IGMPlot
Version 2.6.7 – June 2020
Documentation

A CPU implementation of the IGM approach
using either promolecular density or
from wave-function calculations
for detecting, quantifying and plotting interactions
from electron density

by Johanna Klein, Emmanuel Pluot, Gaëtan Rubez, Jean-Charles Boisson\textsuperscript{a} and Eric Hénon\textsuperscript{b}


Contributors
Miguel Ponce-Vargas, Hassan Khartabil, Corentin Lefebvre, Julia Contreras-García

Referencing IGMPlot

The IGM methodology basis using promolecular electron density
\textsuperscript{1}Lefebvre C., Rubez G., Khartabil H., Boisson J.C., Contreras-García J., Hénon E.
doi: http://doi.org/10.1039/c7cp02110k

The quantum-mechanical theory of IGM
\textsuperscript{2}Lefebvre C., Khartabil H., Boisson J.C., Contreras-García J., Piquemal J.-P., Hénon E.
Chem. Phys. Chem. 2018, 19, 1
doi: http://doi.org/10.1002/cphc.201701325

The atomic decomposition scheme of the interaction
\textsuperscript{3}Miguel Ponce-Vargas, Corentin Lefebvre, Jean-Charles Boisson, Eric Hénon
J. Chem. Inf. Model. 2020, 60, 1, 268
doi: http://doi.org/10.1021/acs.jcim.9b01016

The Intrinsic Bond Strength Index (IBSI)
\textsuperscript{4}Johanna Klein, Hassan Khartabil, Jean-Charles Boisson, Julia Contreras-García, Jean-Philip Piquemal, Eric Hénon
J. Phys. Chem. A 2020, 124, 9, 1850
doi: http://doi.org/10.1021/acs.jpca.9b09845

Bond Asymmetry Factor (BAF)
in preparation

\textsuperscript{a}CReSTIC EA 3804, University of Reims Champagne-Ardenne, 51687 Reims, France
\textsuperscript{b}Institut de Chimie Moléculaire de Reims, CNRS UMR 7312, University of Reims Champagne-Ardenne, 51687 Reims, France
**IGMPlot** is a free software. It is released under the CeCILL-c license [http://cecill.info/index.en.html](http://cecill.info/index.en.html)
Table des matières

I. Introduction ..................................................................................................................................................4

II. Installation..................................................................................................................................................10

III. Running the program ................................................................................................................................11

IV. Inputs - Outputs ........................................................................................................................................12

V. KEYWORD documentation ..........................................................................................................................13

VI. Examples........................................................................................................................................................18

   Example 1 (promolecular density): to use one or two fragments? ([test1],[test2]) ........................................18
   Example 2 (promolecular density): vdw interactions ([test3]) ....................................................................20
   Example 3 (promolecular density): quantification of non-covalent interactions ([test4 to test10]) ............21
   Example 4 (promolecular density): Monitoring intramolecular interactions with IGM-δg_intra in a peptide ([test11],[test12]) ...............................................................25
   Example 5 (promolecular density): smart use of the uncoupling scheme in a trimer ([test13]) .................26
   Example 6 (promolecular density): ligand – protein ([test14]) ..................................................................27
   Example 7 (promolecular density): when to use the CUTPLOT_IGM keyword? ([test15]) ......................28
   Example 8 (promolecular density): atomic decomposition scheme of non-covalent interactions applied to host-guest assemblies ([test16, test17]) ..................................................................................29
   Example 9 (QM): H₂O ... H₂O dimer, FRAG1 keyword ([test18]) ..........................................................30
   Example 10 (QM): B₂H₆ taken as one piece ([test19]) .............................................................................31
   Example 11 (QM): Agostic interaction in the metallic complex [Ti Cl₂C₆H₅]⁺ ([test20]) ............................32
   Example 12 (QM): FRAG1 and FRAG2 keywords: quantifying intramolecular π-stacking ([test21]) ....33
   Example 13 (QM): CUBEFRAG keyword ([test22]) ..............................................................................34
   Example 14 (QM): IBSI keyword: a new way for probing bond strength ([test23]) ...............................35
   Example 15 (QM): IBSI keyword with cutoff ([test24]) .........................................................................37
   Example 16 (QM): Bond δg_pair signature and atom pair isosurface ([test25]) ..........................................37
   Example 17 (QM): Bond Asymmetry Factor ([test26 to test29]) ............................................................38

VII. Future prospects and outlook ................................................................................................................43
The IGMPlot package contains the code source together with sample runs, this documentation, the software reference documentation (html), installation instructions, and Copyright.

I. Introduction

By using IGMPlot you can identify and quantify molecular interactions over a broad range: from non-covalent to covalent bonding. It allows you for:

- Studying molecular system from a wave function (QM treatment)
- Studying large systems like ligand-protein complexes from the atomic cartesian coordinates only (promolecular ED)
- Distinguishing between intramolecular and intermolecular interactions, covalent or hydrogen-bonding or vdW interactions
- Probing and quantifying interactions between two given fragments in a molecular system (both QM and promolecular modes)
- Determining the strength of a given bond and its polarization
- Estimating the atomic contributions to an intermolecular interaction
- Building isosurfaces representing the region of space where the interactions take place
- Monitor interactions along molecular dynamics or along a reaction path

From a practical perspective, an attractive feature of the IGM approach is to provide an automatic workflow delivering data that provides chemists with a visual and quantitative understanding of interactions. Although IGMPlot relies on the ED topology, no topological analysis is required and the IGM analysis can be achieved with little preparation.

History

IGMPlot is based on the electron density-based descriptor called $\delta_g$. The IGM-$\delta_g$ approach was initially designed to work with promolecular density (non-relaxed electron density, sum of spherically averaged neutral atomic densities). But in 2018, we proposed the Gradient-Based Partitioning (GBP) that extends the IGM concept to electron density derived from a wave function (SCF). Thus, thanks to the new IGM approach, detailed information can be directly obtained either on the covalent or on the non-covalent domain, for small and larger molecular systems. It is worth noticing that the promolecular version of IGM is also implemented in the Multiwfn software (http://sobereva.com/multiwfn).

In 2020, a new extension was developed able to emphasize the most relevant atomic contributions to the noncovalent interactions occurring between two fragments. This development proves to be an appealing tool to shed light on the guest accommodation on a per atom basis. Also, in 2020 was proposed the new Intrinsic Bond Strength Index IBSI. This score, (not a bond order), is obtained in an automatic workflow and is very efficient to internally probe the strength of a given atom pair, over a wide range (non-covalent to covalent, transition metal bonding, agostic interactions, …). An IBSI scale has been proposed to range two-centre chemical bonds by their intrinsic strength.

In 2020 was proposed a new release allowing for detecting the interaction between two given sub-fragments of a single molecule, using QM electron density. This possibility is particularly attractive to assess the role of non-covalent intramolecular interactions (for instance intramolecular $\pi-\pi$ stacking or hydrogen-bonding between two part of a single molecule along a reaction path). Also, the quantification of the detected interactions has been implemented using integration schemes and summing local quantities over the space representing the interactions. From a technical perspective, the .wfx file format is now supported by IGMPlot (in addition to .wfn).

Presentation

For more detailed information we refer the interested reader to the original papers (above-mentioned).

The IGM-$\delta_g$ approach is based on a new electron density (ED) reference model, the independent gradient model (IGM): really getting rid of interactions, which was presented in 2017. Compared to an interacting situation, this
non-interacting reference enables quantifying the net drops observed in NCI’s 2D plots both in region of high and low ED.

** IGM principle in a nutshell **

The \( \delta_{\text{IGM}} \) approach relies on the new definition of a non-interacting system in term of ED contragradience ("ED clash" between two sources). More precisely, the new descriptor, \( \delta \), represents locally the difference between a virtual upper limit of the ED gradient (\( \nabla p_{\text{IGM}} \)) representing a non-interacting system and the true ED gradient (\( \nabla p \)). The theoretical background can be found in the papers 1 and 2. The program IGMPlot, written in C++, is a custom implementation using either promolecular density or relaxed electron density (SCF). It is designed to leverage the computational power of Multi-core CPU through parallel OpenMP programming.

** Automatically separating intramolecular from intermolecular interactions **

An attractive feature of the IGM methodology is to provide an uncoupling scheme intra/inter that automatically extracts the signature of intra- (\( \delta_{\text{intra}} \)) and inter-(\( \delta_{\text{inter}} \)) molecular interactions (see Fig. 1, water dimer) for drawing the corresponding 3D iso-surface representations in real space with software like VMD. This automatic intra/inter separation can be carried out either using promolecular ED or ED coming from QM calculations (wave function mode). The user has only to supply the definition of two fragments in terms of atom indexes.

The sign of the second eigenvalue of the ED hessian matrix serves to differentiate repulsive (\( \lambda_2 > 0 \)) from attractive (\( \lambda_2 < 0 \)) interactions. The nature of the interaction is color-coded: blue for attractive interactions, green for weakly repulsive or attractive interactions or red for repulsive situations.

The use of promolecular ED is particularly suited to describe ligand-protein interactions while relaxed ED (SCF) obtained from wave function calculations extends the range of applicability to organic chemistry, inorganic, chemical reactivity, ... Of course, the study of strong interactions (like covalent bonding) requires the use of relaxed ED obtained by means of QM calculations. So far, promolecular electron density computation is limited to atoms of periods 1,2,3,4. Concerning the wave-function mode, the needed wave function file can be obtained from a Gaussian run for example. IGMPlot supports both the .wfn and .wfx input file formats.

![Fig. 1 Water dimer treated with promolecular electron density (dedicated to weak interactions); NCI plot on top left panel (global signature); intramolecular IGM-\( \delta_{\text{intra}} \) signature on top right panel; automatically extracted hydrogen-bonding IGM-\( \delta_{\text{inter}} \) signature at bottom left; isosurfaces bottom right](image)

**Fig. 1** Water dimer treated with promolecular electron density (dedicated to weak interactions); NCI plot on top left panel (global signature); intramolecular IGM-\( \delta_{\text{intra}} \) signature on top right panel; automatically extracted hydrogen-bonding IGM-\( \delta_{\text{inter}} \) signature at bottom left; isosurfaces bottom right
The detection of intermolecular interaction regions takes place automatically from the moment the two fragments definition (atom indexes) has been supplied. Since the intra/inter interactions are well separated by the IGM-δg approach, the program is then able to deliver intermolecular interactions (generally weak interactions) deprived of intramolecular interactions (covalent bonding, but also weak intramolecular interactions like intramol. H-bond or ring closure interactions) without any specific manipulation in the input parameters. Both in the “wave function” and “promolecular” modes, it is not necessary to choose appropriate ED cutoffs in a way that produces well-separated strong/weak interaction domains. But this option is still available.

