CUTOFF

Philippe B. Wilson and Ian H. Williams

University of Bath

CUTOFF is a utility program enabling a user to "cut off" elements of a large Hessian matrix (typically from an electronic-structure calculation of second derivatives for a large molecular system) in order to produce a smaller Hessian corresponding to a specific subset of atoms selected by the user.

Written originally by IHW (University of Bath, UK), CUTOFF deletes those rows and columns of the full input Hessian which do not involve the selected atoms and retains those that do. The output from program CUTOFF is used as input for program UJISO, available elsewhere at the following Github repository: https://github.com/pbw20/SULISO).

CUTOFF is written in Fortran90, with some elements of earlier versions, but compiles fluidly under standard GNU compilers (gfortran…).

The aim of this user guide is to gently introduce the individual to the functionalities of CUTOFF and provide explicit descriptions of the input and output one can expect from standard CUTOFF runs. We will begin by discussing the format of the input, suggesting possible edits to the source code, which is available from the SULISO_source repository on Github: https://github.com/pbw20/SULISO_source

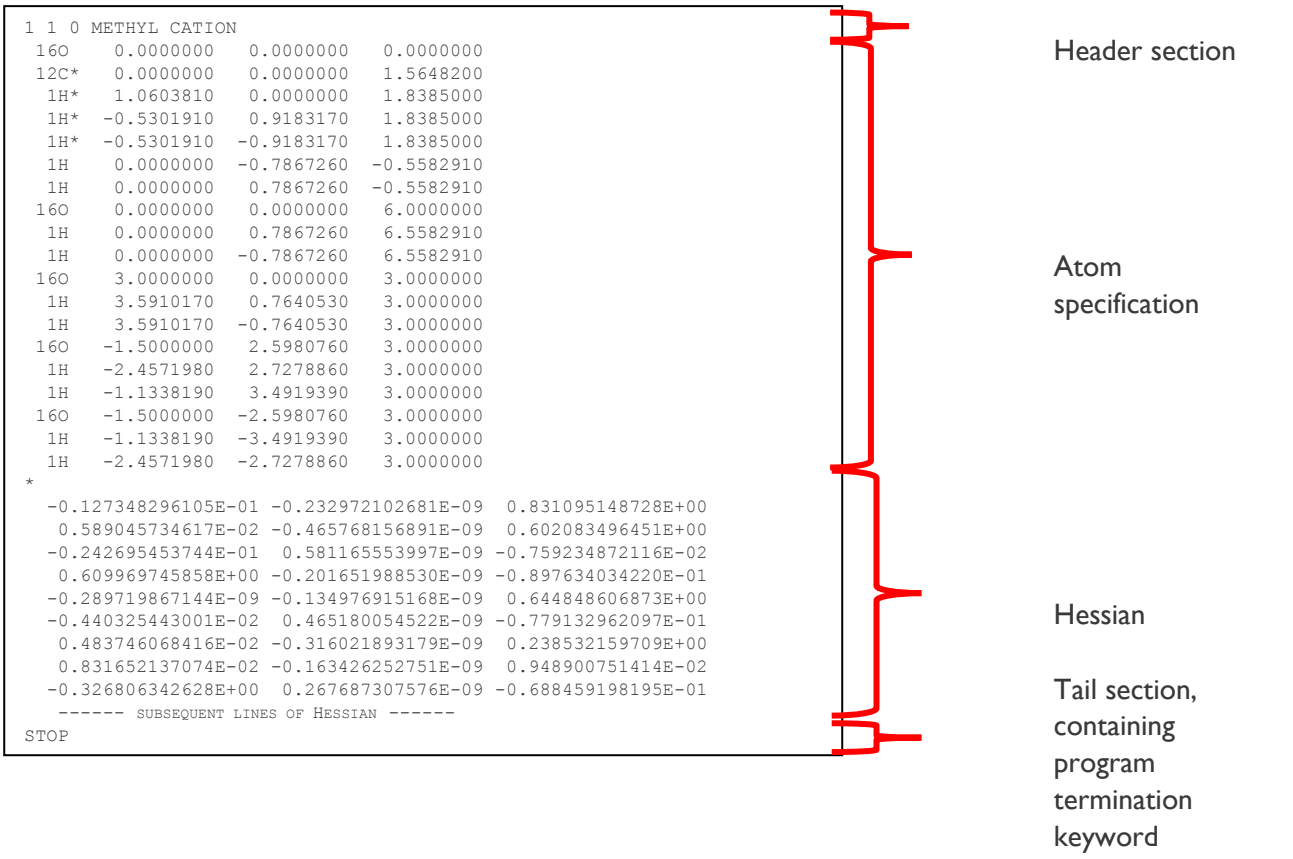

Figure 1. Sample input file for CUTOFF. In the interest of space, only the first few lines of the Hessian are included, with the "------ SUBSEQUENT LINES OF HESSIAN ------" line indicating their original position. This example is for a molecular complex (19 atoms) comprising a methyl cation inside a cage of five water molecules; the four atoms of the methyl moiety are selected and the output Hessian (as its lower triangle) will contain 78 unique elements as compared with 1653 on input.

Overview of CUTOFF input

1. Header section, line 1: (FORMAT 3I2,74A)

   ```
   IS,IG,IF,TITLE
   ```

   The three integer values (3I2) are not used by CUTOFF at all. (They are data required by a post- CUTOFF program.) The alphanumeric title (74A) is also not used by CUTOFF but does allow for descriptive comment.

