**Chemical insight structures Set code**

Sourav Roy & Avital Shurki

School of Pharmacy, Hebrew University of Jerusalem, Jerusalem, Israel

What is Chemical Insight Structures Set (CISS)?:

There are several ways to construct Valence Bond (VB) wave functions, such as Kotani’s branching method, Serber’s method, Geonological methods, Spin-paired spin function method etc. However, the spin-paired spin function has become popular due to its simple Kekule-type representation. Unlike all other methods, the number of spin-paired spin functions is much bigger than the dimension of the space. In the early 19th century, Rumer developed a diagrammatic method to find a linearly independent spin-paired spin function set. The diagrammatic method was very effective as it has no computational cost, but the drawback is its vast restrictions. Due to these restrictions, users cannot select the structures according to their interests. The CISS method provides much more freedom to the users to select the set of structures according to their interests. The CISS method produces all possible spin-paired spin functions and then evaluates their chemical qualities. The user can also define which chemical qualities are more important and which are less. They can also put some structures that should always be in the set (without any logic).

**The code can provide:**

1. A single highest-quality set of spin-paired spin functions.
2. All highest-quality sets together.
3. All possible sets together with their quality values.
4. All possible sets together with their Rumer Non-Rumer Identity.
5. Rumer set for that particular orbital numbering
6. All possible unique Rumer sets with the idea of their corresponding orbital numberings.
7. All sets with equally distributed bonds.

Command to compile: gfortran -o ex\_file\_name CISVB.f90

Command to Run: ./ex\_file\_name inputfile.xmi

X1e.int and INFO files need to be present in the folder.

**Input file:**

The input file is the same ‘.xmi’ file for xmvb code. It reads all the information from the ‘.xmi’ file such as control section, orbital section and bfi section. At the end of the xmi file we need to put the spatial keywords for the code, such as keyword for priority of quality factors, symmetric or non-symmetric structures etc. To read this spatial keyword we need to put a command “chinst=1” in the control section. Otherwise, the code will provide Rumer structures only. All the set of structures will be printed on ‘structure\_set\_n.dat’ file.

**Keywords:**

1. **Quality keywords:** There are five quality keywords.

**iab** (intra-atomic bond): nonphysical covalent bond within the same atom.

**nnb** (near neighbor bond): if the bond distance is less than or equal to the covalent bond distance, we take it as nearest neighbor bond. Intra atomic bond is also taken as nearest neighbor bond.

**sbb** (symmetry breaking bond): If the two bonded orbitals are possessed two different symmetries, we called it is a symmetry breaking bond.

**udr** (user’s defined radicle): For a non-singlet system we have unpaired electrons. User can put their preference to have these electrons in some orbitals.

**udb** (user’s defined bond): User can put their preference of bonds. The syntax is given below:

iab a

nnb b

sbb c

udr d

udb e

a, b, c, d and e are the serial number of importance of the quality factors. The default numbering is given below

iab 1

nnb 2

sbb 3

udr 0

udb 0

1. **Priority radical keyword:** If user choose a no-zero priority number for ‘udr’ we need to specify ‘prad’ keyword.

’prad’ specifies on which orbitals we expect singly occupied electrons to be preferential

sytax:

prad=n where n is the number of orbitals

I

J

K

Where I, j, and k are the identity of the preferred orbitals (given by their numbers).

1. ’lpst’ can be used if the user wishes to specify the radicals, mentioned above, should appear in specific lone-pair sets.

Syntax:

lpst n

L1 I, j, k …..

L2 I, j, k, ….

………………..

‘n’ is the number of lone paired given below. ‘n’=0 means no lone-pairs opted.

L1, L2 …. are the lone pairs.

I, j, k, ….. are the serial number of the radicals given in ‘prad’ section.

‘n’=0 if the default value.

1. **Priority bond keywords (imbd):** If user choose a no-zero priority bond keyword that is ‘udb ≠ 0’ we need to specify ‘imbd’ keyword.

