

autoVB 使用文档

1. 简介

autoVB (当前版本 v0.1.4) 是一款能够从分子结构出发、全自动完成 XMVB 计算的智能化工具。其工作流程涵盖：自动调用 Gaussian 程序执行 NBO 计算、智能选取活性空间、提供可靠的初始猜测，随后自动生成 XMVB 输入文件并执行计算，最终输出可视化结果分析。整个过程无需人工干预，显著降低了价键理论计算的门槛。

如果你对价键理论还不了解，可以查看 [XMVB 中文教程](#)。

2 输入文件

2.1 输入文件示例

输入文件格式包括两种类型，默认类型与 Gaussian .gjf 格式高度相似：

```
%mem=4GB
%nprocshared=4
#p vbscf(4,4)/cc-pVDZ

autovb{}

0 1
C   -3.329219900    2.091826940    0.000000000
H   -2.839967430    1.120094350   -0.000000040
H   -4.414989510    2.089928580    0.000000000
C   -2.624068780    3.231943550    0.000000010
H   -3.149818500    4.187367240    0.000000030
C   -1.168018280    3.297558130    0.000000020
H   -0.642289190    2.342122530    0.000000040
C   -0.462840870    4.437656730   -0.000000010
H   -0.952055180    5.409407090   -0.000000040
H    0.622929910    4.439514340   -0.000000010
```

另一种输入文件格式与 XMVB .xmi 文件格式类似，你可以将输入文件后缀名改为 .xmi 以使用该类型，如果后缀名不正确，autoVB 也将会自动检查你是否使用了该格式。具体格式参见 [XMVB 文档](#)：

```
C4H6
$ctrl
vbscf
nao=4
nae=4
basis=cc-pvdz
$end

$actorb
1
4
6
8
$end

$geo
C   -3.329219900    2.091826940    0.000000000
H   -2.839967430    1.120094350   -0.000000040
H   -4.414989510    2.089928580    0.000000000
C   -2.624068780    3.231943550    0.000000010
H   -3.149818500    4.187367240    0.000000030
C   -1.168018280    3.297558130    0.000000020
H   -0.642289190    2.342122530    0.000000040
C   -0.462840870    4.437656730   -0.000000010
H   -0.952055180    5.409407090   -0.000000040
H    0.622929910    4.439514340   -0.000000010
$end
```

2.2 控制命令行 (# 所在行)

命令行的格式需要是 # 方法(nae,nao)/基组 或 # 方法/基组, 如果采用 .xmi 格式则是

在 \$ctrl 块内填写参数。

2.2.1 方法

目前支持的计算方法包括 vbscf, vbpt2, lam-dfwb, blw, tbvbscf。blw 方法没有活性空间选择相关选项。

2.2.2 显式指定活性空间

你可以直接在方法名后面使用括号显式地声明活性电子数和活性轨道数。例如：vbscf(4,4)。如果不指定，autoVB 会自动确定活性空间。这会覆盖下方在 autovb{}中的设置！

2.2.3 基组

计算使用的基组，支持任何在 Gaussian 和 XMVB 中都支持的基组。

2.3 标题行控制参数

在输入文件部分的标题行位置，你可以写入类似 autovb{...}的语法来精准控制价键计算。

2.3.1 活性空间选择相关参数

自动选择活性空间时，autoVB 会先读取 Gaussian NBO 结果中的轨道类型、占据数和原子连接关系，然后按以下优先级确定活性轨道：

1. 如果给出 aat，按指定原子搜索相关 NBO 轨道；如果给出 aoa_bond，按指定原子对搜索相关成键轨道。
2. 如果给出 aoi，直接使用指定的 NBO 轨道编号。
3. 如果只给出 nae/nao，程序会在占据轨道中按 NBO 占据数从低到高挑选 NBO 轨道。
4. 如果给出 threshold，程序会先选择占据数满足 $1 < occ \leq threshold$ 的对应轨道。
5. 如果以上参数都未给出，则使用默认规则：选择低占据的 BD 轨道、满足 BD 低占据且对应 BD*有一定占据的成键/反键对，以及低占据的 LP 轨道。默认阈值为 $BD < 1.96$ 、 $BD < 1.99$ 且 $BD^* > 0.06$ 、 $LP < 1.96$ ；若设置了 nolp，则跳过 LP 轨道。

