An Alternative Strategy for Count and Storage of Kekulé and Longer Range Resonance Valence Bond Structures

Fei Cai, Heqing Shao, Chungen Liu,* and Yuansheng Jiang

Institute of Theoretical and Computational Chemistry, Key Laboratory of Mesoscopic Chemistry, Department of Chemistry, Nanjing University, Nanjing, 210093, China

In this paper, we suggest a simple representation for notation of Kekulé valence bond (KVB) structures and longer range resonance valence bond (RVB) structures, which is called "the adjacency bonding array". In this representation, only an *N* component one-dimensional array is needed for inscribing each KVB or longer range RVB structure for an *N*-carbon system. Based on the adjacency bonding arrays, we develop very efficient algorithms for the systematic search of KVB and RVB structures as well as evaluation of the basis set overlap and Hamiltonian matrices.

1. INTRODUCTION

For a long time, enumeration of Kekulé valence structures has been one of the main focuses in graph theoretical study of conjugated hydrocarbons. Aside from the systematic search method of Kekulé valence structures suggested by Pauling,[1] people have also found analytical formulas for specific kinds of systems[2−7] or derived recursion relations between smaller fragments and larger systems.[8−11] There were also particular techniques developed to work on this problem,[12−25] which made it possible to enumerate Kekulé valence structures for systems with hundreds of atoms, such as nanotubes. Readers may refer to several review articles for the detailed description of the enumeration techniques.[26−28]

Despite the impressive achievements they have made, topological indices[28] built on the basis of the Kekulé valence structure count have unreasonably implied the equal contributions from different Kekulé valence structures.[29] Such differences were partly taken into considerations in Clar's aromatic sextet[30] and Randić's conjugated circuits[31] theories, while the full consideration of it is reflected in the Kekulé valence structure based Valence Bond theory (KVB)[29,32−34] performed on the space spanned by the Kekulé structures. In this case, the wave function as well as the resonance energy of the system are expressed as a summation over all the Kekulé structures with different weights.

For the calculations on the space spanned by the Kekulé valence structure, it is necessary to develop a representation for inscribing and coding the Kekulé structures. The most trivial representation that can be used in computer programming might be the slightly modified adjacency matrix, with the double bonds denoted by number "2" instead of "1". Another representation is much simplified by recording only the adjacency matrix (which is the same for all the Kekulé structures) and the π electron pairs (double bonds).[35] The above two representations record all necessary information for the recovery of the topological structures of Kekulé valence structures. Recently, Randić[36−41] suggested an elegant algebraic representation for Kekulé structures based on the

count of π-electrons associated with individual rings. It has been proven to have a one-to-one correspondence with the Kekulé structure for cata-condensed benzenoids,[42] while for pericondensed benzenoids, this relationship is violated,[36−43] which introduces some inconvenience in computer programming.

In this paper we suggest an alternative scheme, which is named "the adjacency bonding array", for the search and storage of Kekulé or other resonance valence bond (RVB) structures. Actually, this representation is based on an idea similar to that suggested by Klein et al.,[35] though we will focus on developing an economical storage scheme that could as well be highly efficient for a systematic searching of KVB or longer range RVB structures. In our representation, only an $N \times 1$ array is needed for inscribing each KVB or longer range RVB structure for an *N*-carbon system. Based on the adjacency bonding array, we develop fast algorithms for searching KVB and RVB structures as well as for computing the elements of overlap and Hamiltonian matrices.

The paper is organized as follows. A brief introduction to KVB and RVB theories is shown in section 2. Section 3 summarizes miscellaneous "geometrical representations" (using this term in contrast to the algebraic representation) and introduces "the adjacency bonding array". Next, in section 4, we presented the universal algorithms for the search and storage of the Kekulé, S_3 and $D_{3,3}$ structures and discussed calculation of KVB and RVB basis set overlap matrix elements as well as Hamiltonian matrix elements. Finally, in section 5, the efficiency of our programs is discussed based on calculations on 72 examples of moderately sized conjugated hydrocarbons, among which the most complicated one has as many as 132 carbons and more than 10^8 Kekulé structures.

2. KVB AND RVB THEORY

The traditionally used isotropic VB Heisenberg Hamiltonian for a conjugated π-network reads

$$\hat{H} = J\sum_{i-j}\left(2S_i - S_j - \frac{1}{2}\right) \qquad (1)$$

* Corresponding author phone: +86-25-83596756; e-mail: cgliu@netra.nju.edu.cn.

where S_i denotes the total spin-1/2 operator for the site i, the symbol $i{-}j$ indicates summation over adjacent atom pairs and J denotes a positive exchange parameter, which is simply treated as a multiplicative constant in KVB and RVB models.

Since the configuration space goes up exponentially with the increase of number of atoms, an exact solution of this Hamiltonian is practical at this time only for systems consisting of 30 or fewer atoms. Among the various approximation techniques developed, a chemically reasonable approach is to solve the Hamiltonian in the subspace spanned by the Kekulé structures, while the ground-state wave function could be written as a linear combination of the Kekulé structures[29,32,34]

$$\Psi = \sum_i C_i K_i \tag{2}$$

The coefficients and eigenvalues can be evaluated by solving eq 3

$$(\boldsymbol{H} - \boldsymbol{ES})\boldsymbol{C} = 0 \tag{3}$$

where $\boldsymbol{H}$ is the Hamiltonian matrix and $\boldsymbol{S}$ is the overlap matrix.