** “wave function” or “promolecular mode” **

What is the difference between the “promolecular” and “wave function” modes? The promolecular electron density (ED) is the sum of simple exponential atomic functions. These exponential functions are tabulated and stored in the IGMPlot program for atoms of periods 1-4. It lacks ED relaxation, but this approximate density has however shown to provide similar results to the relaxed one as long as computations remain in the noncovalent domain. Thus, promolecular densities are useful to study large systems, typically ligand–receptor interactions, because its calculation is very fast, and it only require the geometry as input (the ED is calculated by IGMPlot). Similar topological features can be obtained from both promolecular and quantum EDs, as can be observed from Fig. 2 showing the interaction between a trinuclear copper complex and a fullerene molecule.

![Fig. 2 Trinuclear copper complex; comparison of the “promolecular” and “wave function” (M06-2X/6-311G(d)) modes to obtain the intermolecular interaction signature](image)

Of course, the promolecular ED cannot be employed to study strong interactions (covalent bonding, metal coordination). As can be seen on Fig. 3 for the water dimer, the promolecular ED clearly underestimates the covalent O-H interaction (left panel): δg_{intra} (promol.) = 0.4 versus δg_{intra} (QM) = 0.9 for the two covalent peaks. Conversely, both levels of theory give δg_{inter} = 0.06 for the intermolecular signature peak height (on the right).
The promolecular ED is limited to the 36 first elements of the periodic table (H to Kr). In this case, only the atomic coordinates have to be supplied. The promolecular density is obtained from a sum over simple atomic exponential functions. It lacks SCF relaxation. The parameterized exponents and coefficients of sphercialized atomic densities for atoms of periods 1, 2 and 3 have been taken from the program NCIPlot. These densities are developed over a sum of 3 exponential atomic pieces. For H and He, only the first coefficient has a non-zero value. For the atoms of period 2 the third coefficient is set to 0. In order to provide such an analytical function for atoms of period 4, without any additional CPU cost, an expression of only 3 exponential atomic pieces has been tailored to match the neutral averaged spherical atomic ED (DFT B3LYP/6-311+G*) for ED values below 0.1 a.u. This was done to describe atoms of period 4 (such as Br) that can be involved in non-covalent interactions within ligand-protein complexes for instance. However, let’s recall that such promolecular density is able to predict low-density only; i.e., to describe the non-covalent domain. It cannot be used to describe covalent (nor metal coordination) situations.

**Interaction strength**

The $\delta_{\text{inter}}$ descriptor is not dimensionless, linking its value to the familiar concept of interaction strength as illustrated on Fig. 4. The hydrogen-bond spike emerges at a $\delta_{\text{inter}}$ value (around 0.065 a.u.) significantly greater than for van der Waals contacts (0.025 a.u.).

In the series of 4 dimers (Fig. 5), we observe that the peak heights are related to the dimer stabilization energies obtained from CCSD(T) ab initio calculations (see Table below). The values fit relatively well to a linear correlation for $\delta_{\text{inter}}$. 
The IGM reference gives a unique definition of the interaction region: non-zero values of $\delta_{\text{inter}}$ exclusively correspond to interaction situations. As a consequence, the IGM approach overcomes difficulties to define regions within integration procedures to quantify interactions. An integration scheme has been proposed for the first time in this release to quantify intermolecular interactions (see the IGMPlot output on Fig. 6). It is explained in more details in the example section.

**Fig. 5** Correlation between $\delta_{\text{inter}}$ taken at the peak and hydrogen-bond strength

<table>
<thead>
<tr>
<th>Complex</th>
<th>$d$ (Å)</th>
<th>$\rho$</th>
<th>$\delta_{\text{inter}}$</th>
<th>$\Delta E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FH$_2$...H$_2$O</td>
<td>2.65</td>
<td>0.016</td>
<td>0.027</td>
<td>$-2.5^a$</td>
</tr>
<tr>
<td>H$_2$O...H$_2$O</td>
<td>1.93</td>
<td>0.032</td>
<td>0.069</td>
<td>$-5.0^b$</td>
</tr>
<tr>
<td>HCN...HF</td>
<td>1.84</td>
<td>0.044</td>
<td>0.090</td>
<td>$-7.5^b$</td>
</tr>
<tr>
<td>HF...NH$_3$</td>
<td>1.68</td>
<td>0.063</td>
<td>0.121</td>
<td>$-12.6^b$</td>
</tr>
</tbody>
</table>

$^a$ CCSD(T)-F12a/VQZ-F12, $^b$ CCSD(T)/CBS + REL + CV.

**Fig. 6** IGMPlot output in the case of the water dimer using the $\delta_{\text{inter}}$ partition scheme and promolecular ED
The development of an even more accurate energy function is under progress, using machine learning possibilities.

Also, the integration of \( \delta_{\text{g intra}} \) representing the intramolecular interactions present in fragment(s) is reported.

**Atomic decomposition scheme**

In the “promolecular” mode, a compelling feature of the IGMPlot program is to carry out an atomic decomposition scheme in order to estimate the influence of a given atom in the intermolecular region between two user-selected fragments.

When an intermolecular interaction occurs, a new \( \delta_{\text{g atom}} \) index has been derived from the IGM approach. Its value quantifies locally the contribution of atom \( \text{At} \) to the interaction between the two fragments \( \text{A} \) and \( \text{B} \). It can be integrated over the IGM grid to get the associated relative atomic contribution, noted \( \Delta g_{\text{inter,At}} \).

Atoms are then colored according to their \( \Delta g_{\text{inter,At}} \) score (expressed as a percentage) to obtain a coupled analysis: isosurfaces/atomic participation, providing a rapid and rich picture of their role in the formation of the host-guest assembly. A VMD visualization state session \texttt{atContr.vmd} input file is automatically prepared to run vmd program and to get these pictures.

![Fig. 7 Ligand – Protein example; atomic contribution to intermolecular interaction](image)

This complementary analysis brings an additional quantitative side to the initial IGM analysis. This possibility is illustrated on Fig. 7 where we have reported an analysis of non-covalent interactions between a kinase protein and a ligand. The atoms of the complex are colored according to their contribution to the iso-surfaces (summing \( \Delta g_{\text{inter,At}} \) over the grid and using a Blue-to-Red color scheme, with blue: no contribution to the interaction, red: largest relative contribution to the interaction).

**IBSI**

The IBSI (Intrinsic Bond Strength Index) is a score very efficient to internally probe the strength of a given atom pair in molecular situation. It stems from the integration of the IGM-\( \delta_{\text{g pair}} \) signature obtained for a given atom pair using the new bond descriptor \( \delta_{\text{g pair}} \) (see examples of such signatures on Fig. 8).
Fig. 8 IGM is able to focus on a selected atom pair

Based on a large set of compounds (235 species, 677 atom pairs), we suggest to distinguish between weak interactions, coordinate covalent and covalent bonds according to this scale (the limits being only indicative on Fig. 9):

![IBSI scale for indicative purposes](image)

The IBSI does not belong to the conventional class of bond order (like Mulliken, Wiberg or Mayer, giving the number of electron pairs shared between two atoms), but rather assesses the intrinsic bond strength.

The IBSI is defined in a parameter-free manner and its implementation is very easy with IGMPlot: it only requires a wave function file as input. It has been shown little dependent on the level of theory and basis set.

In addition to this score, the spatial representation of $\delta g_{\text{pair}}$ isosurfaces reflect the common chemical concept of bond in a clear and intuitive manner, being distinctly different for single or double bonds. It paves the way for probing specific interactions like agostic interactions in transition metal compounds, three-centers two electrons bond, but also for targeted mechanistic exploration of reactions by monitoring selected covalent pair-interactions.

II. Installation

IGMPlot is written in C++. It has been installed and tested on several platforms: computational centers (linux), MacOS, Windows10, and several compilers and versions (GNU, Intel, PGI). It is a standalone program. Default is to use the OpenMP programming interface. OpenMP has to be installed prior to IGMPlot for OpenMP compilation.

- **General recommendations** (Linux, MacOS, Windows):
  1. For OpenMP compilation, first, make sure having openMP installed on your machine (http://www.openmp.org).
  2. Uncompress the IGMPlot archive: `tar -xvjf IGMPLT-x.y.tbz2`. Change to directory `IGMPLOT-x.y/source`.
  3. Edit the Makefile and select:
     - the family of your C++ compiler (CppCompilerFamily): three choices are proposed (GNU, INTEL, PORTLAND)
     - the version of your C++ compiler (CppCompilerVersion)
     - if you want to use OpenMP or not (OpenMP): the use of openMP can be disabled in the file ‘Makefile’ by (1) commenting out the line ‘OpenMP=YES’ while (2) uncommenting #OpenMP=NO’.
• the path to your compiler (including its name)

• Make sure having the appropriate compiler environment. Type ‘make’ in a terminal. The executable name is ‘IGMPlOt’. Make sure that this new command can be reached (adjust your PATH variable).

• On MacOs machines, a sequential version of IGMPloT can be obtained with the Clang compiler. In the Makefile choose the options:
  1. CppCompilerFamily=GNU
  2. CppCompilerVersion=5_and_above
  3. OpenMP=NO
  4. CC=g++

• On MacOS machines, to leverage OpenMP multicore execution, you must install a gcc (g++) version different from the one provided within the compiler front end “Clang” which until now has not built-in support for OpenMP. You might install gcc with the command: brew install gcc (or brew upgrade gcc). Prior to that, make sure to have installed the latest Xcode and Command Line Tools for Xcode. This way, the g++ compiler will be installed somewhere like /usr/local/Cellar/gcc/7.1.0/bin/g++-9. In this example, make sure the g++-9 command be available with your PATH and adjust the IGMPloT makefile accordingly:
  o CppCompilerFamily=GNU
  o CppCompilerVersion=5_and_above
  o OpenMP=NO
  o changing the g++ command with g++-9 (CC=g++-9 for instance).

• Windows: only the Windows10 version has been successfully tested. The linux bash shell on Windows10 has been employed. Indeed, a full Ubuntu-based Bash shell that can run Linux software directly on Windows is now available on Windows 10. It isn't a virtual machine. It allows you to run the Bash shell and the exact same binaries you would normally run on Ubuntu Linux. After having enabled the “Windows Subsystem for Linux (Beta)”, and installed a gcc developing environment (gcc, makefile), just follow the above-mentioned recommendations (and finally, type ‘make’).

IGMPloT is supplemented with the small shell script “vmdpath”. This shell converts .vmd visualization states produced by IGMPloT (igm.vmd, nci.vmd, atContr.vmd) to a new local.vmd file by automatically adjusting file paths to current path (were cube files have been stored). This is useful since, in the visualization state files (.vmd), absolute file paths are used (which can differ from the real ones after downloading the result from a distant computational center for instance).

III. Running the program

IGMPloT needs a parameter input file param.igm. In addition to specify the filename describing the molecule (xyz, or wfn or wfx), it contains specific IGMPloT keywords. The list of keywords is described hereafter. In addition, example jobs are made available in directory ‘tests’. Make use of the run.sh shell to execute them sequentially.

To execute IGMPloT, in a terminal, type “IGMPloT param.igm” to run the job. In the OpenMP mode the default is to use the maximum of CPU cores available on the machine. This can be changed by setting the OMP_NUM_THREADS variable. For instance (linux, MacOs): export OMP_NUM_THREADS=8.