默认规则会根据选出的电子数和轨道数自动收紧阈值，尽量避免生成过大的活性空间；如果仍然过大，程序会给出警告，建议手动指定 aat、aoi 或 nae/nao。

2.3.1.1 nae=n

显式指定活性电子数。如果不指定，autoVB 会自动根据 NBO 占据数选定活性空间。

2.3.1.2 nao=n

显式指定活性轨道数。如果不指定，autoVB 会自动根据 NBO 占据数选定活性空间。

2.3.1.3 aat=(n1, n2, ...)

活性原子序列 (Active Atom)。指定活性空间所在涉及的原子的索引。比如 aat=(1,2,3,4) 会自动搜索 1-4 号原子的 NBO 轨道作为活性轨道。原子编号从 1 开始，范围在 [1, 分子总原子数] 的整数之间。对于 .xmi 格式的输入文件，该参数将使用 \$actorb 部分的数据，格式见 [输入文件示例](#)。

2.3.1.4 aoi=(n1, n2, ...)

活性轨道的索引编号 (Active Orbital Indices)。直接指定活性空间所使用的 NBO 轨道索引 (从 1 开始)。比如 aoi=(1,2,3,4) 会将 1-4 号 NBO 轨道作为活性轨道。

2.3.1.5 threshold=n

根据 NBO 占据数挑选活性轨道的阈值。程序基于 NBO 占据数自动挑选活性空间。必须大于等于 1。

2.3.1.6 rethre=n

按照修正占据数进行筛选，修正占据数是 NBO 的成键-反键对占据数修正为加和为 2 的占据数。

2.3.1.7 active_order=n

决定如何重新排列活性轨道的顺序，默认值是 rumer，如果设置了 [aat](#) 参数，则默认值是 aoa。它有如下选项：

aoa: 按照用户给定的 aat 顺序排序。

rumer: 按照选择的活性空间原子在分子中的拓扑顺序排序，会尽量避免交叉。

seq: 按照数字从大到小排序。

none: 不进行排序。

2.3.1.8 aoa_bond=(n1, n2, ...)

成键活性轨道原子序列。指定活性空间所在涉及的原子的索引。比如 aoa_bond=(1,2,3,4)表示原子 1 和 2 成键, 3 和 4 成键, 将会选取 1 与 2 的键和 3 与 4 的键为活性空间。如果输入奇数个索引, 则最后一个原子索引将会被识别为孤对电子。原子编号从 1 开始, 范围在 [1, 分子总原子数] 的整数之间。注意如果期望的活性空间中有超过 1 对孤对电子, 它将无法正常工作 (因它会认为两个原子成键), 因此不推荐使用该选项。

2.3.1.9 bond_first

[aat](#) 参数的子选项, 开启该选项后会优先搜索成键轨道而非孤对电子。如果你期望的活性空间内没有占据数较低的孤对电子时可开启它。

2.3.1.10 nolp

不选择孤对电子, 支持 [aat](#) 参数和无参数的默认方式挑选轨道。它和 [bond_first](#) 参数的区别是 nolp 会强制忽略孤对电子, 而 [bond_first](#) 在没有可用的成键轨道时依然会选择孤对电子。

2.3.2 NBO 计算相关参数

2.3.2.1 nbo=n

使用的 NBO 轨道计算方法, 默认为 hf, 使用不同的方法可能会获得不同的活性空间大小和初猜。因未经验证, 不建议修改为 hf 和常见 DFT 泛函 (如 B3LYP) 以外的方法。

2.3.2.2 nbo_file=n

读取已经计算完成的 NBO 文件, n 为带后缀的文件名, 需要该文件和输入文件在同一工作目录下。这会从.fch 文件中读取分子信息, 从.out 文件中读取 NBO 输出信息, 从.37 文件中读取 NBO 轨道, 从.36 文件中读取 PNBO 轨道。此时 autoVB 将不进行 NBO 计算。