It was found that this approach gives a reasonable approximate description of the exact ground state.[29,44] To get an even better approximation to the exact VB solution, multiply excited RVB bases should be included besides the Kekulé basis set.[34,44] In these bases, each overall singlet function ϕ is written as a product of individual singlet spin pairs (p,q), not necessarily adjacent atom pairs

$$\phi = \prod_{p,q}^{N/2} \frac{1}{\sqrt{2}}(\alpha_p\beta_q - \beta_p\alpha_q) \tag{4}$$

This is known as the general form of a singlet RVB basis set, with the KVB being a special case where all the (p,q) pairs are adjacent atoms. In a previous paper,[34] the longer range RVB functions were classified into different groups according to their degree of excitation e depending on how many singlet pairs in ϕ are between the nonadjacent atom pairs. Furthermore, within each set ϕ^e of equally excited RVB functions we can label each $\phi \in \phi^e$ by the bond distances d_1, d_2..., d_e associated with the first, second, ..., eth excited singlet pair in ϕ^e. Hence we can formally decompose our RVB basis in the following way

$$\phi = \oplus_{e=0}^{max} \oplus_{d=e\times2}^{e\times d^{max}}\{\phi_d^{\ e}\} \tag{5}$$

where $e \times 2$ and $e \times d^{max}$ denotes e excited singlet pairs of distance 2 and d^{max}, respectively. Hence, one may designate the RVB subspaces corresponding to $e=0$, 1 and 2 by the special symbols K, S and D, respectively, writing d values as suffixes, of which the nonexcited K-space is identical to the space spanned by all Kekulé functions and the most important excited subspaces are S_3, $D_{3,3}$.[34] The Kekulé and S_3, $D_{3,3}$ structures are shown in Figure 1.

3. TOPOLOGICAL INDEXES FOR CONJUGATED HYDROCARBONS

3.1. Adjacency Matrix. The adjacency matrix is widely utilized in chemical graph theory. The simplest form of the
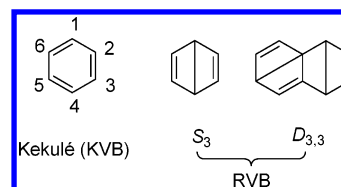
Figure 1. Kekulé, S_3 and $D_{3,3}$ structures.

adjacency matrix A is just an atom connectivity table in matrix form. Thus, $A_{ij} = A_{ji} = 1$ if atoms i and j are directly bonded; all other elements are zero. For a Kekulé structure, $A_{ij} = A_{ji} = 2$ if atoms i and j are doubly bonded.

3.2. Extended Adjacency Matrix. An extended adjacency matrix is built as a *generalized* adjacency matrix by adding an extra column and row to the original adjacency matrix, whose elements have slightly different definitions for the notation of the double bond positions.

It is known that each carbon atom in conjugated hydrocarbons will connect to three carbon atoms, two carbon and one hydrogen atoms, or one carbon and two hydrogen atoms. According to the definition of the adjacency matrix, for the three cases mentioned above the summation of the correspondent column or row in the adjacency matrix will be 3, 2, or 1 respectively, while for Kekulé structures, they will be 4, 3, or 2, respectively.

For the sake of simplicity in computer programming, in the extended adjacency matrix representation, the matrix is extended to rank $(N + 1) \times (N + 1)$, so as to treat these three cases uniformly. The elements in the added column and row are either set to 1 (or 2) when there is 1 (or 2) hydrogen atom connections to the corresponding atom, and 0, otherwise. For instance, the respective extended adjacency matrices for the skeleton and the Kekulé structure of benzene (see Figure 1) are

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} j/i & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 1 & 0 & 0 & 0 & 2 & 1 \\ 2 & 1 & 0 & 2 & 0 & 0 & 0 & 1 \\ 3 & 0 & 2 & 0 & 1 & 0 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 & 2 & 0 & 1 \\ 5 & 0 & 0 & 0 & 2 & 0 & 1 & 1 \\ 6 & 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ 7 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

3.3. Bond List. Since the adjacency matrix is a sparse matrix, it is not economical in computer programming. The technique of the *so-called* "bond list", which records the position and value of the nonzero elements in the lower triangle of the extended adjacency matrix, can be employed to save memory. Each member in the list reads (i, j, A_{ij}), with i, j referring to the position, and A_{ij}, the corresponding value of the matrix element. The bond list (columnwise) for

the above-mentioned Kekulé structure is as follows:

$$(1,2,1), (1,6,2), (1,7,1), (2,3,2), (2,7,1), (3,4,1), (3,7,1),$$
$$(4,5,2), (4,7,1), (5,6,1), (5,7,1), (6,7,1)$$
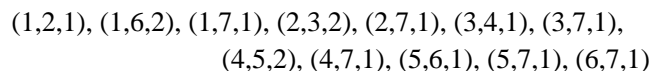
In this representation, the memory cost for the notation of one Kekulé structure is cut down from $(N + 1)^2$ to $6N$ or less, where N is the number of carbon atoms in the system.