2. Coordinate input, one line per atom, from line 2 to line 2 + *N*, where *N* is the number of atoms: (FORMAT A4,A1, 3F12.8)

   ```
   IA,IFK,(X(J,IOLD),J=1,3)
   ```

   IA is the element symbol (right-justified in the A4 field) for the IOLDth atom;

   IFK is *either* a space character, signifying that this atom is to be cutoff,
   (so that its corresponding rows and columns in the Hessian are excluded from the output file)
   or an asterisk (*) - which signifies that this atom is to be retained (kept);

   X are the *x*, *y*, and *z* Cartesian coordinates in Ångström (3F12.7).

   The following element symbols are currently accepted, although the UJISO source code may be edited to make additions or changes:

   1H, 2H, 3H, 12C, 13C, 14C, 14N, 15N, 16O, 17O, 18O, F, 32S, 34S, 35CL, 37CL, 6LI, 7LI, 9BE, 10B, 11B, SI, P, 79BR, 81BR, NA, SP, MG, AL, CL, K, CA, ZN, BR, RB, I, CS, B, H, C, N, O, S, HE, NE, AR, KR, 33S, 36S, 11C

   An element symbol preceded by an atomic mass specify a particular isotope (e.g. ^{1}H) whereas an element symbol without a mass specifies the chemical atomic weight (e.g. H = 1.00797). The special symbol 'SP' specifies a mass of 999.

3. Data separation line, line 3 + *N* : (FORMAT A4)

   Consists of an asterisk '*' (left-justified in the A4 field) to separate the coordinate input from the Hessian. (This character is a special value for 'IA' in item 4.)

4. Hessian input lines, from line 4 + *N* to line 4 + *N* + *m*: (FORMAT: 3E20.10)

   ```
   (F(I),I=1,NNC)
   ```

   This is the lower triangle of a square symmetric Hessian matrix, stored as a linear array. The number of unique elements is NNC = *N* × (*N* + 1)/2. The number of three-column lines of data in this section is therefore given by:

   $$\text{Number of lines of Hessian, } m = \frac{3N}{2} \times \frac{3N+1}{3}$$

   The number of lines of Hessian output from a calculation can be used to verify the reasonableness of the input, by comparing this with the number of atoms using the formula above.

5. Final line: (FORMAT A4)

   The final line of the program consists of the word STOP, followed by a line break, in order to inform the program that no more input remains to be read.