注意: autoVB 暂时会忽略 NBO 输出文件中的坐标, 而是使用 autoVB 输入文件中给出的坐标, 所以请确保 NBO 计算使用的坐标与 autoVB 输入文件中给出的坐标一致。

2.3.3 XMVB 计算相关参数

2.3.3.1 str=n/stru=n

配置 XMVB 将要在计算中生成的价键结构类型，可选值包括"full", "cov", "ion(a,b,c)", "ion(a-b)"。默认值为"full"，具体含义参见 [XMVB 文档](#)。

2.3.3.2 int=n/inte=n

配置 XMVB 计算积分的方式。可选值包括 libcint, ri, cosx。默认值为 libcint，具体含义参见 [XMVB 文档](#)。

2.3.3.3 iscf=n

配置 XMVB 轨道优化的方式。可选值包括 2, 5, 6。默认值为 5，具体含义参见 [XMVB 文档](#)。

2.3.3.4 sort

是否按照结构系数排序价键结构。

2.3.3.5 guess=n

生成的 XMVB 使用的初猜以及如何挑选活性空间的设置，可选项包括 nbo, pnbo。默认值为 nbo。

nbo: 使用 NBO 轨道作为初猜和挑选活性空间的依据。

pnbo: 使用 NBO 轨道作为挑选活性空间的依据，使用 PNBOPNBO 轨道作为初猜。PNBO 轨道是非正交且严格定域的。

2.3.3.6 atom_slice

对分子轨道作单原子截断（将指定原子外的基函数清零），以生成定域的初猜。一般来说，XMVB 会自动进行截断，所以不需要使用此功能。

2.3.3.7 novb

不进行 XMVB 计算，只生成输入文件。当环境中没有可用的 XMVB 或希望手动修改生成的输入文件时推荐使用该选项。

2.3.4 绘图相关

2.3.4.1 draw_xmo

调用 draw_xmo 工具，对.xmo 文件进行解析，生成分子价键结构.svg 图像文件。画出的分子构型仅供参考，可能存在错误。

2.3.4.2 xmo2svg=n

在 XMVB 计算结束后解析.xmo 文件，并生成与.xmo 文件同名的价键结构.svg 图像。支持以下选项：

不设置：不生成.svg 图像。

xmo2svg=rdkit：使用 RDKit 根据分子键连关系生成标准二维构型。适合普通单片段分子，不保留输入文件中的三维构型特征。

xmo2svg=pca：使用主成分分析将输入分子的三维坐标投影到二维平面。它会尽量保留原始三维构型的整体方向，适合希望图像反映输入构型的体系。

xmo2svg=optimized3d：从多个观察方向中搜索原子重叠、键压缩和键交叉较少的三维投影视角，再将分子正交投影到二维平面。适合非平面分子，通常比直接 PCA 投影更清晰。

xmo2svg=contact：优先根据价键结构中的跨片段原子对排列不同分子片段，适合反应物复合物、过渡态等包含多个分子片段的体系。对于只有一个片段的分子，会使用 RDKit 的标准二维构型。

2.3.4.3 hide_svg_labels

隐藏 xmo2svg 图中的原子编号，以及每个结构权重后面的成键原子对标记。该参数需要与 xmo2svg 一起使用。

2.3.4.4 svgweight=n

选择 xmo2svg 图中显示的结构权重，可设为 cc、lowdin 或 both，默认值为 both。当设置为 both 时，CC 权重和 Lowdin 权重分两行显示，结构仍按 Lowdin 权重选择和排序。

2.3.4.5 draw_rumer

使用 Rumer 规则选择的活性空间原子在分子中的拓扑顺序排序时，生成分子.svg 图像文件。

2.3.5 其他参数

2.3.5.1 debug

调试模式，会打印更多输出。

3 示例

3.1 苯分子的自动计算

该例展示不手动指定活性空间时，autoVB 如何完成 NBO 计算、自动选择活性空间、生成 XMVB 输入、运行 VBSCF，并将主要价键结构绘制成图。输入文件如下：

```
# vbscf/6-31g*

autovb{str=cov}

0 1
C      0.6995000584   1.2115696411   0.0000000000
H      1.2460106991   2.1581538376   0.0000000000
C     -0.6995000584   1.2115696411   0.0000000000
H     -1.2460106991   2.1581538376   0.0000000000
C     -1.3990001169   0.0000000000   0.0000000000
H     -2.4920213982   0.0000000000   0.0000000000
C     -0.6995000584  -1.2115696411   0.0000000000
H     -1.2460106991  -2.1581538376   0.0000000000
C      0.6995000584  -1.2115696411   0.0000000000
H      1.2460106991  -2.1581538376   0.0000000000
C      1.3990001169   0.0000000000   0.0000000000
H      2.4920213982   0.0000000000   0.0000000000
```