3.4. Adjacency Bonding Array. Further improvements could be achieved by the introduction of the adjacency bonding array. The adjacency bonding array is based on the concept of quadrivalence of carbon. For a conjugated hydrocarbon, the skeleton (H atoms included) provides three valences while a Kekulé or RVB structure supplies the fourth. Hence for an N-carbon system, a two-dimensional array $a[N]$-$[3]$ (in C Language) is employed to store the atom connectivity of the skeleton, which we refer to as the skeleton adjacency bonding array: $a[i][k]$, $k = 0,1,2$ records adjacent atoms (0 for a H atom) of the ith carbon atom (Ci). An additional array $K[N]$, which is referred to as a KVB or RVB adjacency bonding array, is required for notation of each KVB or RVB structure. $K[i]$ records the atom paired with Ci in the Kekulé or RVB structure. Accordingly, the above-mentioned KVB structure of benzene is denoted by

$$a[6][3] = \begin{vmatrix} 0 & 2 & 6 \\ 0 & 1 & 3 \\ 0 & 2 & 4 \\ 0 & 3 & 5 \\ 0 & 3 & 5 \\ 0 & 4 & 6 \\ 0 & 1 & 5 \end{vmatrix} \text{ combined with } K[6] = \begin{vmatrix} 6 \\ 3 \\ 2 \\ 5 \\ 4 \\ 1 \end{vmatrix}$$

The arrays $a[N][3]$ and $K[N]$ are sometimes called the connection matrices in the form of eq 6 and eq 7.

$$C(i, c) \equiv$$
$$\begin{cases} i_c \equiv c^{th} \text{ site to which site } i \text{ is attached, } c \leq d \\ 0 \qquad \text{when } c < d_i \leq 3 \end{cases}, c = 1,2,3$$
$$(6)$$

$$K(i) \equiv j \text{ for } \{i, j\} \in \text{Kekulé structure under consideration}$$
$$(7)$$

Since the skeleton adjacency bonding array is the same for all the KVB and RVB structures in one molecule, one only needs to save a one-dimensional array $K[N]$ for each structure.

4. ALGORITHM

The program dynamically allocates a number of one-dimensional arrays for storing KVB or RVB structures and declares a pointer array $p[Q]$ for storing the start addresses of these arrays (Q slightly exceeds K considering the temporary storage of unreasonable structures (adjacency bonding arrays) generated in the process). $p[m][i]$ (i.e., $*p[m] + i$ in C language) records the sequence number of the atom paired with Ci in structure m. The generated adjacency bonding arrays would be handled more flexibly with the pointer array.

4.1. Kekulé Structure Count and Storage. The steps for Kekulé structure count and storage are illustrated in Figure 2 and Table 1 with the simple example of naphthalene. The
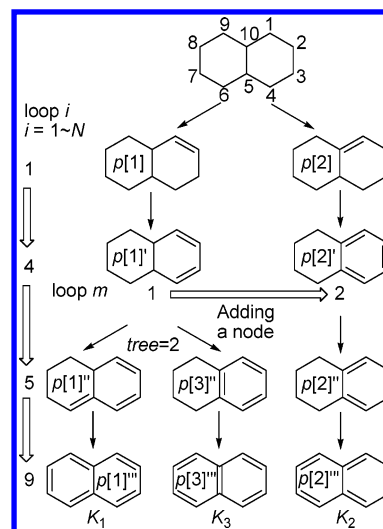
Figure 2. Illustration of the procedure of Kekulé structure count and storage.

algorithm could be arranged in a very simple nested loop structure. As illustrated in Figure 2, loop i is the outer loop for treating all the carbon atoms one by one from 1 to N; and loop m is the inner loop for dealing with each KVB structure candidate generated in the last cycle of loop i. This algorithm is similar to the original graph theoretical algorithm[45] and is feasible to be implemented using computer languages.

As for the two loops, the outer loop is not difficult to understand; in the following, we would like to give a brief introduction to the loop m. In the ith cycle of loop i, "Saturating" the ith atom is done for each KVB structure candidate (denoted by the variable m) searched out in the last cycle ($i-1$ th) of loop i. The so-called "Saturating" is to form a double bond between Ci and one of its neighbors, so that the four valences of Ci are used up. For example, in the first step, saturation of C1 is to form a double bond between C1 and, either C2 or C10. Accordingly, two KVB structure candidates are generated and recorded in two adjacency bonding arrays to which $p[1]$ and $p[2]$ point, respectively.

Besides "Saturation", another two operations need to be mentioned. One is the "Deleting" operation. In some cases, we may arrive at an unsaturated atom, all of whose neighbors are already saturated, which indicates that this structure is not a reasonable candidate of KVB structure. We should delete it from the lists. The other operation is "Appending". When the atom in consideration has only one unsaturated neighboring atom ahead, there is only one choice to saturate it, see $p[1]$ to $p[1]'$ or $p[2]$ to $p[2]'$. However, we may often face the other case when there is more than one unsaturated neighbor. In this case, more than one KVB candidate is generated from the original candidate, of which we can assign one to the original position of the structure list, i.e., $p[1]''$, while appending the others to the end of the list, e.g., p[3]$''$.
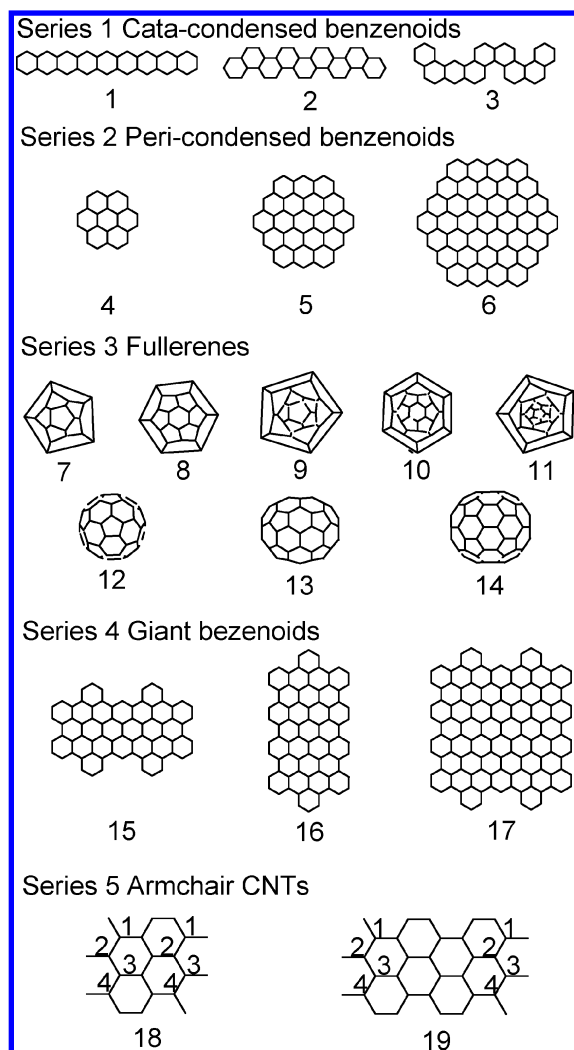
To check whether an atom is saturated seems to be a trivial problem. However, in the extended adjacency matrix representation, it is checked by the following discriminant