Note that now, a “runinfo” file is generated after a while and regularly updated to display timing information (Fig. 10).
IV. Inputs - Outputs

The IGM-δg approach is based on the possibility to define two fragments whose interaction is studied (even though intramolecular interactions will be also delivered).

Two kinds of files are needed:
- the parameter input file \textit{param.igm}
- either:
  - QM mode: one single file to specify both the geometry and the wave function (\textit{.wfn} or \textit{.wfx}),
  - Promolecular mode:
    - Two .xyz files to specify the geometry of the two fragments
    - One .xyz file: the system will be considered as a whole (δg\textit{inter}=0)

In the QM mode, the definition of the two fragments will be achieved through keywords (\textit{FRAG1}, \textit{FRAG2} or \textit{CUBEFRAG}). Otherwise, the system will be considered as a whole (δg\textit{inter}=0).

\textit{IGMPlot} proceeds with inputs and outputs in a very simple way. Once the fragment definition is provided in the \textit{param.igm} input file, \textit{IGMPlot} only requires a very limited set of keywords in the standard case.

\textbf{Input}

- \textit{from a wave function file:} in the standard case two input files are needed: one \textit{wfn} or \textit{wfx} file describing the wave function (and the geometry) and in addition, a ‘igm’ parameter file containing the keywords. The definition of the two fragments to be studied can be set is set through the keyword \textit{FRAG1}, \textit{FRAG2} or \textit{CUBEFRAG} (see below).

- \textit{with promolecular ED:} in the standard case three input files are needed: two .xyz file describing the cartesian coordinates of the two interacting partners; in addition, a ‘igm’ file containing the program keywords is needed. It is also possible to use the program to study non-covalent interactions within one single molecule: one single .xyz input file in that case. The electron density (ED) is then calculated directly by \textit{IGMPlot} using the promolecular approximation.

\textbf{WFN/WFX remarks:}

- the WFN file generated from the Gaussian program does not support ECP basis properly (the ECP data is treated as a second “basis set”, and it is necessary to remove it manually from the wfn file so as to describe valence electrons only). You should prefer in that case the Output=WFX option in Gaussian.

- Note however that, upon using a WFX file, \textit{IGMPlot} does not make use of the additional core density function data to represent the electron density of ECP-modeled core electrons (when pseudo-potentials are used). According to the AIMAll recommendations (http://aim.tkgristmill.com), for an atom for which a "small-core" or "medium-core" ECP is used, ignoring this core ED will not affect significantly the topology of the ED in the valence region and therefore the IGM-δg analysis can be performed. For an atom for which a "large-core" ECP is used, modelling the core ED cannot be ignored and the use of \textit{IGMPlot} is not yet adapted to this situation.

- IGM-δg\textit{inter} calculations have been successfully carried out using semi-empirical methods like PM7.
• Restricted and open shell IGM calculations have been successfully tested; excited state IGM calculations remain to be tested.

**param.igm** constrained format:

- The first line must contain the number \( n \) of files describing the molecular system. Currently, \( n \) is in the range \([1-2]\). \( n=1 \) if you intend to use a wave function file (wfn or wfx), \( n=2 \) if you intend to use promolecular density (xyz mode). According to the filename extension supplied (xyz or .wfn or .wfx) \textsc{igmplot} will appropriately proceed with a promolecular or wave function calculation.
- The second line must contain the name of the first .xyz coordinate file (or of the wfn file or wfx file).
- If 2 files have to be read (promolecular mode), the second and third lines must contain the two names of the first and second .xyz files.

**Example**, minimum set of input:

```
2
ligand.xyz
protein.xyz
```

The rest of the keywords is optional.

**Output**

Output files are generated within the IGM approach (details are given hereafter in the keyword section OUTPUT).

**V. KEYWORD documentation**

Examples given below can be found in the “test” folder provided with this release.

Note that, in addition to the IGM calculations, ‘NCI’ calculations (RDG descriptor) are also performed within an \textsc{igmplot} run.

Two “modes” are made available in \textsc{igmplot}: (1) using promolecular ED or (2) using ED coming from a single wave function file. \textsc{igmplot} will detect the running mode according to the file extension of the input file(s) you will provide: .xyz (promolecular ED calculated by \textsc{igmplot}) or .wfn/.wfx (quantum calculations).

A set of options is available. The corresponding keywords are described below.

The default IGM grid is rectangular cuboid (Fig. 11). It permits the generation of \( \delta g \) isosurfaces.

**Fig. 11** The rectangular grid type is used for every calculation except for bond studies (IBS, BAF) where an ultrafine cylindrical grid is automatically determined by \textsc{igmplot} for each specified bond.
Optional keywords:

Keywords marked with an asterisk * are to be used only within the promolecular mode.

There are three possibilities to position the IGM rectangular grid, with LIGAND, CUBE or RADIUS keywords.

If none of these 3 keywords is given the default is to build this grid based on coordinates of the atoms (given in Angströms). Then, in QM mode (wfn or wfx), the whole system will be considered to determine the grid encompassing all the atoms. The same thing with promolecular mode using only one single .xyz file. Within the promolecular mode using two .xyz coordinate files, the default is to position the grid such as to encompass the first molecule specified in the param.igm input file. In every case, in order to ensure a correct grid in planar or linear cases, a minimum default buffer distance of +/- 3 Å is added to the axes in all directions.

*LIGAND n r
Promolecular mode only. Only when two .xyz files are provided.

- n is the index of the xyz file used to position the grid. The order of introduction in the .igm file: 1 or 2, is used to designate the molecule encompassed by the grid.
- r is a buffer distance in Å (taken around the molecule used to define the calculation grid)

Generally, within protein-ligand systems, the ligand is used as reference to locate the grid because it is the smallest molecule. More generally, it is recommended to select the smallest molecule to lighten the calculation.

After reading the atomic positions the program determines the dimension of the grid increased by a “r” Å buffer that encompasses one of the molecules. Default index is 1 and default buffer is 3 Å.

CUBE x1 y1 z1 x2 y2 z2
6 real numbers (defining two points in real space)

This option is used to specifically set the IGM 3D rectangular grid by using the two extremities of a cube diagonal (x1,y1,z1) to (x2,y2,z2), all in Å.

RADIUS x y z r
x y z determine the position around which interactions are represented for a radius r (all in Å)

The IGM 3D rectangular grid will be determined to encompass the given sphere of radius r and center (x,y,z).

ONAME name
name stands for the prefix be passed to the output file names (default is ‘out’).

OUTPUT n
- n is an integer in the range 1 – 5; default is 5.

- 1 will print 2 files:
  - nci.dat: 2-columns file (signed ED and RDG)
  - igm.dat: 4-columns file (signed ED, RDG, δg_intra, δg_inter)

- 2 will print two .cube files (NCI mode only):
  - dens.cube: signed ED
  - RDG.cube: reduced density gradient
✓ 3 will print three files (IGM mode):
  ▪ dens.cube: signed ED
  ▪ dgInter.cube: $\delta g_{\text{inter}}$ values
  ▪ dgIntra.cube: $\delta g_{\text{intra}}$ values

✓ 4 will print three files:
  ▪ atContr.vmd: vmd session file to display atomic contributions to interaction
  ▪ complex.xyz: concatenation of two fragment xyz coordinate files
  ▪ percent.dat: 1-column file

This possibility is only available in the promolecular mode. A compelling feature of the IGM model is to enable designing other references where interactions partially cancel, for instance interactions due to a single given atom. This gives rise to an estimate of the relative contribution of atoms in the non-covalent interaction regions obtained by summing each descriptor $\delta g$ over the grid (no distinction is made between attractive and repulsive points). Thanks to IGM reference, this confers a quantitative aspect to the analysis. The atoms of the complex are colored according to their contribution to the $\delta g_{\text{inter}}$ interaction iso-surfaces (the atomic contributions are stored in file ‘percent.dat’ and read from the vmd session file ‘atContr.vmd’).

✓ 5 will print 11 files: full output
  ▪ nci.dat
  ▪ igm.dat
  ▪ dens.cube
  ▪ RDG.cube
  ▪ nci.vmd: vmd session file to display RDG iso-surfaces
  ▪ dgInter.cube
  ▪ dgIntra.cube
  ▪ igm.vmd: vmd session file to display separately $\delta g_{\text{inter}}$ and $\delta g_{\text{intra}}$ iso-surfaces
  ▪ atContr.vmd: vmd session file to display atom contributions to $\delta g_{\text{inter}}$
  ▪ complex.xyz
  ▪ percent.dat (atom contributions to $\delta g_{\text{inter}}$)

**INCREMENTS** $i_1$ $i_2$ $i_3$
  ▪ 3 real numbers

This option sets the increments (Å) taken along the x, y, z directions of the IGM rectangular grid. The default is set to 0.1, 0.1, 0.1 Å.

**FRAG1 selection pattern** (IGM, WFN/WFX mode)
QM mode only. Enables the $\delta g_{\text{inter}}$ analysis.
  ▪ a selection pattern to specify the atoms belonging to fragment 1 (see examples below)
Example: FRAG1 2-5;7 --&gt; atoms 2,3,4,5,7 form the fragment 1.

**FRAG2 selection pattern** (IGM, WFN/WFX mode)
QM mode only. Enables the $\delta g_{\text{inter}}$ analysis.
  ▪ a selection pattern to specify the atoms belonging to fragment 2 (see examples below)
Example: FRAG2 1;6;8-10 --&gt; atoms 1,6,8,9,10 form the fragment 2. The remainder forms the fragment 2.
When only FRAG1 (FRAG2) is specified, the remaining atoms in the system form the fragment 2 (fragment 1).

Example, for a 10-atoms system: FRAG2 1;6;8-10 -- Fragment 2 = atoms 1,6,8,9,10 | Fragment 1 = atoms 2,3,4,5,7

When both FRAG1 and FRAG2 are specified, then, only the atomic contributions of the sub-system made of FRAG1+FRAG2 are considered. This option is particularly useful to study the interaction between two small fragments of a very large molecular system. This option speeds up the calculation. Illustrative examples are given below.

### CUBEFRAG (IGM, WFN/WFX mode)
QM mode only. Requires CUBE keyword to be also defined.

- Only atoms located in the rectangular grid defined by the CUBE parameters will be considered in the IGM analysis.

This option is particularly useful to study interactions within a given small region of a very large system. Then, the calculation is speeded up. Illustrative examples are given below.