Overview of CUTOFF output

```
1 1 0 METHYL CATION
 12C*    0.0000000    0.0000000    1.5648200
  1H*    1.0603810    0.0000000    1.8385000
  1H*   -0.5301910    0.9183170    1.8385000
  1H*   -0.5301910   -0.9183170    1.8385000
*
    0.6099877817E+00   -0.1171338636E-09    0.6448654246E+00
    0.4840786899E-02   -0.3147298145E-09    0.2385321597E+00
   -0.3268000661E+00    0.2858926120E-09   -0.6884236124E-01
    0.3321574387E+00   -0.1852785845E-09   -0.5617904058E-01
    0.5950832891E-10    0.3954643328E-10    0.5015007827E-01
   -0.5139403942E-01    0.4079144595E-09   -0.5571562347E-01
    0.6400141536E-01   -0.5244803832E-10    0.7420862240E-01
   -0.1217786016E+00    0.1209486864E+00    0.3357360460E-01
   -0.7046772982E-02    0.2323484398E-01    0.7471911850E-02
    0.1194377085E+00    0.1188381201E+00   -0.2580124756E+00
   -0.6034560022E-01   -0.2740164517E-02    0.4730794331E-02
    0.1370561097E-02   -0.1239870550E+00    0.2602347003E+00
    0.2656729821E-01   -0.4111474045E-01   -0.5719017164E-01
   -0.2476619352E-02    0.7289448018E-02    0.7454326008E-02
   -0.3177909150E-01    0.5614667348E-01    0.7269179063E-01
   -0.1217786012E+00   -0.1209486864E+00    0.3357360411E-01
   -0.7046773192E-02   -0.2323484397E-01    0.7471912093E-02
    0.1027445066E-01    0.1297117056E-01   -0.5645164448E-02
    0.1194377085E+00   -0.1188381201E+00   -0.2580124758E+00
    0.6034560003E-01    0.2740164570E-02    0.4730794300E-02
   -0.1370561162E-02   -0.1297117052E-01   -0.1286299711E-01
    0.5950210916E-02    0.1239870550E+00    0.2602347003E+00
    0.2656729844E-01    0.4111474097E-01   -0.5719017204E-01
   -0.2476619444E-02   -0.7289448007E-02    0.7454326137E-02
   -0.5645164546E-02   -0.5950210891E-02    0.6684117355E-02
   -0.3177909145E-01   -0.5614667346E-01    0.7269179058E-01
```

(2) Atom specifications

(3) Subset Hessian

Figure 2. CUTOFF output generated from the file in Figure 1.

1. CUTOFF output begins by printing the title line exactly as input by the user.

2. The isotopic element symbol and Cartesian coordinates for the atoms in the subset are printed, each on a separate line.

3. The lower triangle of the subset Hessian is printed out in the same format as in the input. The rows and columns present in the full square symmetric Hessian matrix correspond to the *x*, *y* and *z* coordinates of the subset atoms.

CUTOFF Troubleshooting guide

There are a number of errors that users can come across, which will hopefully decrease in number as they familiarise themselves with the code. The most common of these are input errors. Luckily, the program will inform the user within the UNIX environment of the last read format and line number in the fortran source code, and therefore a likely area linking to the error. Throughout this manual, the formats of input have been specified to allow the user easy identification of the areas which require specific types. This should be the first reference point in the case of incomplete termination.

One error that has been noted is the format of certain coordinate/Hessian outputs from electronic structure codes. Some codes simply print out the coordinates, followed directly by the Hessian. However, sometimes the gradient is printed following the coordinates and preceding the Hessian. This consists of *n* lines of a similar format to the Hessian, where *n* is the number of atoms in the structure. This is often noticed upon processing, where erroneous frequencies are printed from the program and the input needs to be re-inspected. The simplest solution is to use the equation in Part 4 of the *Overview of CUTOFF input* to deduce the expected number of lines of Hessian, and a program such as Notepad++ with a line count to identify if lines of gradient are present. When located, these can simply be deleted from the input and the programs rerun, which should fix the errors.