$$a(i,1,m) + a(i,2,m) + \cdots + a(i,i-1,m) + (a(i,i,m) + a$$
$$(i + 1,i,m) + \cdots a(n + 1,i,m) == 4$$

which involves N operations on discontinuously distributed

Table 1. Generation of Kekulé Structures for Naphthalene

i	$a[i][0]$	$a[i][1]$	$a[i][2]$	$p[1][i]$	$p[2][i]$	$p[1]'[i]$	$p[2]'[i]$	$p[1]''[i]$	$p[2]''[i]$	$p[3]''[i]$	$K_1[i]$	$K_2[i]$	$K_3[i]$
1	0	2	10	2	10	2	10	2	10	2	2	10	2
2	0	1	3	1	0	1	3	1	3	1	1	3	1
3	0	2	4	0	0	4	2	4	2	4	4	2	4
4	0	3	5	0	0	3	5	3	5	3	3	5	3
5	4	6	10	0	0	0	4	6	4	10	6	4	10
6	0	5	7	0	0	0	0	5	0	0	5	7	7
7	0	6	8	0	0	0	0	0	0	0	8	6	6
8	0	7	9	0	0	0	0	0	0	0	7	9	9
9	0	8	10	0	0	0	0	0	0	0	10	8	8
10	1	5	9	0	1	0	1	0	1	5	9	1	5

Figure 3. Representative conjugated systems.

data in memory. In the bond list representation, this is not effectively optimized; one still has to check all the bond lists, while in the adjacency bonding array representation the discriminant is simplified to check whether $p[m][i]$ is assigned a nonzero value (referring to paired atom).

For illustration we show in Figure 3, nineteen representative conjugated systems grouped in 5 classes, which include cata-condensed benzenoids, peri-condensed benzenoids, fullerenes, giant benzenoids and armchair carbon nanotubes (CNT). The KVB structure numbers are presented in Table 2. All of the results are in accord with the reported values.[2,18,25,34,46,47]

4.2. RVB (S_3 and $D_{3,3}$) Structure Count and Storage. According to the definition of S_3 and $D_{3,3}$, the S_3 algorithm

is to search all of the $d = 3$ pairs in the skeleton, and the $D_{3,3}$ algorithm is to find all combinations of two $d = 3$ pairs, followed by the determination of K for each residual fragment. Actually, searching of the $d = 3$ pairs is trivial by inspection of the skeleton adjacency bonding array for the specific molecule.

Combining the most important subspaces S_3, $D_{3,3}$ with the K-space, we designed the $K + S_3 + D_{3,3}$ algorithm. Calculations are performed on 72 selected conjugated systems including carbon cages and planar hydrocarbons, as shown in Figure 4, and the results are presented in Table 3. Most of the structures ($K + S_3$: $C_{40}-C_{60}$ and $K + S_3 + D_{3,3}$: $C_{20}-C_{40}$) have been studied in the literature,[34,44] and the reported data are in accord with our results.

The results presented in Table 3 indicate that for most planar conjugated hydrocarbons, $K + S_3$ is around $10-20$ times as large as K and $K + S_3 + D_{3,3}$ is $5-15$ times $K + S_3$. However, these could only be considered as a minor portion of the complete VB space, furthermore, the ratio of the dimension of these subspaces against the complete space falls off exponentially with the system size. On the other hand, it has been claimed[44] that using K-space alone in a RVB calculation (KVB) one recovers typically 90% of the exact ground-state energy of eq 1, 98.5% using $K + S_3$ and 99.5% using $K + S_3 + D_{3,3}$.

4.3. Calculation of Matrix Elements. The Hamiltonian matrix and overlap matrix elements $H_{k\lambda}(\langle\phi_k|H|\phi_\lambda\rangle)$ and $S_{k\lambda}(\langle\phi_k|\phi_\lambda\rangle)$ in eq 3 between the resonance valence bond structure ϕ_k and ϕ_λ can be computed easily in terms of Pauling's diagrammatic rules[32,48,49]

$$S_{k\lambda} = (-1)^{n_v} \cdot 2^{n_i - N/2} \qquad (8)$$

$$H_{k\lambda} = JS_{k\lambda}\sum_{i-j}\left(\frac{3}{2}\delta_{i,j} - \frac{1}{2}\right) \qquad (9)$$

where n_i and n_v denote the number of closed islands and reversals needed for forming standard islands; δ_{ij} is nonzero only if i and j belongs to the same island, in which case its value is $+1$ or -1 as the distance between i and j is even or odd.