---

#### FRAG DEFINITION SUMMARY

Five situations are considered: no FRAG definition, FRAG1 solely defined, FRAG2 solely defined, FRAG1+FRAG2 defined, CUBEFRAG definition supplied:

<table>
<thead>
<tr>
<th>CASE1</th>
<th>CASE2</th>
<th>CASE3</th>
<th>CASE4</th>
<th>CASE5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword</td>
<td>no FRAG</td>
<td>FRAG1</td>
<td>FRAG2</td>
<td>FRAG1+FRAG2</td>
</tr>
<tr>
<td>Fragment 1</td>
<td>All atoms</td>
<td>FRAG1 selection</td>
<td>remainder</td>
<td>FRAG1 selection</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>0</td>
<td>remainder</td>
<td>FRAG2 selection</td>
<td>FRAG2 selection</td>
</tr>
<tr>
<td>Primitives considered</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Primitives on Frag1 &amp; Frag2 atoms</td>
</tr>
</tbody>
</table>

### IBSI $r$ (IGM, WFN/WFX mode)
QM mode only. Requires ENDIBSI keyword to be also defined at the end of the IBSI section.

- $r$ : optional, designates a cut-off radius (in Å); primitives beyond $r$ Å of the given atom pair are not considered in the IBSI calculation. $r$ can be omitted (full calculation in that case).

Example:

```
IBSI 4.0
1 2
12 87
6 3
ENDIBSI
```

This IBSI option will calculate the intrinsic bond strength index (IBSI) for the atom pairs: 1-2, 12-87, 6-3, using primitives within a distance cut-off of 4.0 Å of the atom pair considered. Each IBSI index is calculated
for a given A-B atom pair from equation 1 of paper:

\[ \frac{\int_V \delta g_{\text{pair}} \, dV}{\int_V \delta g_{\text{pair}}^2 \, dV} = IBSI \]

\( \delta g_{\text{pair}} \) is the descriptor quantifying the ED contragradience between atomic sources A and B. d is the internuclear distance between A and B. IBSI has been normalized to 1 for H\(_2\) at the M06-2X/6-31G** level of theory. IBSI has been shown to be little dependent on the level of theory, but we advise not using the HF method nor the STO-3G basis set. We recommend the DFT M06-2X/6-31G** level of theory for its performance/price ratio and using the same method/basis set for comparative studies. Adding diffuse functions can be relevant for weakly bonded systems. In extreme cases, in very polar bonds like LiH, AlO, AlCl, MgC, … a larger dependence can be observed.

Unless r is specified, the calculation of IBSI for a given atom pair accounts for all the primitives of the molecular system. Otherwise, only primitives located within r Å of the atom pair will be accounted for. This option is particularly useful to speed up the calculation for large systems. Actually, an ultrafinite cylindrical grid is built around each bond in this calculation, which can be quite expensive.

In addition to IBSI, a Bond Asymmetry Factor (BAF) is provided which describes the ED contragradience asymmetry in-between the given atom pair.

**INTERMOLECULAR** (NCI model)
Promolecular mode only
- r in the range 0-1. For good results, r should be 0.8-0.9.

In the original NCIPlot code it is possible to discard the grid nodes for which more than a fraction (default threshold value is 0.95) of the total promolecular density comes from only one molecule (A or B). Typically, this turns off the intramolecular interactions in the resulting files for NCI calculations only. IGM result files won’t be affected by this parameter.

**CUTOFFS** r\(_1\) r\(_2\) (NCI model)
- 2 real numbers

Defines the density (r\(_1\)) and RDG (r\(_2\)) cutoffs used to print .dat files:
- .dat, \(\text{sign}(\lambda_2)\rho\) range: \([-r_1 \ r_1]\): impacts both nci.dat and igm.dat
- .dat, RDG range: \([0 \ r_2]\): impacts nci.dat

r\(_1\) default is 0.3 u.a. r\(_2\) default is 10.0 u.a.

**CUTPLOT** r\(_3\) r\(_4\) (NCI model)
- 2 real numbers; only concerns the NCI method (impacting both cube and vmd sessions files)

r\(_3\): density threshold, the unique criterion used to filter values written to the RDG cube file. Points with \(\text{sign}(\lambda_2)\rho\) in the range \([-r_3 \ r_3]\) will be stored in the cube file.

r\(_4\): the RDG iso-value used in the nci.vmd session script generated by IGMPlot; RDG iso-surfaces will be colored according to a BGR scheme over the ED range \(-r_4 < \text{sign}(\lambda_2)\rho < r_4\) a.u.

This option might be useful to display 3D iso-surfaces coming from a specific peak-region previously identified on the 2D plot. Default values are: r\(_3\)=0.3, r\(_4\)=0.3.
**CUTPLOT_IGM** \( r_5 \) \( r_6 \) (IGM model)

- 2 real numbers; only concerns IGM results (impacting the cube files content)

\( r_5 \): density threshold used to filter values written to the \( \delta g_{\text{intra}} \) cube file. Default: 1.2. Points with \( \text{sign}(\lambda_2)\rho \) in the range \([-r_5 : r_5]\) will be stored in the \( \delta g_{\text{intra}} \) cube file.

\( r_6 \): density threshold used to filter values written to the \( \delta g_{\text{inter}} \) cube file. Default: 0.3. Points with \( \text{sign}(\lambda_2)\rho \) in the range \([-r_6 : r_6]\) will be stored in the \( \delta g_{\text{inter}} \) cube file.

*This option is normally not necessary* since \( \delta g_{\text{intra}} \) and \( \delta g_{\text{inter}} \) cubes files are naturally separated within the IGM model (without any specific user input).

**VMD_COLRANG_IGM** \( r_7 \) \( r_8 \) (IGM model)

- 2 real numbers; only concerns the VMD visualisation state file generated by IGMPLOT (the way the isosurfaces will be colored)

\( r_7 \): \( \delta g_{\text{intra}} \) iso-surfaces will be colored according to a BGR scheme over the ED range \(-r_7 < \text{sign}(\lambda_2)\rho < r_7 \) a.u.; default = 0.3.

\( r_8 \): \( \delta g_{\text{inter}} \) iso-surfaces will be colored according to a BGR scheme over the ED range \(-r_8 < \text{sign}(\lambda_2)\rho < r_8 \) a.u.; default = 0.08.

*This option is normally not necessary* since

---

**VI. Examples**

**Example 1 (promolecular density):** to use one or two fragments? \((\text{test1,test2})\)

**Water dimer: considered with two fragment.xyz files**

The input param.igm file is reduced to 3 lines as shown below. Type the following command in a terminal:

```
IGMPLOT param.igm > igm.log
```

Using two fragment .xyz files in this example has enabled the IGM approach to automatically separate the intra/inter-fragment contributions. Using the generated file ‘out-igm.dat’, we can construct the following 2D plots signatures (Fig. 12), with gnuplot for instance: `gnuplot> plot “out-igm.dat” u 1:4` --&gt; gives the \( \delta g_{\text{inter}} \) signature (frame of right below).

```
2
frag1.xyz
frag2.xyz
```
\(\delta_{\text{g\_intra}}\) (for weak interactions) is generally one order of magnitude smaller than \(\delta_{\text{g\_intra}}\) (covalent). The default increment has been used here (0.1 Å), which can be reduced to 0.05 Å for instance to get more accurate interaction signatures (using the \textsc{INCREMENTS} keyword).

\textsc{IGMPlot} has also generated the result cube files (‘out-dens.cube’, ‘out-dgInter.cube’ and ‘out-dgIntra.cube’) to construct a colored 3D representation of these interactions using for instance the program VMD.

To make it easier the generation of these 3D representations using the VMD program, the igm.vmd session is prepared automatically for you by \textsc{IGMPlot}: two representations (intra and inter) are then available in the VMD graphical interface, leading to the following Fig. 13. \textsc{IGMPlot} has automatically detected the maximum of \(\delta_{\text{g\_intra}}\) (and \(\delta_{\text{g\_inter}}\)) peak values stored in the cube file such as to set the \(\delta_{\text{g\_intra}}\) (and \(\delta_{\text{g\_inter}}\)) isovalue to 40% of this \(\delta\) maximum value (in the generated vmd session file). Load this file from the VMD GUI (menu File/Load Visualization state).

Don’t forget to adjust manually (in this .vmd file) the absolute file path for vmd to access the data. Or, you can use the \texttt{vmdpath} bash which is made available in the \textsc{IGMPlot} distribution and which will transform the .vmd file to adjust the absolute path to your current path.

The \(\delta\) index gives rise to a coherent picture with stabilizing interactions accumulated in the center of the envelop enclosing the bond critical point (BCP). Of course, depending of the iso-value chosen for \(\delta\) index, iso-surfaces will appear more or less large. The interesting thing is that using the “2-fragments mode”, the intermolecular interaction signature \(\delta_{\text{g\_inter}}\) has been automatically extracted. Let us recall however that a promolecular ED cannot describe covalent interactions in a rigorous way. Thus, the \(\delta_{\text{g\_intra}}\) signature (2D or 3D isosurface) is not really relevant in this case. Please, use an ED derived from wave function calculations instead to study strong interactions.
(see hereafter). However, in the low electron density (ED) domain, it has been shown (see for instance ref.3) that promolecular and quantum ED IGM signature are very similar both in order of magnitude and in the peak height position; accordingly, it is very attractive for studying non-covalent interactions within large systems.

**Water dimer considered with one fragment.xyz file**

The input param.igm file is now reduced to 2 lines. Using the ‘out-igm.dat’ file generated by IGMPlot, we can build (for instance using the gnuplot program) the following 2D plots (Fig. 14).

1
dimer.xyz

![NCI](image1.png) ![IGM δg_intra](image2.png) ![IGM δg_inter](image3.png)

*Fig. 14 water dimer – a single .xyz file has been used; δg_intra (blue, middle) contains both intra and inter peak signatures in that case, and δg_inter = 0 here (blue, right)*

Within the IGM-δg approach, using the single-fragment mode is not useful since intra/inter contributions are not separated in that case. However, in some cases (see below, molecular dynamics of a single peptide), some information can be extracted from the IGM approach, yet.

**Example 2 (promolecular density): vdW interactions**

Van der Waals non-covalent interactions are also well detected by the IGM-δg_inter approach. In the Uracyl dimer, thanks to the uncoupling “2-fragments” scheme and employing promolecular ED, intramolecular ring closure is naturally separated from intermolecular weak interactions occurring between the two monomers (Fig. 15 and Fig. 16).

2
frag1.xyz
frag2.xyz

![Covalent signature](image4.png) ![Internal mol. signature](image5.png)

*Fig. 15 Uracyl dimer δg_intra and δg_inter signatures*
Fig. 16 uracyl (parallel displaced) dimer; $\delta_g^{\text{inter}}$ 0.010 a.u. isosurface (left) with BGR color code in the range -0.08 < $\text{sign}(\lambda_2)$ $\rho$ < 0.08 a.u; atomic decomposition scheme (right) with BGryR color scale (Blue=0%, Red=8.6%).