该例使用 VBSCF/6-31G* 计算电荷为 0、自旋多重度为 1 的苯分子。由于没有指定 nae、nao、aat 或 aoi，autoVB 会根据 NBO 分析结果自动选择活性空间。

3.1.1 活性空间选择结果

程序首先得到 42 个电子、21 条占据轨道和 81 条虚轨道，然后按照默认阈值选中三条低占据的 BD 轨道及其对应的 BD*轨道：

NBO 轨道	涉及的原子	BD 占据数	BD*占据数
4	C1-C11	1.66476	0.33237
7	C3-C5	1.66476	0.33237
12	C7-C9	1.66476	0.33237

每条占据 BD 道与一条对应的 BD*轨道共同进入活性空间，因此最终得到 6 个活性电子和 6 条活性轨道，即(nae, nao)=(6,6)。21 条占据 NBO 中有 3 条作为活性占据轨道，其余 18 条作为非活性轨道。随后，程序根据 Rumer 规则将活性原子排列为：

[1, 3, 5, 7, 9, 11]

这些编号对应输入文件中的 6 个碳原子。生成的 XMVB 输入文件在\$ctrl 中写入 nae=6 和 nao=6，并将上述原子顺序写入\$orb 的活性轨道部分。

3.1.2 XMVB 计算结果

XMVB 计算成功收敛，关键结果为：

项目	结果
计算方法	VBSCF
基组	6-31G*
活性空间	(6,6)
收敛状态	已收敛
迭代次数	51
VBSCF 能量	-230.56875969 a.u.

生成的价键结构数	5
----------	---

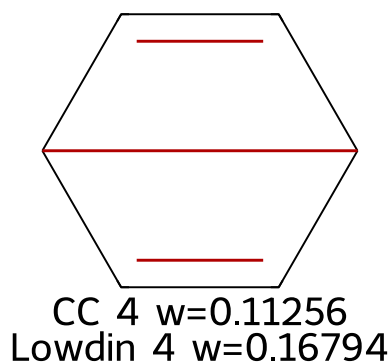
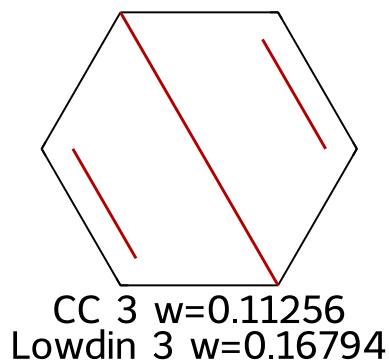
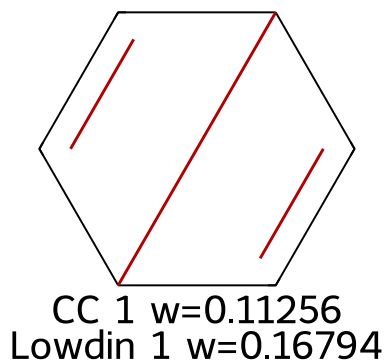
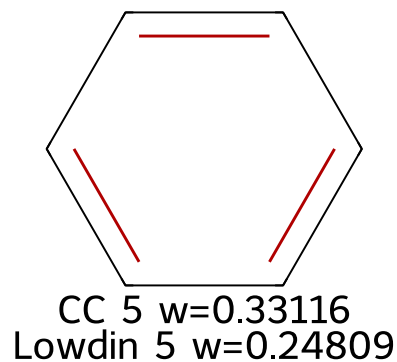
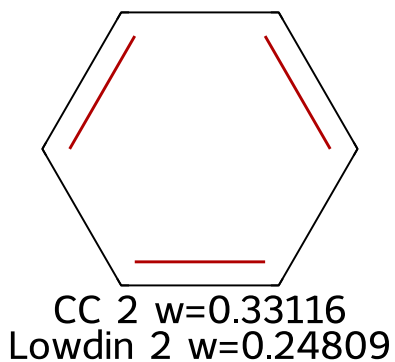
按权重排序时，最主要的两个结构为：