With the structures generated in the form of adjacency bonding arrays, it is convenient to design a fast algorithm to determine the number of islands, reversals and $H_{k\lambda}$. An example of Kekulé structures K_2 and K_3 (in Figure 2 and Table 1) of naphthalene is illustrated in Figure 5a and Table 4. Let us begin with $i = 1$. Putting together the two columns of adjacency bonding arrays for K_2 and K_3 in Table 4, we find that C1 is paired with C2 (in K_3) and C10 (in K_2). Then we look at the smaller indexed atom C2 and find that C2 is
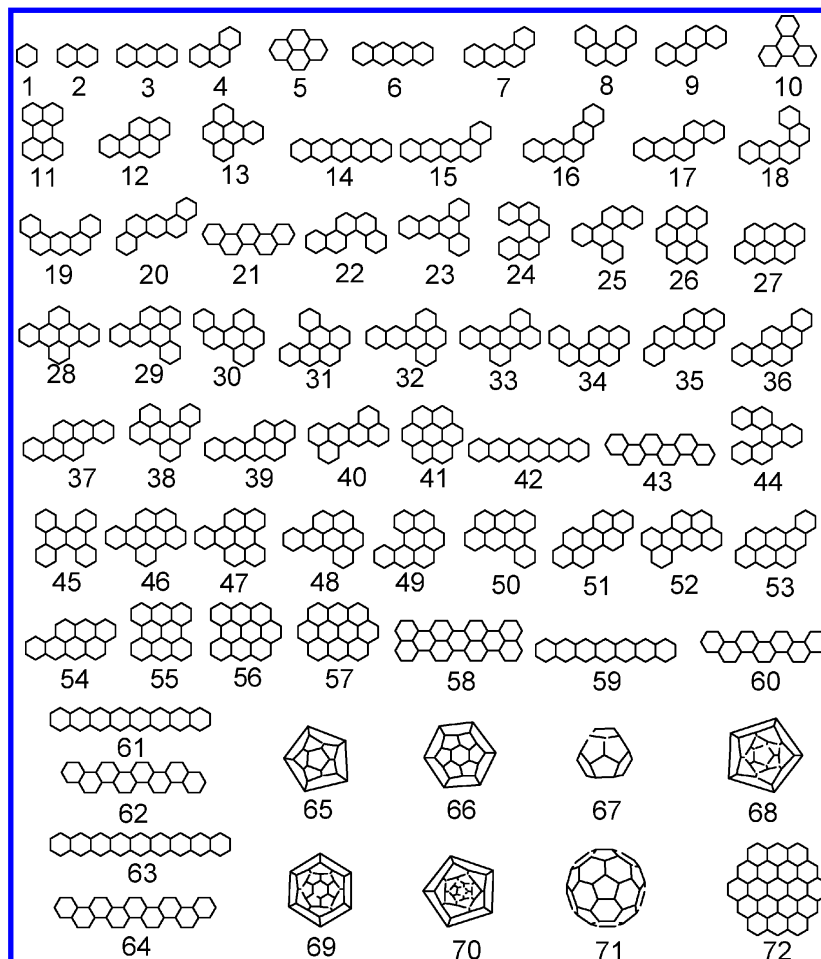
Table 2. Kekulé Structure Number of Molecules in Figure 3

compd	1	2	3	4	5	6	7	8	9	10
n	38	38	38	24	54	98	20	24	30	36
K	10	89	81	20	980	232 848	36	54	151	272

compd	11	12	13	14	15	16	17	18	19
n	40	60	70	84	72	78	132	12	24
K	701	12 500	52 168	394 944	16 100	42 500	125 820 416	13	41

Figure 4. Selected conjugated systems.

further paired with C3 in K_3. Thereafter, C3 is paired with C4 in K_3; C4 is paired with C5 in K_2. Finally, an island is formed (loop C1−C2−C3−C4−C5−C10−C1). All of the islands n_i would be found by repeating the above procedure for $i = 1 \sim$ N, skipping the atoms that belong to a previous island.

Two additional arrays *spin*[i] and *island*[i] are employed in the algorithm so as to obtain the number of reversals n_v and δ_{ij} in S_{23} and H_{23}. During the formation of an island, mark the atoms on the island with 1 and −1 alternatively (recorded in the array *spin*[i] as the notation of spin α and β, respectively—see column 5 in Table 4) to denote the spin distribution in a standard island, besides, the island sequence number that the atom belongs to is recorded in another array *island*[i]. All spin pairs (i, j) in the superposition pattern, with $i < j$, constitute all arrows $i \rightarrow j$ in the Rumer Diagram. While the direction in a standard island is from spin α to β, so a reversal is required when *spin*[i] < *spin*[j], i.e., the arrow should be $i \leftarrow j$. As mentioned above, δ_{ij} is nonzero when i and j belong to the same island. In this case, obviously, δ_{ij}

could be given by a simple expression, *spin*[i] × *spin*[j].

In fact, for bipartite graphs such as compounds 1−6 (including the aforementioned naphthalene), and 15−19 in Figure 3, or compounds 1−64 and 72 in Figure 4, the sign of $S_{k\lambda}$ is positive, $S_{k\lambda}$ being $2^{n_i-N/2}$. The above procedure might be simplified considerably. The arrows may be replaced by lines (see Figure 5b), that is, the array *spin*[i] is not necessarily. And since the distance between i and j is always even, δ_{ij} is −1 or 0 as i and j belongs to the same island or not.

In this way, the matrix elements necessary for the KVB or longer range RVB calculations could be obtained, and the only thing left is to find a high-efficient algorithm for the eigenvalue problem on the nonorthogonal basis set. Flocke et al. used a modified Davidson method.[44] Alternatively, we used Nesbet's iterative algorithm.[50] During the computations on the 72 molecules in Figure 4, we find that the convergence of eigenvalue calculation is not strongly related to the initial guess, which makes this algorithm easy to implement, while maintaining a satisfactory speed of convergence.