As above-mentioned, in the promolecular mode, a $\delta_g$ decomposition scheme is automatically performed by IGMPlot in order to give an estimation of the contribution of each atom in the peaks appearing in $\delta_g^{\text{inter}}$ 2D plots. After summing over the grid ($\sum_{\text{grid}} \delta_g^{\text{at}}$, without distinction between the repulsive and attractive parts of the 2D plot), each atom is given a score (%) and colored according to this percentage. These data are stored to the 'percent.dat' file. The generated file 'percent.dat' is employed by the 'atContr.vmd' session file to generate the picture reported on the right of Fig. 16 uracyl (parallel displaced) dimer; $\delta_g^{\text{inter}}$ 0.010 a.u. isosurface (left) with BGR color code in the range -0.08 < $\text{sign}(\lambda_2)$ $\rho$ < 0.08 a.u; atomic decomposition scheme (right) with BGryR color scale (Blue=0%, Red=8.6%). As can be seen, discrepancies occur between atoms showing that they contribute differently in the flat iso-surface representing the van der Waals interaction between the two monomers here. This tool can be considered as a complement to the $\delta_g^{\text{inter}}$ analysis and is mainly interesting in the case of host-guest interaction as will be shown hereafter.

Example 3 (promolecular density): quantification of non-covalent interactions (test4 to test10)

Since non-zero values of $\delta_g$ exclusively corresponds to molecular interaction situations, the IGM-$\delta_g$ approach is very convenient and suitable for integration schemes. Actually, the question of how defining the integration volume does not arise, one simply need to sum $\delta_g$ over the entire grid. It is proposed for the first time in this release an integration scheme developed to relate the sum of $\delta_g^{\text{inter}}$ to the strength of interaction between fragments.

**Link between energy stabilization and $\int_{\text{entire grid}} \delta g^{\text{inter}} dV$ for selected dimers**

For comparison, four dimers of varying strength are described hereafter (see Fig. 17). First, we observe that the peak heights are related to the dimer stabilization energies at the CCSD(T) level of theory (see paper ...). Furthermore, the integrated values $\int_{\text{entire grid}} \delta g^{\text{inter}} dV$ fits rather well to a linear correlation for QM calculations.
Towards the development of a scoring function from promolecular ED

The idea of integrating the local descriptor $\delta g_{\text{inter}}$ has been extended to a larger dimer along a molecular dynamic trajectory in an attempt to find out whether it was possible to establish a relationship between the integrated $\delta g_{\text{inter}}$ and the stabilization energy assessed at the quantum mechanical level of theory. To do so, the interaction between the cucurbit[7]uril (CB7) with the guest bicyclo[2.2.2]octane (B2) has been examined (Fig. 18).

In order to explore an extend range of host-guest binding energies, a steered molecular dynamic was performed allowing us to observe a substrate being pulled from the binding cavity in explicit water at 300 K. From the trajectory, 4000 frames were extracted and subjected to both an IGM-$\delta g_{\text{inter}}$ calculation and a semi-empirical PM7 quantum calculation (actually, three QM calculations are required to get the binding energy for a given dimer geometry). Several integration schemes have been considered. It turns out that by focusing on the peaks of the 2D-
IGM-\(\delta_{\text{inter}}\) signature (by taking those points for which the ratio \(\frac{\nabla \rho_{\text{IGM}}}{\nabla \rho}\) is greater than 1.2) and considering the grid points for which \(\lambda_2 < 0\) (attractive regions), a significant correlation is observed between \(\Delta E\) PM7 (kcal/mol) and \(\int \delta g_{\text{inter}}\) as illustrated on Fig. 19.

![Fig. 19 PM7 stabilization energy versus \(\int \delta g_{\text{inter}}\) obtained from the IGM-2-fragments approach reported for 4000 frames extracted from a steered molecular dynamics at 300K](image)

A cubic polynomial expression has been obtained by regression that fits well with the PM7/IGM data \((R^2=0.989)\). This model has been used to get a simulation of the binding energy from the IGM-\(\delta_{\text{inter}}\) approach, leading to the next graph (Fig. 20) which shows the host-guest interaction energy along the studied molecular dynamic trajectory:

![Fig. 20 B2-CB7 interaction energy along the 2 ns steered MD trajectory; left panel: PM7 QM binding energy and integrated IGM-\(\delta_{\text{inter}}\) score, right panel: PM7 and IGM smoothed curves; \(\delta_{\text{inter}}\) isosurfaces with BGR color code](image)

As can be seen, the three stages of the substrate being pulled from the binding cavity are well recovered by the integrated IGM-\(\delta_{\text{inter}}\) score: (1) cage-guest interaction during the extraction (2) ligand-cage surface interactions before (3) the final release.

Applying this IGM-\(\delta_{\text{inter}}\) scoring function to the water dimer for instance leads to a consistent result \((\Delta E = -5.5\) kcal/mol), and the following output is generated (Fig. 21):
Fig. 21 Example of IGMPlot output for the water dimer with the 2-fragments approach and promolecular ED

The IGM-δg_site integration scheme is not affected by the keywords CUTOFFS, CUTPLOT, CUTPLOT_IGM (which only affect the final .dat and .cube output files).

Although consistent results have been obtained here, some limitations exist with this first model using promolecular ED. First, it only describes non-covalent interactions in the calibration range 0-40 kcal/mol. Only non-covalent interactions can be addressed by this model (covalent and strong repulsive steric cannot be described of course). Warnings will be delivered by the code in these situations. Also, so far, it is based on the existence of ED contragradience. Thus, long-range electrostatic interactions (which may not involve significant ED clash) cannot be taken into account (note that, in the presence of the water solvent with large dielectric constant, this issue is reduced). Also, this score is found to overestimate π-stacking interactions over hydrogen bonding. This is a proof of concept, which is currently being developed using machine learning possibilities employing a variety a different molecules to remedy these lacks and to provide a more general and still better fit with QM interaction energy.

Of course, the IGM-δg_site approach using promolecular electron density cannot be expected to lead as accurate results as QM calculations, but it offers a fast and robust way to get an order of magnitude of interactions comparing different systems or during their evolution.
Example 4 (promolecular density): Monitoring intramolecular interactions with IGM-$\delta g^{\text{intra}}$ in a peptide (test11, test12)

In addition to the IGM-$\delta g^{\text{inter}}$ local descriptor describing the interactions between two fragments A and B, the IGM approach also delivers the IGM-$\delta g^{\text{intra}}$ descriptor. It describes all the interactions present inside each fragment. Generally, these $\delta g^{\text{intra}}$ interactions correspond to covalent bonding (Case 1 on Fig. 22).

<table>
<thead>
<tr>
<th>Fragments considered</th>
<th>$\delta g^{\text{intra}}$</th>
<th>$\delta g^{\text{inter}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>A .... B</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>Case 2</td>
<td>A .... B</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>Case 3</td>
<td>A -- B</td>
<td>&gt; 0</td>
</tr>
<tr>
<td></td>
<td>C .... D</td>
<td></td>
</tr>
</tbody>
</table>

*Fig. 22 Several scenarios of fragmentations within IGMPlot*

When we consider the system A…B taken as one piece (one single fragment, Case 2 on Fig. 22), the intra/inter uncoupling scheme automatically leads to IGM-$\delta g^{\text{inter}} = 0$ (at every grid node), and IGM-$\delta g^{\text{intra}}$ reveals every interaction present in the system: both covalent bonding inside A and B and weak interaction between A and B. Then, the IGM-$\delta g^{\text{intra}}$ signature displays two domains, strong and weak interactions. Such a use (defining a single fragment) of the IGM approach does not make sense if you are interested in the intermolecular A…B interaction. In that case, it is much better to define two fragments to separate automatically the two signatures.

A last case may occur when one single molecule involves both covalent and intramolecular non-covalent interactions (between C and D, Case 3). In that case, one may be interested in extracting and quantifying the intramolecular weak interaction part of the whole IGM-$\delta g^{\text{intra}}$ signature. Indeed, for instance, it may be worthwhile to get information on folding patterns of a single peptide (like alpha helices and beta sheets), which results from intramolecular weak interactions between neighbouring amino-acids and makes up the secondary structure of a protein.

That is the reason why IGMPlot automatically provides the attractive contribution to the weak interaction IGM-$\delta g^{\text{intra}}$ signature. This way, the IGM-$\delta g^{\text{intra}}$ can then serve to study energy changes in a single molecule which are only due to weak interactions.

This possibility is illustrated below during a molecular dynamic trajectory of a peptide. In this simulation, due to fluctuating weak interactions between side chains of amino acids, the non-covalent contribution to IGM-$\delta g^{\text{intra}}$ changes. Then, IGM-$\delta g^{\text{intra}}$ (weak interaction part) can be employed to monitor energy changes in a single molecule that results from non-covalent interactions. The result is shown on Fig. 23 which reports the back-bone RMSD of a peptide (PDB ID 1dep) and the evolution of the IGM-$\delta g^{\text{intra}}$ (weak interaction domain) curve along a 2 ns MD trajectory at 300K. As can be seen, starting from a quasi-linear structure, the peptide adopts an organized structure (a turn) due to non-covalent interactions between amino acids. The RMSD plot reveals the presence of this organized state after 0.7 ns. Accordingly, the $\int -\delta g^{\text{intra}}$ sum (weak interaction domain) decreases in the range [0: 0.7 ns] and reveals the gradual accumulation of stabilizing weak interactions, accurately reflecting the formation of this “turn” state.
Fig. 23 Membrane protein (PDB ID 1dep), sequence: ARG-SER-PRO-ASP-PHE-ARG-LYS-ALA-PHE-LYS-ARG-LEU-LEU-CYS-PHE; backbone RMSD (Å, green) and $\int -\delta g^{\text{intra}}$ (a.u., scaled by 5, weak interaction domain, purple curve) along a 2ns MD trajectory (smoothed curves on the right).

This example on a peptide proves that the integral of $-\delta g^{\text{intra}}$ obtained in the non-covalent domain and from promolecular electron density can be used to obtain meaningful chemical predictions so as to produce output data representative of the time evolution of the molecular system, with the added flexibility of focusing on a subdomain of a biomolecular system.

Work is on progress to still improve the sensitivity of this approach and to provide energetics data in conventional units (kcal/mol).

As already pointed out, the non-relaxed electron density obtained by means of the promolecular approach is not appropriate to describe strong interactions (covalent domain).

Example 5 (promolecular density): smart use of the uncoupling scheme in a trimer (test13)

Let us consider the following cyclic cluster, and say we are interested in the interaction energy between (H$_2$O)$_2$ and HF:

Fig. 24 The two fragments defined for the IGM study of a cyclic cluster

To access the interaction between just HF and the rest of the cluster we choose the two chosen fragments as following: (H$_2$O)$_2$ + HF (Fig. 24). The resulting 2D plots are reported on Fig. 25.
As can be seen, the OH…O and OH…F hydrogen bonds are automatically and naturally separated using the IGM approach leading to the following $\delta_{\text{inter}}$ representation (Fig. 26), not polluted with the water dimer (internal) hydrogen-bonding:

As already pointed out by the $\delta_{\text{inter}}$ 2D-plot (peak heights on Fig. 25), it is clear from this 3D-picture that the O…HF hydrogen bond strength is larger (larger volume enclosed by iso-surface) than that of the OH…F hydrogen bond. It worth noticing that using this partition removes the ring closure in the $\delta_{\text{inter}}$ representation since this feature results from the overall cluster (taken as one piece).