结构编号	成键原子对（输入文件原子编号）	CC 权重	Lowdin 权重
2	3-5、7-9、1-11	0.33116	0.24809
5	1-3、5-7、9-11	0.33116	0.24809
1	5-7、3-9、1-11	0.11256	0.16794
3	3-5、1-7、9-11	0.11256	0.16794
4	1-3、7-9、5-11	0.11256	0.16794

这 5 个结构分别对应苯环中两种 Kekulé 形式以及三种对称等价的 Dewar 形式，且两种 Kekulé 结构具有相同的最大权重，符合苯分子的对称性。

3.1.3 价键结构图

绘图默认使用 CC 权重和 Lowdin 权重。红线表示当前价键结构中的活性成键关系。



图中展示了 5 个主要共价结构，其中 2 和 5 为两种 Kekulé 结构，权重最高；1、3 和 4 为三种 Dewar 型结构，权重较低。该结果说明在完整的 5 结构模型中，苯分子的共振描述依然以两种 Kekulé 结构为主，但也包含少量 Dewar 贡献。

3.2 指定结构的价键计算

如果想要指定某些特定的价键结构，通常需要使用 XMVB 的.xmi 输入文件格式，而不是普通的 Gaussian 格式。该例展示了如何在苯分子的计算中指定两种 Kekulé 结构。

输入文件如下：

```

$ctrl
autovb
vbscf
basis=6-31g*
$end

$str
1:18 19 20 21 22 23 24
1:18 19 24 20 21 22 23
$end

$geo
C      0.6995000584   1.2115696411   0.0000000000
H      1.2460106991   2.1581538376   0.0000000000
C     -0.6995000584   1.2115696411   0.0000000000
H     -1.2460106991   2.1581538376   0.0000000000
C     -1.3990001169   0.0000000000   0.0000000000
H     -2.4920213982   0.0000000000   0.0000000000
C     -0.6995000584  -1.2115696411   0.0000000000
H     -1.2460106991  -2.1581538376   0.0000000000
C      0.6995000584  -1.2115696411   0.0000000000
H      1.2460106991  -2.1581538376   0.0000000000
C      1.3990001169   0.0000000000   0.0000000000
H      2.4920213982   0.0000000000   0.0000000000
$end

```

这是 XMVB.xmi 格式的输入文件：\$ctrl 块用于控制计算方法、基组和图像输出，\$str 块给出两种价键结构的连接模式，\$geo 块给出分子坐标。

3.2.1 活性空间选择结果

与 [3.1.1](#) 中的自动选择方式一致，程序仍会在苯环中选出 3 对低占据的 BD/BD*轨道，最终得到 (nae, nao)=(6,6)的活性空间。

3.2.2 XMVB 计算结果

XMVB 计算成功收敛，关键结果为：

项目	结果
计算方法	VBSCF
基组	6-31G*
活性空间	(6,6)
收敛状态	已收敛
迭代次数	49
VBSCF 能量	-230.56408042 a.u.
生成的价键结构数	2

两种价键结构权重几乎完全相同，分别对应苯环的两种 Kekulé 成键方式：

结构编号	成键原子对（输入文件原子编号）	CC 权重	Lowdin 权重
1	1-3、5-7、9-11	0.49999	0.49999
2	1-11、3-5、7-9	0.50001	0.50001

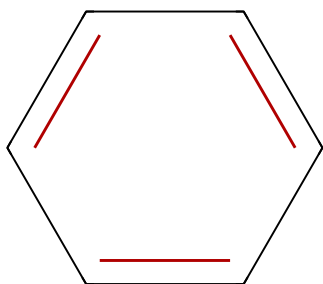
这说明在这种两结构限制下，苯分子的两种 Kekulé 结构在价键波函数中具有几乎等价的贡献。进一步，与 [3.1.1](#) 中包含 5 个价键结构的计算结果对比：

$$\Delta E = E_{4.2} - E_{4.1} = (-230.56408042) - (-230.56875969) = 0.00467927 \text{ a.u.}$$