’imbd’ specifies the bond user wants to have as much as possible in the set. The syntax is written below:

Imbd X

a – b c – d

‘X’ stands for the number of bonds. As an example, x=2 will be in the above.

a, b, c and d are the orbital numbers. a – b represents the bond between a and b orbitals.

1. **Priority structures keywords(imps):** If the users want some structures must be present in the set, they should use the keyword ‘imps’, but the structures must be linearly independent. The syntax is given below.

**Imps X**

**a – b c – d e – f**

**a – c b – e d – f**

‘X’ represents the number of structures the user wants to have in the set. In the above X=2. a, b, c, d…. are the orbital numbering and ‘a – b c – d e – f ‘ represents the structure**.**

1. **Set number keywords:** The set number keyword is ‘nset’ and the syntax is written below.

nset X Y

‘X’ stands for the type of sets

X=0 the first best set will be given (default).

X=1 all the highest quality sets will be given.

X=2 all the possible independent sets will be given.

X=3 all possible Rumer set has been printed in the file ‘sym\_str\_set.dat’. in the output file ‘structure\_set.dat’ it will write all possible sets and mentioned the Rumer’s sets.

X=4 all equally distributed bond sets will be given.

X=5 all the Rumer structures will be given in the ‘structure\_set.dat’ file.

‘Y’ is the upper limit of quality values of the printed structures. Which set of structures have greater values than Y will not be given in the ‘structure\_set.dat’. It will work for x=1, 2 and 3.

Default is X=0 and Y=1000 (means all allowed)

1. **Maximum output printing keyword:** For large systems (like 8o8e) the total number of sets (keyword ‘nset 0’) can be in the order of billions. Billions of sets cannot be printed in one file. Users can choose any number of the output file (not bigger than 1000) with the keyword ‘mout’. The syntax is given below.

**mout n**

n must be an integer between 1 and 1000.

n=1 stands for the single output file (default)

each output file contains maximum 75000 sets.

1. **Overlap keyword:** The overlap keyword is ‘ovlp’. Ovlp estimate the overlap of the structures in a set. Overlap works only with the keyword ‘nset 1’. nset 1 provides all the sets having Best qualities. User can choose the lowest overlap structure among equally qualified structures.
2. **Symmetry keyword:** The symmetry keyword is ‘sym’.

The syntax is: sym a b

a=loose is for the beginning of the calculation of symmetry. If it is found that the code could not be able to provide appropriate symmetric sets. Then a=tight can be applied. a=check is for checking some sets of structures are symmetric or not. In that case, the sets of structures need to write in ‘sym\_str\_set.dat’ file.

Default option for ‘a’ is loose.

b=stob will arrange the symmetric set from smaller to bigger quantity.

b=btos will arrange the symmetric set from bigger to smaller quantity.

b=qult will arrange the symmetric set according to structure quality.

Default option for ‘b’ is qult.

\*\* It is recommended to the user that the two types of pi orbitals should be given separately () in the orbital section for proper work of the program.

(a doubt arises: need to check the printing of final quality values. )