**Example 6 (promolecular density): ligand –protein (test14)**

For large systems, the use of the fast promolecular density (computed by IGMPplot) is advised.

From a practical point of view, studying ligand-host interactions requires choosing the ligand as the first molecule specified in the param.igm input file in order to reduce the grid size and avoid calculations. Depending on the size of the ligand and the grid step value, the computation time might be large (timing predictions are regularly updated in the runinfo file). On Fig. 27 is reported the interaction between a pyridazinone derivative and an isoform of the phosphodiesterase IV (PDE4) protein.
In tandem with $\delta_{g\text{inter}}$ isosurfaces, the **atomic scheme decomposition** $\Delta g_{\text{inter/At}}$ brings supplementary information. For instance, here, it can be seen that the indole moiety of the ligand primarily contributes to the interactions in this complex with protein “hotspots” well identified (red color).

From a practical point of view, the IGM workflow exhibits some facilities. Drawing the corresponding ligand-protein interactions and performing the atomic contribution analysis is straightforward and parameter free.

**Example 7** (promolecular density): when to use the **CUTPLOT_IGM** keyword? *(test15)*

Let us consider the single molecule on **Fig. 28**:

![Fig. 27](image)

**Fig. 27** ligand – protein example; three analyses from left to right: $\delta_{g\text{inter}}$ 2D-signature, $\delta_{g\text{inter}}$ 0.013 a.u. isosurfaces with BGR color code in the range -0.08 < sign($\lambda_2$) $\rho$ < 0.08 a.u., host-guest colored according to the $\Delta g_{\text{inter/At}}$ atomic contributions to the intermolecular interactions, using a BGryR color scale (blue=0%, red=2.9%).

We are interesting in this case by non-covalent interactions located between the two substituents (-Cl and the group -CH$_2$OH). To get this interaction using the promolecular ED model, the first idea is to proceed with two fragments: -Cl on one hand, and the rest of the molecule one the other hand. Unfortunately, in addition to the desired non-covalent Cl…CH$_2$OH interaction we will also get the covalent C-Cl interaction peak. Then, exceptionally, in addition, we can use the keyword **CUTPLOT_IGM** (in the param.igm input file) to tell IGMPPlot to write in the resulting cube file only the points corresponding to non-covalent interactions present on the 2D-plot reported below on **Fig. 29**.

The second number given after the keyword **CUTPLOT_IGM** (0.03) means that only $\delta_{g\text{inter}}$ points for which the signed ED is in the range [-0.03:0.03] will be printed to file ‘dgInter.cube’. This will get rid of the C-Cl covalent bond still present in the $\delta_{g\text{inter}}$ representation. This ensures having the desired “intramolecular” non-covalent interaction in the resulting 3D picture reported on the right of **Fig. 29**:

```
2
frag1.xyz
frag2.xyz
INCREMENTS 0.05 0.05 0.05
CUTPLOT_IGM 0.3 0.03
```
Using **IGM_CUTPLOT** keyword allows for extracting and visualizing the desired weak contribution.

Note that in this very special 2-fragments case, the “deltaE Inter Estimate” delivered by IGMplot (not affected by the **CUTPLOT_IGM** keyword) includes C-Cl covalent bonding and is not meaningful (promolecular ED is not appropriate for strong interaction). In order to have an estimate of this weak intramolecular interaction between the chlorine atom and the hydroxyl group, you can opt for the 1-fragment IGM scheme and get the \( \Delta g[\text{intra_weak}] \) score.

**Example 8 (promolecular density):** atomic decomposition scheme of non-covalent interactions applied to host-guest assemblies (**test16, test17**)

The atomic scheme decomposition \( \Delta g_{\text{inter/At}} \), is an extension of the original IGM-\( \delta g \) approach and is able to emphasize the most relevant contributions to the non-covalent interactions occurring between two fragments, after the integration of the local \( \delta g_{\text{inter/At}} \) index, complementary to \( \delta g_{\text{inter}} \). We highlight in Fig. 30 this possibility on a trimuclear copper complex serving as a buckycatcher.

```
2
frag1.xyz
frag2.xyz
INCREMENTs 0.1 0.1 0.1
```

**Fig. 30** a) Optimized structure of the \( \text{[Cu}_3(3,5-\text{CF}_3\text{pz})_3]^{-}\text{C}_60 \) 1:1 system; b) IGM analysis with a isosurface \( \Delta g_{\text{inter/At}} \) of 0.01 a.u., BGR color code in the range \(-0.06 < \rho \times \text{sign} \times (\lambda) < 0.06 \) a.u. c) Host colored according to the \( \Delta g_{\text{intra/At}} \) score using a BGrYR color scale.
A $\Delta g_{\text{inter}/\text{At}}$ value of 23.1% has been obtained for the three copper atoms confirming their significant participation in the C$_{60}$ capture. However, the high overall value $\Delta g_{\text{inter}/\text{At}}$ value associated with the three pyrazolate rings (48.0%) and the lateral CF$_3$ groups (28.9%) show that these fragments are also relevant in the C$_{60}$ inclusion.

An additional example is given below in the field of supramolecular systems based on host-guest chemistry. On Fig. 31 a macrocyclic oligothiophene incorporates C$_{60}$ in its inner cavity to form a unique Saturn-like complexes. The IGM-$\delta g$ reveals van der Waals interactions between the fullerene and the host sulfur atoms, as well as atomic contributions to this interaction.

![Fig. 31 Host-guest example (oligothiophene-C60 complex); $\delta g_{\text{inter}}$ isosurfaces and $\Delta g_{\text{inter}/\text{At}}$ scores](image)

This tool can be useful for teams devoted to the synthesis of functional supramolecular materials, where a judicious structural modification may result in an advantageous tuning of the host-guest capabilities.

**Example 9 (QM):** H$_2$O ... H$_2$O dimer, **FRAG1 keyword** *(test18)*

In the QM mode (WFN/WFX) the electron density is obtained from a wave function, prior to the IGM calculation. The name of a single .wfn or .wfx file has to be supplied. Without any specific keyword, the whole system will be considered as a single molecule, and the resulting $\delta g_{\text{intra}}$ cube will be 0.0 (all interactions will be present in the $\delta g_{\text{intra}}$ cube, like in the 1-fragment approach using the promolecular ED).

Instead, in order to separate the two domains (intra/inter), you can use the **FRAG1** keyword and supply an atom selection pattern (see example below). For instance, in the provided .wfn file, if the first three atoms of the water dimer belong to the first water molecule and we want to study the interaction between the two water molecules, we then define the fragment 1 using the **FRAG1** pattern: ‘FRAG1 1-3’.

Be aware that if only FRAG1 keyword is specified in the param.igm input file, fragment 2 will be made of the rest of the atoms, such that frag1+frag2 represents the whole system.

```
1
  dimer.wfn
  INCREMENTS 0.07 0.07 0.07
  ONAME dimer
  FRAG1 1-3
```
Among other things, it can be noted (orange plot in Fig. 32) that the promolecular ED (also reported here for comparison) clearly underestimates ED gradient drops for covalent bonding ($\delta g_{\text{max intra, pro}}^{\text{intra}} = 0.4$) vs ($\delta g_{\text{max intra, QM}}^{\text{intra}} = 0.9$). In contrast, in the domain of weak interactions, SCF quantum-mechanical ED and frozen promolecular ED have very similar features, which have served to validate the promolecular approach for weak interactions.

**IGMPlot** allows you a bit of flexibility to handle cases where atoms in fragment 1 are not written in a contiguous way in the .wfn file. We give below some general guidelines for specifying atom numbers through **FRAG1** keyword:

- You can specify a single atom number: **FRAG1 5**
- You can specify a group of non-sequential atoms by separating them by semicolons: **FRAG1 4;8;19**
- You can specify a sequential range of atoms by putting a dash between them: **FRAG1 4-7**
- You can mix the possibilities: **FRAG1 6;2-4;22**
- You can specify the range: **FRAG1 4-** (from atom 4 to the last atom in the wfn file).
- Any whitespace character will be ignored.

**Example 10 (QM):** $\text{B}_2\text{H}_6$ taken as one piece (test19)

The typical case of a three-center two-electron bond can hardly be partitioned in two separated fragments.
Example 11 (QM): Agostic interaction in the metallic complex [Ti Cl₂ CH₂ CH₃]⁺ (test20)

Here is a smart use of the inter-uncoupling scheme: fragment 1 = Titanium atom (number 1), fragment 2 will be the rest of the system. This allows you to probe every interaction involved by the metallic centre. In addition to the 3 coordination bonds we can observe on Fig. 34 that the δg_intra index is able to detect the agnostic interaction between Ti and C-H.
Example 12 (QM): FRAG1 and FRAG2 keywords: quantifying intramolecular $\pi$-stacking (test21)

The combination of FRAG1 and FRAG2 keywords allows to probe interactions between two sub-parts of a given molecule. This possibility is very convenient to study intramolecular $\pi$-stacking. In order to do so, both FRAG1 and FRAG2 atom selections must be supplied in the param.igm input file.

Be aware that, when FRAG1 and FRAG2 are specified, the sum (FRAG1 + FRAG2) can be smaller than the whole system.

In the example below (Fig. 35), FRAG1 and FRAG2 have been set to the phenyl ring and indole-based ring:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mol.wfn</td>
<td></td>
</tr>
<tr>
<td>FRAG1 11-21;35-42</td>
<td>FRAG2 1-6;26-30</td>
</tr>
</tbody>
</table>

Then, only the atomic orbitals of the FRAG1/FRAG2 selections are considered to achieve the IGM-$\delta g_{\text{inter}}$ calculation. Accordingly, the ED is limited to fragments 1 and 2 atomic sources.

It is strongly advised not to use diffuse functions to perform a transition state IGM-$\delta g_{\text{inter}}$ analysis.

As can be seen, the flat region of interaction between the two aromatic rings can be extracted from the wave function information. IGMPlot also quantifies this interaction through the integration of the $\delta g_{\text{inter}}$ descriptor. The corresponding output emphasizes how many orbitals have participated to probe this internal interaction (Fig. 36). This option should be useful in organic chemistry for mechanistic studies often revealing such non-covalent but intramolecular interactions during reactions.
The corresponding \( \pi \)-stacking interaction is quantified by IGMPLOT thanks to the integral: \[ \int \delta g_{\text{inter}} \, dV \] (given in the output). Hence, it is now for instance possible to monitor the effect of internal \( \pi \)-stacking during a chemical reaction across the stationary points of a given reaction path.

Example 13 (QM): **CUBEFRAG** keyword (test22)

Performing an IGM analysis on a large system from a WFN (or WFX) input file, may be CPU expensive. That is the reason why the option **CUBEFRAG** is proposed. It is intended to limit the atomic orbitals considered in the calculation to a reduced portion of the system. **CUBEFRAG** means that only the atoms within the working grid defined by the keyword **CUBE** (or **RADIUS**) will be considered. Hence, the fragment 1 will be a part of the whole system, and the fragment 2 will be empty. Accordingly, the \( \delta g_{\text{inter}} \) descriptor will be 0, and \( \delta g_{\text{intra}} \) will describe all the interactions present in fragment 1. Note that the ED is then limited to fragments 1 and 2 atomic sources.

