by the inverse iteration [18] after the eigenvalues have been obtained. The approximate density of states $\bar{\rho}(E)$ can be calculated as

$$\bar{\rho}(E) = \frac{\eta(E + \Delta E) - \eta(E)}{\Delta E}, \tag{78}$$

where $\eta(E)$ is the number of eigenvalues of eqs. (2)–(4) in the interval $(-\infty, E]$.

4. Discussion

The application of the ENFC theorem has now made it possible for us to solve the eigenvalue equation derived for a native protein in which cross-links exist with chemical bonds or van der Waals forces. The advantages of this method are that only non-zero matrix elements have to be treated. Pig insulin, one of the smallest native proteins in which there are two chains, three disulfur bonds (cross-links with chemical bonds), 51 amino acid residues and 782 atoms, has been taken as the first example of this method in an ab initio scheme. In the calculation 2418 basis functions were used applying a minimal basis set. The bisect width of the eigenvalue was taken as 10^{-6} and $\delta = \{[(\mathbf{H} - E\mathbf{S})C]^{+}[(\mathbf{H} - E\mathbf{S})C]\}^{1/2} \leqslant 2 \times 10^{-6}$ was the error test of the eigenvectors. The electronic DOS of insulin confirmed the conclusions obtained previously by model proteins. Further, the frontier orbitals showed that there are some intrinsic relationships between the electronic structures and the biological activities of insulin. The details of the results are reported in the subsequent paper [17].

Acknowledgement

The author would like to express his gratitude to Professor J. Ladik at the University of Erlangen-Nürnberg and to Professor Lemin Li and Professor Guan-Xian Xu at Peking University for their encouragement and helpful discussions.

References

[1] R. Srinwasan, N. Yathindra and E. Subramanian, *Biomolecular Structure, Comformation, Function, and Evolution*, Vol. 2, Physico-Chemical and Theoretical Studies (Pergamon Press, Oxford, 1981).
[2] D.L. Oxender and C.F. Fox, *Protein Engineering* (Liss, New York, 1987);
A. Fersht, *Enzyme Structure and Mechanism*, 2nd Ed. (Freeman, New York, 1985).
[3] P.-O. Lödwin and B. Pullman, *New Horizons of Quantum Chemistry* (Reidel, Dordrecht, 1983).
[4] P. Dean and J.L. Martin, Proc. Roy. Soc. A259 (1960) 409.
[5] P. Dean, Rev. Mod. Phys. 44 (1972) 127.
[6] R.S. Day and F. Martino, Chem. Phys. Lett. 84 (1981) 86.
[7] J. Ladik and J.-M. André, *Quantum Chemistry of Polymers, Solid State Aspects* (Reidel, Dordrecht, 1983).
[8] A.K. Bakhshi and J. Ladik, Chem. Phys. Lett. 129 (1986) 269.
[9] J. Ladik, M. Seel, P. Otto and A.K. Bakhshi, Chem. Phys. 108 (1986) 203.

[10] A.K. Bakhshi, P. Otto, J. Ladik and M. Seel, Chem. Phys. 108 (1986) 215.
[11] P. Otto, A.K. Bakhshi, J. Ladik, M. Seel and S. Chin, Chem. Phys. 108 (1986) 223.
[12] A.K. Bakhshi, J. Ladik, M. Seel and P. Otto, Chem. Phys. 108 (1986) 233.
[13] J. Ladik, P. Otto, A.K. Bakhshi and M. Seel, Int. J. Quant. Chem. 29 (1986) 597.
[14] A.K. Bakhshi, P. Otto, C.-M. Liegener, E. Rehm and J. Ladik, Int. J. Quant. Chem. 38 (1990) 573.
[15] C.-M. Liegener, A. Sutjianto and J. Ladik, Chem. Phys. 145 (1990) 385.
[16] J.-B. Li, Y.-J. Ye, A.-Q. Tang, Chinese Sci. Bull. 34 (1989) 1882.
[17] Y.-J. Ye and J. Ladik, J. Math. Chem. 14 (1993) 141.
[18] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1965).