**Input-output file examples:**

**Ex 1.**

**Reaction N2+OH=N2O+H**

$ctrl

nao=6 nae=7

str=cov nmul=2

iscf=1 itmax=8000 bprep

guess=read

orbtype=oeo

**chinst=1**

$end

$bfi

3 60

1 2 3

2-19 21-38 40-63

$end

$orb

26 13 10 13 5 6 13 13 13 13 13

19 20 21 23-25 27-29 31-34 37-39 41-43 45-47 49-52 ;N1 N2 sxy

20 19 21 23-25 27-29 31-34 ; N1 xys

22 26 30 35 36 40 44 48 53 54 ; N1 N2 z

1 3 2 5-7 9-11 13-16 ; O syx

4 8 12 17 18 ; O z

55-60 ; H

3 1 2 5-7 9-11 13-16 ; O ysx

37 38 39 41-43 45-47 49-52 ; N2 syx

2 3 1 5-7 9-11 13-16 ; O xys

21 19 20 23-25 27-29 31-34 ;N1 ysx

39 37 38 41-43 45-47 49-52 ; N2 ysx

$end

**iab 1**

**udr 0**

**nnb 1**

**udb 2**

**sbb 0**

**prad 6**

**6**

**7**

**8**

**9**

**10**

**11**

**imbd 4**

**10 - 11 6 - 9 8 - 9 9 - 11**

**nset 0**

**output:**

inputfile :lscf\_ts\_R4\_3.xmi

\*\*\*\*\*\*\* CHEM. QUAL. STRUCTURES \*\*\*\*\*\*\*\*\*\*\*\*\*\*

30 covalent structures

IAB NNB SBB fqual

1 1 0 5 | 1: 5 10 10 7 8 9 11 6

1 1 0 5 | 1: 5 10 10 7 11 8 9 6

1 2 0 8 | 1: 5 10 10 6 7 8 9 11

1 2 0 8 | 1: 5 10 10 6 7 9 11 8

1 2 0 8 | 1: 5 10 10 6 9 7 8 11

quality value 34 bnd\_std 0.894

Set\_number= 1

1 2 0 8 | 1: 5 11 11 6 7 8 9 10

1 2 0 8 | 1: 5 11 11 6 9 7 8 10

1 2 0 8 | 1: 5 11 11 6 9 8 10 7

1 2 0 9 | 1: 5 11 11 6 7 8 10 9

1 3 0 11 | 1: 5 11 11 7 10 8 9 6

quality value 44 bnd\_std 0.894

Set\_number= 1

1 1 0 5 | 1: 5 9 9 7 8 10 11 6

1 1 0 6 | 1: 5 9 9 7 11 8 10 6

1 2 0 8 | 1: 5 9 9 6 7 10 11 8

1 2 0 9 | 1: 5 9 9 6 7 8 10 11

1 4 0 15 | 1: 5 9 9 6 8 7 11 10

quality value 43 bnd\_std 0.894

Set\_number= 1

1 2 0 7 | 1: 5 8 8 6 9 10 11 7

1 2 0 8 | 1: 5 8 8 6 7 9 11 10

1 2 0 8 | 1: 5 8 8 6 7 10 11 9

1 2 0 8 | 1: 5 8 8 6 9 7 11 10

1 3 0 11 | 1: 5 8 8 7 10 9 11 6

quality value 42 bnd\_std 0.894

Set\_number= 1

1 1 0 4 | 1: 5 7 7 8 9 10 11 6

1 1 0 5 | 1: 5 7 7 8 10 9 11 6

1 2 0 7 | 1: 5 7 7 6 9 10 11 8

1 2 0 8 | 1: 5 7 7 6 9 8 10 11

1 4 0 14 | 1: 5 7 7 6 8 9 11 10

quality value 38 bnd\_std 0.894

Set\_number= 1

1 1 0 4 | 1: 5 6 6 8 9 10 11 7

1 1 0 5 | 1: 5 6 6 7 8 9 11 10

1 1 0 5 | 1: 5 6 6 7 8 10 11 9

1 1 0 5 | 1: 5 6 6 7 11 8 9 10

1 1 0 5 | 1: 5 6 6 8 10 9 11 7

quality value 24 bnd\_std 0.894

Set\_number= 1

# different lone pairs belong to different spaces. Therefore, calculated and printed separately.

# IAB, NNB and SBB are different qualities of each structure. ‘fqual’ is the final quality value of the structure.

# ‘quality value 24 ‘ means the total quality value of the set. It is basically the addition of the ‘fqual’.

# ‘bnd\_std’ is the standard deviation of the bonds present in the sets. It would be 0 for equally distributed sets.

**EX 2:**

**molecule C2**

$ctrl