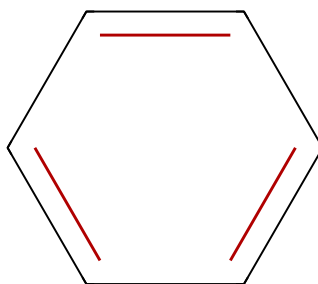
其中， $E_{4.1} = -230.56875969 \text{ a.u.}$ 是 5 个价键结构（2 个 Kekulé 结构加上 3 个 Dewar 结构）的最终能量， $E_{4.2} = -230.56408042 \text{ a.u.}$ 是 2 个 Kekulé 结构的最终能量。该差值约等于 2.94 kcal/mol，说明三个 Dewar 苯结构提供约 2.94 kcal/mol 的共振能。

3.2.3 价键结构图

该图展示了两种 Kekulé 结构的绘图结果，可以看到，两种 Kekulé 结构的权重几乎完全相同。



CC 2 w=0.50002
Lowdin 2 w=0.50002



CC 1 w=0.49998
Lowdin 1 w=0.49998

3.3 复杂体系的指定活性空间计算

该例展示当默认自动选择不能准确表达目标反应区域时，如何通过原子编号指定活性空间。输入文件如下：

```
#p vb3cf/6-31g(d)

autovb{str=cov,nolp,aat=(5,4,2,1,6,7,8,9)}

0 1
C      -0.61461      1.64866      -0.14318
C       0.48763       0.94566      -0.65308
H       0.71579       0.95698      -1.71111
C       1.41105       0.27025       0.26981
O       1.10384       0.15207       1.45186
C      -2.15769       0.56469      -0.13681
C      -1.90990      -0.62090       0.58183
C      -0.95283      -1.51089       0.05185
O      -0.34442      -1.22918      -1.00876
C       2.69763      -0.28811      -0.29293
H      -2.89530       1.24987       0.27046
H      -0.66170      -2.39630       0.63398
H      -1.01367       2.43637      -0.77523
H      -2.26681      -0.75770       1.59602
H      -2.20092       0.45166      -1.21511
H      -0.55594       1.89304       0.91218
H       3.28877      -0.73227       0.50695
H       2.46925      -1.04416      -1.04890
H       3.27747       0.50325      -0.77898
```

该分子为 Diels-Alder 反应过渡态，使用 VBSCF/6-31G(d)计算。控制参数的含义如下：

- aat=(5,4,2,1,6,7,8,9): 指定活性空间涉及的原子, 并规定活性轨道的原子排列顺序。
- nolp: 搜索 NBO 轨道时排除孤对电子, 避免氧原子的孤对电子被选入目标活性空间。
- str=cov: 只生成共价型价键结构, 不展开离子型结构。

3.3.1 活性空间选择结果

NBO 计算得到 68 个电子、34 条占据轨道和 121 条虚轨道。程序按照用户给定的 aat 原子序列和 nolp 规则选出 4 条占据活性轨道, 并为每条占据轨道加入相应的未占据轨道, 最终得到(nae, nao)=(8,8):

项目	结果
活性电子数	8
活性轨道数	8
活性占据轨道数	4
非活性占据轨道数	30
活性原子顺序	5、4、2、1、6、7、8、9

3.3.2 XMVB 计算结果

XMVB 计算成功收敛，关键结果为：

项目	结果
计算方法	VBSCF
基组	6-31G(d)
活性空间	(8,8)
结构类型	共价结构（str=cov）
收敛状态	已收敛
迭代次数	97
VBSCF 能量	-420.26804610 a.u.
生成的价键结构数	14