Table 3. K, $K + S_3$ and $K + S_3 + D_{3,3}$ Structure Numbers of Molecules in Figure 4

compd	n	K	$K + S_3$	$K + S_3 + D_{3,3}$	compd	n	K	$K + S_3$	$K + S_3 + D_{3,3}$
1	6	2	5	5	37	24	13	224	1687
2	10	3	17	39	38	24	12	209	1594
3	14	4	40	144	39	24	12	213	1668
4	14	5	39	145	40	24	9	156	1245
5	16	6	62	301	41	24	20	332	2690
6	18	5	77	406	42	26	7	205	1994
7	18	7	80	413	43	26	21	325	2420
8	18	8	84	412	44	26	22	335	2447
9	18	8	84	412	45	26	24	346	2469
10	18	9	84	411	46	26	22	405	3309
11	20	9	117	722	47	26	23	391	3224
12	20	9	121	747	48	26	22	380	3190
13	20	11	131	785	49	26	20	363	3092
14	22	6	131	759	50	26	19	333	2827
15	22	9	146	990	51	26	18	326	2794
16	22	10	147	1007	52	26	18	318	2733
17	22	11	159	1019	53	26	16	295	2606
18	22	11	159	1019	54	26	14	284	2546
19	22	12	159	1021	55	28	16	388	4030
20	22	12	159	1021	56	30	30	702	7711
21	22	13	168	1042	57	32	50	1210	14 181
22	22	13	168	1042	58	40	81	2241	30 492
23	22	13	164	1028	59	30	8	302	3768
24	22	13	168	1042	60	30	34	611	5323
25	22	14	173	1043	61	34	9	425	6613
26	22	14	204	1449	62	34	55	1127	11 233
27	22	10	162	1209	63	38	10	577	10 945
28	24	20	270	1898	64	38	89	2047	22 980
29	24	17	250	1813	65	20	36	516	3696
30	24	17	254	1838	66	24	54	1134	10 950
31	24	16	246	1794	67	26	63	1611	17 778
32	24	16	243	1812	68	30	151	3891	48 726
33	24	15	226	1678	69	36	272	9872	167 648
34	24	15	234	1742	70	40	701	25 821	461 171
35	24	15	234	1742	71	60	12 500	722 300	20 633 840
36	24	14	223	1702	72	54	980	43 190	938 270
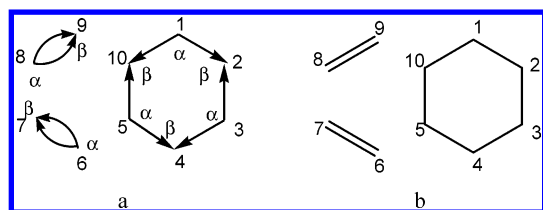
Figure 5. The superposition pattern formed by Kekulé structures K_2 and K_3 of naphthalene.

Table 4. Searching for Islands Formed by Superposition of Kekulé Structures K_2 and K_3

i	$K_2[i]$	$K_3[i]$	*island*[i]	*spin*[i]
1	10	2	1	1
2	3	1	1	-1
3	2	4	1	1
4	5	3	1	-1
5	4	10	1	1
6	7	7	2	1
7	6	6	2	-1
8	9	9	3	1
9	8	8	3	-1
10	1	5	1	-1

5. PROGRAM EFFICIENCY

In the algorithm based on adjacency bonding arrays, for an $p[m][N]$ system, type of one-byte unsigned character is enough for the dynamically allocated one-dimensional array $p[m][N]$, which saves memory dramatically. Taking C_{60} for example, only $60 \times 12500 \times 1 = 732$ K computer memory is required for storing all its 12,500 Kekulé structures.

Table 5. Time Cost (in seconds) for Computing K, $K + S_3 + D_{3,3}$ Structures of C_{60} on a SGI Origin3000 (400 MHz)

structure	number of structures	based on adjacency matrices	based on adjacency bonding arrays
K	12 500	56.7	0.1
$K + S_3$	722 300	12800.4	9.7
$K + S_3 + D_{3,3}$	20 633 840		322.4

According to this algorithm, it is found that the computational time consumption in structure counting is approximately proportional to the product of atom number N and structure count, whenever for K, $K + S_3$, or $K + S_3 + D_{3,3}$. The data listed in Table 5 indicate that, with the employment of adjacency bonding arrays, structure counting is greatly accelerated. Besides, our representation of resonant valence structures has the following benefits:

1. The discrimination of whether a carbon atom is saturated is much simplified by using a simple operation, i.e., to inspect whether $p[m][i]$ is positive or zero. A positive value refers to the sequence number of the atom, which forms a spin pair with Ci in structure m; while zero means in structure m, atom Ci is still unpaired.

2. Since the storage structure of adjacency bonding arrays does not distinguish between double bonds and $d > 1$ bonds, Kekulé, S_3 and $D_{3,3}$ structures can be treated uniformly.

3. As we show in the last section, with the adoption of the adjacency bonding arrays, it is very convenient and highly efficient to search the islands and determine the reversals as

NOTATION OF KEKULE VALENCE BOND STRUCTURES

J. Chem. Inf. Model., Vol. 45, No. 2, 2005 **377**

well as the interatomic topological distances during the computation of overlap matrix elements and Hamiltonian matrix elements.

Owning to the above-mentioned advantages of adjacency bonding arrays, as has been shown in Table 5, the program runs much faster than that employs adjacency matrices. Another property of equally importance is that our strategy is memory saving, which makes it possible to deal with systems having as many as 10^7 $K + S_3 + D_{3,3}$ RVB structures on a typical 64 bit UNIX workstation, with less than 10^3 seconds of CPU time.