As above-mentioned, the complementary keyword **CUBE** (or **RADIUS**) must be supplied in order to position the grid in the system.

In this example, the **CUBE** keyword sets the grid around the two aromatic rings of the system. Then, specifying **CUBEFRAG** will limit the IGM analysis to those atoms inside the rectangular box. The covalent bonding isosurfaces are obtained by choosing a relative large isovalue (0.3 a.u. on Fig. 37); furthermore, to get rid of the strong interaction domain, one could use the keyword **CUTPLOT_IGM**, and then obtain the non-covalent interactions like \( \pi \)-stacking (not represented here).

As above-mentioned, a new $\Delta g_{\text{pair}}$ descriptor has been built to focus the IGM analysis solely on the atomic ED gradients $\nabla \rho_i$ brought by a given atom pair. A bond by bond picture appears (see **Fig. 8**), and it could be tempting to associate the strength of the bond with peak height of the 2D-plot $\Delta g_{\text{pair}}$ signature. However, each bond has its own individual profile suggesting that correlating the $\Delta g_{\text{pair}}$ descriptor should not be limited to the use of a local information (bond critical point for instance). The IBSI integrated score stems from this research and provides a quantitative bridge between a local electron density-based descriptor and the physically grounded bond strength concept. The results obtained on a large set of bonds underpin the IGM formalism and its two key-components: the non-interacting reference definition and the gradient based partition (GBP).

We refer the reader to the corresponding study, where numerous examples are given. Here, only one example is given, covering all the interaction domains: non-covalent, metal coordination and covalent bonding. **Fig. 38** shows a transition state structure (TS) taken from a study on the reactivity of thiolates with cisplatin. The mechanism proceeds with the intramolecular substitution of the coordinated water molecule by the $\text{CO}_2$- group of the cysteine, which ultimately will act as a $(S,O)$-bidentate chelating ligand in the final product. The IBSI index and Mayer bond order are reported in bold for some atom pairs.

<table>
<thead>
<tr>
<th>mol.wfn</th>
<th>IBSI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9 10</td>
</tr>
<tr>
<td></td>
<td>2 5</td>
</tr>
<tr>
<td></td>
<td>9 5</td>
</tr>
<tr>
<td></td>
<td>12 6</td>
</tr>
<tr>
<td>ENDIBSI</td>
<td></td>
</tr>
</tbody>
</table>

Theoretically, you can specify more than one bond in the **IBSI/ENDIBSI** section. However, you may prefer to distribute the IBSI calculations of several bonds in parallel on several computation nodes.
Fig. 3B Bond strength Index (IBSI, bold) and \( \delta_{\text{pair}} = 0.045 \) a.u. isosurfaces for selected bonds in a TS involving a platinum complex (M06-2X LANL2DZ); color coding in the range \(-0.3 < \text{sign}(\lambda^2) \rho < 0.3\) a.u; Mayer bond order in parentheses

From this figure we can see that IBSI and the Mayer bond order are complementary. For instance, according to the Mayer analysis (MBO), the C-S and S-Pt bonds exhibit an identical bond multiplicity of about 1.0. The IBSI provides an additional information: the covalent bond C-S is stronger (IBSI=0.520) than the S-Pt coordinate bond (IBSI=0.278). This is consistent with the strength of metal ligand bonds known to be intermediate between that of covalent bonds and non-bonded interactions.

In other respects, the MBO predicts a formal half-bond between the nitrogen and platinum atoms (0.507 pair) and a single bond between the sulfur and platinum atoms (1.023 pair). It may be then naïvely expected that the corresponding metal-ligand bond strengths should correlate with these bond orders at a ratio of 1(N-Pt):2(S-Pt). However, the IBSI predicts an inverse strength ratio with the N-Pt bond being significantly stronger (IBSI=0.338) than the S-Pt one (0.278).

The IBSI stays attached to the restoring force concept, allowing for internally probing bond strength in TS structures and along reaction pathways for breaking and forming bonds.

The IBSI covers a large range of interactions, from hydrogen-bonds to covalent bonding, as long as contragradience domains are well separated. However, there are limitations in interpreting IBSI for interactions with complex bonding scenarios where more than two atoms are involved in a bonding situation. Metalloccenes for instance and even non-covalent stacking interaction, like in the benzene dimer, belong to this category. Actually, the IGM tool is fully able to reveal any kind of interaction resulting from ED clash, however, whether the atom-pair IBSI score can be used is usually obvious by inspection of the molecular geometry and the shape of the contragradience isosurfaces. An indicative scale is provided (obtained for a set of 677 bonds of 235 molecules in their ground state):

From a practical point of view, the IBSI keyword does not generate cube files to visualize bond isosurfaces. Actually, an ultrafine cylindrical grid is employed by IGMPplot to calculate the IBSI score, which is not compatible with visualization program commonly used in theoretical chemistry. Work is on progress to overcome this difficulty by using the ParaView software. Meanwhile, one can use the FRAG1/FRAG2 tandem keywords to define two atomic fragments corresponding to the desired bond to be visualized.
Example 15 (QM): **IBSI keyword with cutoff (test24)**

Although the IBSI analysis is limited to those atomic ED gradients $\nabla \rho_i$ brought by the considered atom pair, the calculation of each atomic term $\nabla \rho_i$ takes into account all the atomic orbitals of the entire system. It can be time-consuming. To speed-up the calculation, a cutoff can be applied around the considered bond (see Fig. 39). That way, only the atoms (and their primitives) within the cutoff radius are considered in the calculation. To do that, simply provide the cutoff value (in Å) after the keywords **IBSI** in the param.igm input file (3.0 Å in the example below).

```
1 mol.wfn
IBSI 3.0
9 10
ENDIBSI
```

The **Table 1** shows the IBSI scores obtained for the single C-C bond in the molecule of **Fig. 39** for several cutoff radii.

![Fig. 39 Cutoff radius taken around each atom of the considered bond for the IBSI calculation](image)

**Fig. 39 Cutoff radius taken around each atom of the considered bond for the IBSI calculation**

Default is to use no cutoff (full calculation), leading here to the IBSI (C-C) = 0.829. However, providing a cutoff radius of 3.5 Å seems reasonable (IBSI = 0.829), saving much time.

<table>
<thead>
<tr>
<th>Cutoff radius (in Å)</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
<th>3.5</th>
<th>4.0</th>
<th>4.5</th>
<th>5.0</th>
<th>No limit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IBSI</strong> (C-C)</td>
<td>0.837</td>
<td>0.835</td>
<td>0.829</td>
<td>0.829</td>
<td>0.829</td>
<td>0.829</td>
<td>0.829</td>
<td>0.829</td>
</tr>
<tr>
<td><strong>primitives</strong></td>
<td>175</td>
<td>210</td>
<td>402</td>
<td>528</td>
<td>640</td>
<td>668</td>
<td>773</td>
<td>843</td>
</tr>
<tr>
<td><strong>Execution time</strong> (s)</td>
<td>148</td>
<td>193</td>
<td>694</td>
<td>1197</td>
<td>1710</td>
<td>1856</td>
<td>2432</td>
<td>2815</td>
</tr>
</tbody>
</table>

**Table 1 Effect of several cutoff radii applied to the IBSI calculation for a single C-C bond in a molecule**

This way, IBSI calculations can be performed on very large systems by focusing on the considered atom pair, taking only a few minutes while maintaining accuracy.

Example 16 (QM): **Bond $\delta g^{\text{pair}}$ signature and atom pair isosurface (test25)**

If you are interested in visualizing the $\delta g^{\text{pair}}$ 2D-plot and associated bond isosurfaces for a given atom pair, you must use the **FRAG1/FRAG2** combination in order to generate a $\delta g^{\text{inter}}$ cube file, and then you must remove the **IBSI** keyword from the param.igm input file. Actually, the IBSI score is obtained using an ultrafine cylindrical grid. Ideally, this grid should be used to generate bond isosurfaces. However, software commonly used in Chemistry like VMD are only able to build three-dimensional isosurfaces from scalar data store in cube file
(gaussian cube file), not a cylindrical one. Work is on progress to generate cylindrical files compatible with the use of ParaView. Meanwhile, you can run a FRAG1/FRAG2 IGM calculation (without the IBSI keyword) where FRAG1 represents the first atom and FRAG2 the second atom of the bond. In that case, it is strongly advised to set the CUTOFFS and CUTPLOT_IGM keywords to: “CUTOFFS 10.0 100.0” and “CUTPLOT_IGM 0.3 10.0” in order to obtain a complete isosurface (even close to the atoms where the electron density is large). This is illustrated on Fig. 40.

One has to keep in mind that using the tandem FRAG1/FRAG2 keywords means that only the primitives coming from the two selected atoms are considered in the calculation of $\delta g_{\text{pair}}$ and of the ED used to color the $\delta g_{\text{pair}}$ isosurface. The visual rendering is however consistent with what is obtained from a full calculation. On the contrary, using the IBSI keyword to calculate accurately the bond strength will make every primitive be considered in the calculation (unless a cutoff radius is employed).

**Example 17 (QM): Bond Asymmetry Factor (test26 to test29)**

For the first time, a Bond Asymmetry Factor (BAF) is proposed in the IGMPplot release. It can be obtained by using the electron density (ED) coming from QM calculations and is delivered with the IBSI value (keyword IBSI). It has been found incidentally by coupling the Gradient Based Partition (GBP) to the IGM approach as is explained below. This index tells us how much the bond strength is asymmetrical and the direction of asymmetry. For instance, looking at the C-O bond in the methanol, we get the following output from IGMPplot:
As can be seen on Fig. 41, in addition to the bond strength (IBSI=1.140), an asymmetry is given by IGMPPlot for the bond strength (BAF=0.339), in the direction from C to O. The BAF would be exactly zero in the case of a purely symmetrical bond like C-C in ethane.

The BAF was developed from the Gradient Based Partition which defines atomic electron density gradients from the expression of a molecular wave function. It involves the use of two new descriptors: \( \delta g_{L_{1}S_{2}} \) and \( \delta g_{S_{1}L_{2}} \). In order to present the BAF index on a simple example, let us take a molecular orbital \( \psi \) occupied by two electrons and written as a linear combination of two atomic orbitals (AO): \( \Psi = C_{1}\phi_{1} + C_{2}\phi_{2} \), on atoms 1 and 2, respectively. The associated ED is: \( \rho = 2(C_{1}^{2}\phi_{1}^{2} + 2C_{1}C_{2}\phi_{1}\phi_{2} + C_{2}^{2}\phi_{2}^{2}) \). The GBP expresses the resulting gradient component \( \frac{\partial \rho}{\partial x} \) (noted \( \nabla \rho \) for the sake of convenience) as the sum of two atomic terms: \( \nabla \rho = \nabla \rho_{1} + \nabla \rho_{2} \).