按 Lowdin 权重排序时，权重最高的五个结构为：

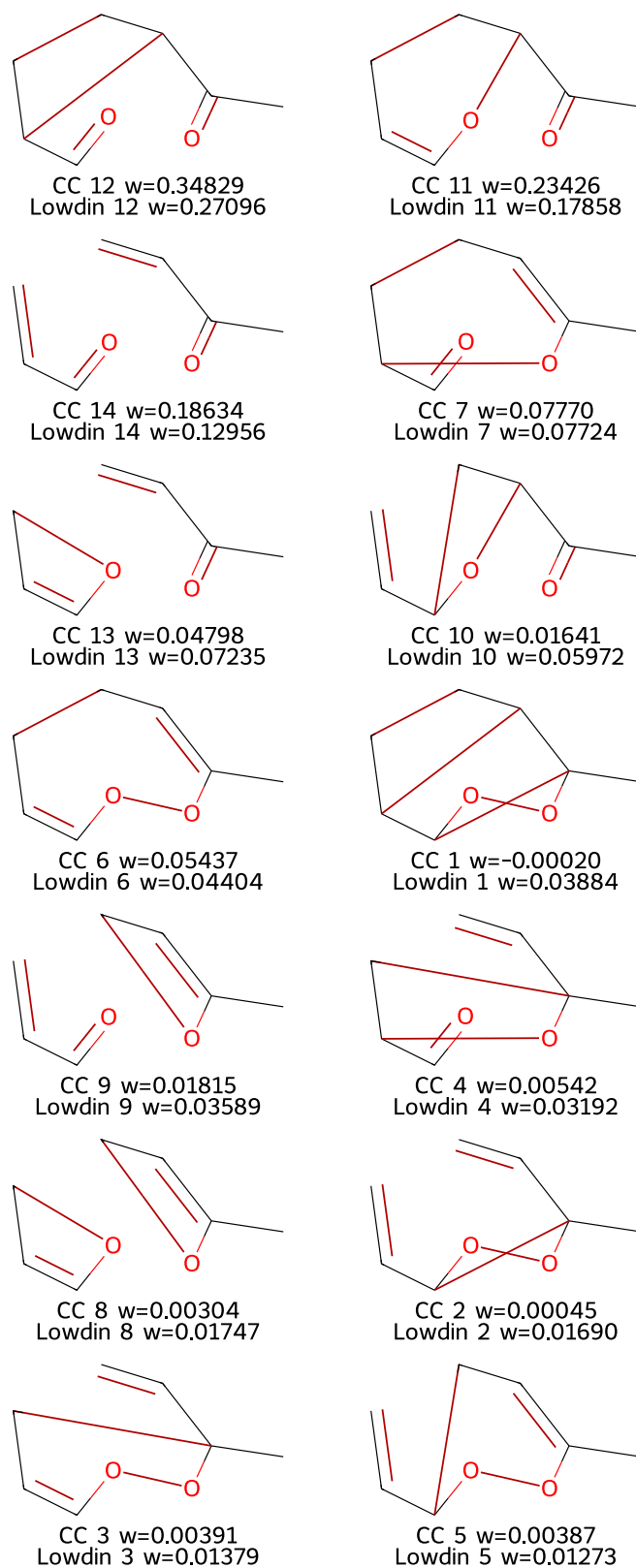
结构编号	成键原子对	Lowdin 权重	CC 权重
12	5-4、1-6、2-7、8-9	0.27096	0.34829

11	5-4、1-6、7-8、2-9	0.17858	0.23426
14	5-4、2-1、6-7、8-9	0.12956	0.18634
7	4-2、1-6、5-7、8-9	0.07724	0.07770
13	5-4、2-1、7-8、6-9	0.07235	0.04798

结构 12 的权重最高，是该计算结果中的主要共价结构；结构 14 也结构 11 和有较大贡献。

3.3.3 价键结构图

计算结束后生成的图片使用 Lowdin 权重排列全部 14 个共价结构，并从输入的三维坐标中搜索较清晰的二维投影视角，从而生成图片。图片下方同时展示 CC 权重和 Lowdin 权重。图中的红线表示活性空间内的电子配对，黑线表示基础分子键连，结构下方依次给出结构编号、权重和成键原子对。



左上角的结构 12 具有最大的权重。其余结构展示了同一组活性原子之间不同的电子重配对方式，可以直观看出各共价结构对总波函数的相对贡献。