6. CONCLUSION

A new representation for the atom connectivity of hydrocarbons has been suggested, which we refer to as the adjacency bonding array. For an N-carbon system, only an $N \times 1$ array is needed for the designation of each KVB or longer range RVB structure. Based on the adjacency bonding array representation, we have developed a fast algorithm for the RVB structure searching and storing. Validation of this algorithm was performed on 19 representative conjugated systems of 5 series, for which the KVB structures had been reported earlier. We also generated the KVB, S_3 and $D_{3,3}$ RVB structures for 72 selected moderately sized conjugated hydrocarbons. In addition, with the KVB or longer RVB structures stored in the form of adjacency bonding arrays, a high-efficiency computer program for the valence-bond Hamiltonian and overlap matrix elements calculation has also been developed.

REFERENCES AND NOTES

(1) Pauling, L. Bond Numbers and Bond Lengths in Tetrabenzo[de,no,st, c1de]Heptacene and Other Condensed Aromatic-Hydrocarbons − A Valence-Bond Treatment. *Acta Crystallogr.* **1980**, *B36*, 1898−1901.

(2) Randić, M.; Guo, X. Resonance Energy of Giant Benzenoid Hydrocarbon $C_{78}H_{26}$. *Int. J. Quantum Chem.* **1999**, *74*, 697−708.

(3) Gordon, M.; Davison, W. H. T. Theory of Resonance Topology of Fully Aromatic Hydrocarbon. I. *J. Chem. Phys.* **1952**, *20*, 428−435.

(4) Cyvin, S. J. The Number of Kekulé Structures for Primitive Coronoids (Cycloarenes). *Chem. Phys. Lett.* **1988**, *147*, 384−388.

(5) Gutman, I.; Klein, D. J. Generalization of Cyvin Formula. *J. Math. Chem.* **1991**, *6*, 41−50.

(6) Sachs, H.; Hansen, P.; Zheng, M. Kekulé Count in Tubular Hydrocarbons. *Match* **1996**, *33*, 169−241.

(7) Klein, D. J.; Zhu, H. Resonance in Elemental Benzenoids. *Discrete Appl. Math.* **1996**, *67*, 157−173.

(8) Lin, C.; Tang, P. Kekulé Count in Capped Zigzag Boron-Nitride Nanotubes. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 13−20.

(9) Fowler, R. H.; Rushbrooke, G. S. An Attempt to Extend the Statistical Theory of Perfect Solutions. *Trans. Faraday Soc.* **1937**, *33*, 1272−1294.

(10) Ivanciuc, O.; Balaban, A. T. Nonisomorphic Graphs with Identical Atomic Counts of Self-Returning Walks-Isocodal Graphs. *J. Math. Chem.* **1992**, *11*, 155−167.

(11) Klein, D. J.; Trinajstić, N. PASCAL Recurrence Algorithm for Kekulé Structure Counts of Benzenoid and Coronoid Hydrocarbons. *J. Mol. Struct. (THEOCHEM)* **1990**, *206*, 135−142.

(12) Platt, J. R. *In Handbuch der Physik*; Flugge, S., Ed.; Springer-Verlag: Berlin, 1961; pp 205−209.

(13) Hall, G. G. The Bond Orders of Alternant Hydrocarbon Molecules. *Proc. R. Soc. London* **1955**, *A229*, 251−259.

(14) Cash, G. G. A Fast Computer Algorithm for Finding the Permanent of Adjacency Matrices. *J. Math. Chem.* **1995**, *18*, 115−119.

(15) Cash, G. G. A Simple Means of Computing the Kekulé Structure Count for Toroidal Polyhex Fullerenes. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 58−61.

(16) Torrens, F. Computing the Kekulé Structure Count for Alternant Hydrocarbons. *Int. J. Quantum Chem.* **2002**, *88*, 392−397.

(17) Klein, D. J.; Hite, G. E.; Schmalz, T. G. Transfer-Matrix Method for Subgraph Enumeration: Applications to Polypyrene Fusenes. *J. Comput. Chem.* **1986**, *7*, 443−456.

(18) Schmalz, T. G.; Seitz, W. A.; Klein, D. J.; Hite, G. E. Elemental Carbon Cages. *J. Am. Chem. Soc.* **1988**, *110*, 1113−1127.

(19) Hite, G. E.; Metropoulos, A.; Klein, D. J.; Schmalz, T. G.; Seitz, W. A. Extended π-Networks with Multiply Spin-Pairing Phases: Resonance Theory Calculations on Poly-Poly Phenanthrenes. *Theor. Chim. Acta* **1986**, *69*, 369−391.

(20) Randić, M.; Klein, D. J.; Zhu, H.-Y.; Trinajstić, N.; Živković, T. Comparative-Study of Large Molecules-Highly Accurate Calculation of A Limit for Infinite Systems from Data on Finite Systems. *Theor. Chim. Acta* **1995**, *90*, 1−26.

(21) Randić, M.; Klein, D. J.; Zhu, H.-Y.; Trinajstić, N.; Živković, T. Aromatic Properties of Fully-Benzenoid Hydrocarbons. *Fizika A* **1994**, *3*, 61−75.

(22) Živković, T.; Randić, M.; Klein, D. J.; Zhu, H.-Y.; Trinajstić, N. Analytical Approach to Very Large Benzenoid Polymers. *J. Comput. Chem.* **1995**, *16*, 517−526.