nao=8 nae=8

str=cov nmul=1

iscf=5 itmax=800 boys

guess=mo

orbtyp=hao frgtyp=sao

**chinst=1**

$end

$frag

1 1 1 1

spzdxxdyydzzdxy 1

pxpydxzdyz 1

spzdxxdyydzzdxy 2

pxpydxzdyz 2

$end

$orb

1\*10

1

3

1

3

1

3

2

4

2

4

$end

$gus

1 1

2 2

3 3

4 4

5 7

6 10

7 5

8 6

9 8

10 9

$end

**nset 0**

**imp\_stru 5**

**3 6 4 5 7 10 8 9**

**3 6 4 8 5 9 7 10**

**3 6 4 10 5 7 8 9**

**3 6 4 5 7 9 8 10**

**3 6 4 10 5 8 7 9**

**Output:**

inputfile :C2\_chem\_IS\_1.xmi

\*\*\*\*\*\*\* CHEM. QUAL. STRUCTURES \*\*\*\*\*\*\*\*\*\*\*\*\*\*

14 covalent structures

IAB NNB SBB fqual

**1 1 1 1 | 1: 2 3 6 4 5 7 10 8 9**

**3 1 1 11 | 1: 2 3 6 4 5 7 9 8 10**

**3 1 3 13 | 1: 2 3 6 4 8 5 9 7 10**

**3 1 3 13 | 1: 2 3 6 4 10 5 7 8 9**

**3 1 3 13 | 1: 2 3 6 4 10 5 8 7 9**

1 1 1 1 | 1: 2 3 4 5 6 7 8 9 10

1 1 1 1 | 1: 2 3 4 5 6 7 10 8 9

1 1 3 3 | 1: 2 3 4 5 8 6 7 9 10

1 1 3 3 | 1: 2 3 4 5 8 6 9 7 10

1 1 3 3 | 1: 2 3 4 5 10 6 7 8 9

1 1 3 3 | 1: 2 3 8 4 5 6 7 9 10

1 1 3 3 | 1: 2 3 8 4 5 6 9 7 10

1 1 3 3 | 1: 2 3 10 4 5 6 7 8 9

1 1 5 5 | 1: 2 3 8 4 7 5 10 6 9

quality value 76 bnd\_std 1.773

Set\_number= 1

## The structures specified in the input file is present at the top of the set in the output file.

# IAB, NNB and SBB are different qualities of each structure. ‘fqual’ is the final quality value of the structure.

# ‘quality value 76 ‘ means the total quality value of the set. It is basically the addition of the ‘fqual’.

**Ex 3:**

**molecule C6H6**

$ctrl

nao=6 nae=6

str=cov nmul=1

iscf=5 itmax=4000 boys

guess=unit

**chinst=1**

$end

$stru

$end

$orb

78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 4 4 4 4 4 4

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

1-4 6-8 10-13 16-21 23-25 27-30 33-38 40-42 44-47 50-55 57-59 61-64 67-72 74-76 78-81 84-89 91-93 95-98 101 102

5 9 14 15

22 26 31 32

39 43 48 49

56 60 65 66

73 77 82 83

90 94 99 100

$end

**Sym**

**nset 0**

**Output file:**

inputfile :c6h6.xmi

\*\*\*\*\*\*\* CHEM. QUAL. STRUCTURES \*\*\*\*\*\*\*\*\*\*\*\*\*\*

5 covalent structures

IAB NNB SBB fqual

1 1 0 1 | 1: 18 19 20 21 22 23 24

1 1 0 1 | 1: 18 19 24 20 21 22 23

1 2 0 2 | 1: 18 19 20 21 24 22 23

1 2 0 2 | 1: 18 19 22 20 21 23 24

1 2 0 2 | 1: 18 19 24 20 23 21 22

quality value 8

Set\_number= 1