(23) Kasteleyn, P. W. Dimer Statistics and Phase Transitions. *J. Math. Phys.* **1963**, *4*, 287−293.

(24) Klein, D. J.; Liu, X. Many-Body Conjugated-Circuit Computations. *J. Comput. Chem.* **1991**, *12*, 1260−1264.

(25) Lukovits, I.; Graovac, A.; Kálmán, E.; Kaptay, G.; Nagy, P.; Nikolić, S.; Sytchev, J.; Trinajstić, N. Nanotubes: Number of Kekulé Structures and Aromaticity. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 609−614.

(26) Rongsi, C.; Cyvin, S. J.; Cyvin, B. N.; Brunvoll, J.; Klein, D. J. Methods of Enumerating Kekulé Structures Exemplified by Applications to Rectangle-Shaped Benzenoids. *Top. Curr. Chem.* **1990**, *153*, 227−253.

(27) Randić, M. Aromaticity of Polycyclic Conjugated Hydrocarbons. *Chem. Rev.* **2003**, *103*, 3449−3605.

(28) Cyvin, S. J.; Gutman, I. *Kekulé Structures in Benzenoid Hydrocarbons*; Springer-Verlag: Berlin, 1988; Chapter 4, pp 33−50.

(29) Li, S. H.; Ma, J.; Jiang, Y. S. Kekulé-Based Valence Bond Model. I. The Ground-State Properties of Conjugated π-Systems. *Chin. J. Chem.* **2002**, *20*(11), 1168−1179.

(30) Clar, E. *The Aromatic Sextet*; Wiley: London, 1972.

(31) Randić, M. Conjugated Circuits and Resonance Energies of Benzenoid Hydrocarbons. *Chem. Phys. Lett.* **1976**, *38*, 68−70.

(32) Pauling, L. The Calculation of Matrix Elements for Lewis Electronic Structures of Molecules. *J. Chem. Phys.* **1933**, *1*, 280−283.

(33) Ma, J.; Li, S. H.; Jiang, Y. S. Kekulé-Based Valence Bond Model. II. Diels-Alder Reactivity of Polycyclic Aromatic Hydrocarbons. *Chin. J. Chem.* **2002**, *20*(11), 1180−1191.

(34) Flocke, N.; Schmalz, T. G.; Klein, D. J. Valence Bond Multiple Excitation RVB Calculations on the Singlet Ground States of Some Carbon Fullerenes C_n with $20 \leq n \leq 60$. *Chem. Phys. Lett.* **1998**, *298*, 71−76.

(35) Klein, D. J.; Zhu, H.; Valenti, R.; Garcia-Bach, M. A. Many-Body Valence-Bond Theory. *Int. J. Quantum Chem.* **1997**, *65*, 421−438.

(36) Randić, M. Algebraic Kekulé Formulas for Benzenoid Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 365−372.

(37) Randić, M.; Balaban, A. T. Partitioning of π-electrons in Rings of Polycyclic Conjugated Hydrocarbons. Part 1. Catacondensed Benzenoids. *Polycyclic Arom. Comput.* **2004**, *24*, 173−193.

(38) Balaban, A. T.; Randić, M. Partitioning of π-electrons in Rings of Polycyclic Conjugated Hydrocarbons. Part 2. Catacondensed Coronoids. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 50−59.

(39) Balaban, A. T.; Randić, M. Partitioning of π-electrons in Rings of Polycyclic Conjugated Hydrocarbons. Part 3. Perifusenes. *New J. Chem.* **2004**, *28*, 800−806.

(40) Vukičević, D.; Randić, M.; Balaban, A. T. Partitioning of π-electrons in Rings of Polycyclic Conjugated Hydrocarbons. Part 4. Benzenoids with More Than One Geometric Kekulé Structure Corresponding to the Same Algebraic Kekulé Structure. *J. Math. Chem.* **2004**, *36*, 271−279.

(41) Balaban, A. T.; Randić, M. Partitioning of π-electrons in Rings of Polycyclic Conjugated Hydrocarbons. Part 5. Nonalternant Compounds. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1701−1707.

(42) Gutman, I.; Vukičević D.; Graovac, A.; Randić, M. Algebraic Kekulé Structures of Benzenoid Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 296−299.

(43) Miličević, A.; Nikolić; Trinajstić, N. Coding and Ordering Kekulé Structures. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 415−421.

(44) Flocke, N.; Schmalz, T. G.; Klein, D. J. Variational Resonance Valence Bond Study on the Ground State of C_{60} Using the Heisenberg Model. *J. Chem. Phys.* **1998**, *109*, 873−880.

(45) Randić, M. Enumeration of the Kekulé Structures in Conjugated Hydrocarbons. *J. Chem. Soc., Faraday Trans. 2* **1976**, *72*, 232−243.

(46) Klein, D. J.; Schmalz, T. G.; Hite, G. E.; Seitz, W. A. Elemental Carbon Cages. *J. Am. Chem. Soc.* **1986**, *108*, 1301−1127.

(47) Randić, M. Giant Benzenoid Hydrocarbons. Supernaphthalene Resonance Energy. *Acta Chim. Slov.* **1997**, *44*, 361−374.

(48) Cooper, I. L.; McWeeny, R. Studies in Configuration Interaction. I. Matrix Elements between Spin-Coupled Functions. *J. Chem. Phys.* **1966**, *45*, 226−234.

(49) Sutcliffe, B. T. Matrix Elements between Bonded Functions. *J. Chem. Phys.* **1966**, *45*, 235−243.

(50) Nesbet, R. K. Algorithm for Diagonalization of Large Matrices. *J. Chem. Phys.* **1965**, *43*, 311−312.