The local descriptor \( \delta g_{pair} \) measuring the ED contragradience (clash) between these two sources 1 and 2 relies on the difference: \( |\nabla \rho_{1}| + |\nabla \rho_{2}| - |\nabla \rho_{1} + \nabla \rho_{2}| \). By considering now each atomic ED gradient made of two sources called \( \nabla \rho_{L} \) (Localised source) and \( \nabla \rho_{S} \) (Shared source) we can also determine the ED contragradience between each source type (see Fig. 42):

![Fig. 42 Origin of \( \delta g_{pair} \), \( \delta g_{L} \) and \( \delta g_{S} \)](image)

leading to two other ED contragradience descriptors called \( \delta g_{L} \) and \( \delta g_{S} \). Fig. 43 represents the isosurfaces of these two descriptors. The \( \delta g_{L} \) isosurface looks like a compact disc. It will not be described further here, but \( \delta g_{S} \), based on purely Localized atomic sources, is similar to what would be obtained using promolecular ED. Conversely,
\( \delta g_S \) represents an ED contragradience more delocalized along the atom pair. The smaller \( \delta g_S \) the less shared electrons.

![Fig. 43 Comparison of \( \delta g_{\text{pair}} \), \( \delta g_L \) and \( \delta g_S \) isosurfaces for the C-O bond of methanol; \( \delta g = 0.015 \text{ a.u.} \) isosurfaces color coding in the range \(-0.3 < \text{sign}(\lambda) \rho < 0.3 \text{ a.u.}\)](image)

The integral of \( \delta g_L \) (0.864) and \( \delta g_S \) (0.367) are displayed in the IGMplot output on Fig. 41. \( \delta g_S \) is small as soon as the atoms are distant (\( \delta g_S (\text{H} \cdots \text{O}) = 0.058 \) in the water dimer for instance) or for ionic compounds (\( \delta g_S (\text{Li} \cdots \text{H}) = 0.032 \)).

If you pay close attention to the \( \delta g \) isosurface for the C-O bond on the left of Fig. 43, you can see that it is slightly egg-shaped, gradually narrowing toward the carbon atom and the surface seems to be slightly shifted toward the oxygen atom. To quantify this \( \delta g \) asymmetry, two asymmetrical “cross-descriptors” have been defined as illustrated on Fig. 44: \( \delta g_{L_1S_2} \) and \( \delta g_{S_1L_2} \).

![Fig. 44 Origin of \( \delta g_{\text{pair}} \), \( \delta g_{L_1S_2} \) and \( \delta g_{S_1L_2} \)](image)

Of course, at a given point taken in between the two atoms, \( \delta g_{L_1S_2} \) and \( \delta g_{S_1L_2} \) differ, even for a purely symmetrical bond. Then, the difference (\( \delta g_{L_1S_2} - \delta g_{S_1L_2} \)) calculated at one point is either positive (red) or negative (blue) as illustrated on Fig. 45 for the C-C bond of ethane.

![Fig. 45 (\( \delta g_{L_1S_2} - \delta g_{S_1L_2} \)) along the C-C bond of ethane; 0.02 a.u. isosurfaces, color coding red; -0.02 a.u. isosurfaces, color coding blue; BAF = 0.0)](image)
But, as expected, along the purely symmetrical C-C bond of ethane, in absolute value, this function is symmetrical. Then integrating the function $\delta g_{L_1S_2} - \delta g_{S_1L_2}$ (or $\delta g_{S_1L_2} - \delta g_{L_1S_2}$) over the grid gives $0$ in that case. But for another bond, such as C-O in methanol, the resulting function is asymmetrical (see Fig. 46).

In that case, the integration of this function gives the so-called Bond Asymmetry Factor (BAF, which is always taken in absolute value): $0.339$. Thus, the BAF approach is capable not only of quantifying the asymmetry in bond strength but also of identifying the direction of this asymmetry (the sign of the integral). In the case of C-O in methanol, in absolute value, $(\delta g_{L_1S_2} - \delta g_{S_1L_2})$ is larger on the oxygen atom side, which is the more electronegative atom of this single bond. This is visually in line with the $\delta g_{\text{pair}}$ isosurface found polarized towards the AIM basin of the oxygen atom (on Fig. 47).

To sum up, the local bond strength (represented by the ED contragradience descriptor $\delta g_{\text{pair}}$) varies in between the two atoms of a bond. If the bond is symmetrical, the contragradience will be symmetrical as well and the BAF index will be zero. Otherwise, one observes a local bond strength polarization toward one of the two atoms and the BAF indicates how much it is asymmetrical and the direction of this asymmetry.

Generally, a bond strength polarization is observed toward the more electronegative atom as illustrated in Fig. 48. This makes sense since the more electronegative atom generally displays the largest electron density in the covalent domain, hence it brings the largest ED gradient source, and then the largest $\delta g_{\text{pair}}$ values are larger in the direction of this atom. The BAF trend is the same as the one deduced from electronegativity difference $\Delta \chi$. Compared with the electronegativity difference, the BAF index has the advantage of taking into consideration the molecular chemical environment of the atoms (through the molecular wave function) and can be applied to multiple bonds. Although the BAF index looks similar to the charge density polarization, it rather measures the tendency of the local bond strength (represented by $\delta g_{\text{pair}}$) to be asymmetric and directed towards one of the two atoms. In some cases BAF and $\Delta \chi$ differ, mainly for ionic compounds like LiH. But since the BAF index is based (among others) on shared electron source (S), it makes little sense to compute it for such ionic compounds, or for distant atoms not sharing electrons or even for metal coordination. **We strongly advise to use the BAF index for bonds.**
involving elements of the p-block (and hydrogen). A warning is delivered by the code when two little electrons are shared between the two atoms (based on the value of $\delta g_0$).

<table>
<thead>
<tr>
<th>$\text{C} - \text{H}$</th>
<th>IBSI = 0.009</th>
<th>BAF = 0.199</th>
</tr>
</thead>
</table>

For each atom (except H, alkalis, halogens and C in C$_2$H$_6$) is surrounded with CH$_3$ group(s); M062X/x2c-TZVP level of theory; the arrow shows the polarization direction.

Interestingly, the BAF index is very sensitive to the chemical environment, and as such, can serve to quantify short-range inductive attractive electronic effect as illustrated on Fig. 49.

![Fig. 49 BAF: quantification of the attractive inductive effect of a fluorine atom on the neighbouring C-C bond; C$_2$H$_6$ on the left, C$_2$H$_5$F on the right](image)

As can be seen in the first series of Table 2, the highly electronegative fluorine atoms exert a local bond strength polarizing effect towards $\text{C}_A$. This is because electron density is removed from $\text{C}_A$ in the fluoro-derivatives, hence reducing the ED gradient source $\text{C}_B$ (more than on $\text{C}_A$, more distant). Since ED contragradience (clash) is achieved only by means of ED sources, therefore ED contragradience decreases on the $\text{C}_B$ side more than on the $\text{C}_A$ side, and consequently we observe a bond strength polarization towards $\text{C}_A$.

![Fig. 48 Trends of the BAF and IBSI across the periodic table](image)

<table>
<thead>
<tr>
<th>$\text{C} - \text{Li}$</th>
<th>BAF = 0.068</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C} - \text{Be}$</td>
<td>BAF = 0.092</td>
</tr>
<tr>
<td>$\text{C} - \text{B}$</td>
<td>BAF = 0.114</td>
</tr>
<tr>
<td>$\text{C} - \text{C}$</td>
<td>BAF = 0.111</td>
</tr>
<tr>
<td>$\text{C} - \text{N}$</td>
<td>BAF = 0.199</td>
</tr>
<tr>
<td>$\text{C} - \text{O}$</td>
<td>BAF = 0.267</td>
</tr>
<tr>
<td>$\text{C} - \text{F}$</td>
<td>BAF = 0.335</td>
</tr>
<tr>
<td>$\text{C} - \text{Na}$</td>
<td>BAF = 0.017</td>
</tr>
<tr>
<td>$\text{C} - \text{Mg}$</td>
<td>BAF = 0.057</td>
</tr>
<tr>
<td>$\text{C} - \text{Al}$</td>
<td>BAF = 0.085</td>
</tr>
<tr>
<td>$\text{C} - \text{Si}$</td>
<td>BAF = 0.038</td>
</tr>
<tr>
<td>$\text{C} - \text{P}$</td>
<td>BAF = 0.070</td>
</tr>
<tr>
<td>$\text{C} - \text{S}$</td>
<td>BAF = 0.172</td>
</tr>
<tr>
<td>$\text{C} - \text{Cl}$</td>
<td>BAF = 0.200</td>
</tr>
<tr>
<td>$\text{C} - \text{Br}$</td>
<td>BAF = 0.277</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\text{C} - \text{Li}$</th>
<th>BAF = 0.068</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C} - \text{Be}$</td>
<td>BAF = 0.092</td>
</tr>
<tr>
<td>$\text{C} - \text{B}$</td>
<td>BAF = 0.114</td>
</tr>
<tr>
<td>$\text{C} - \text{C}$</td>
<td>BAF = 0.111</td>
</tr>
<tr>
<td>$\text{C} - \text{N}$</td>
<td>BAF = 0.199</td>
</tr>
<tr>
<td>$\text{C} - \text{O}$</td>
<td>BAF = 0.267</td>
</tr>
<tr>
<td>$\text{C} - \text{F}$</td>
<td>BAF = 0.335</td>
</tr>
<tr>
<td>$\text{C} - \text{Na}$</td>
<td>BAF = 0.017</td>
</tr>
<tr>
<td>$\text{C} - \text{Mg}$</td>
<td>BAF = 0.057</td>
</tr>
<tr>
<td>$\text{C} - \text{Al}$</td>
<td>BAF = 0.085</td>
</tr>
<tr>
<td>$\text{C} - \text{Si}$</td>
<td>BAF = 0.038</td>
</tr>
<tr>
<td>$\text{C} - \text{P}$</td>
<td>BAF = 0.070</td>
</tr>
<tr>
<td>$\text{C} - \text{S}$</td>
<td>BAF = 0.172</td>
</tr>
<tr>
<td>$\text{C} - \text{Cl}$</td>
<td>BAF = 0.200</td>
</tr>
<tr>
<td>$\text{C} - \text{Br}$</td>
<td>BAF = 0.277</td>
</tr>
</tbody>
</table>

Table 2 Quantification of attractive inductive effect by the BAF index

In the second example, the fluorine atoms intensifies the existing C-O bond polarization in the methanol. Once more, the more fluorine atoms the poorer in electrons the carbon atom, which shifts the ED contragradience towards the oxygen atom. The last example focus on the C-H bond. First, as can be seen from Table 2, this bond
is only weakly polarized (BAF=0.229 in CH₄) compared to C-O (BAF = 0.339 in CH₃OH). This polarization is initially directed towards the carbon atom. Fluorinating the methane results in a carbon atom depleted in electrons (more depleted than hydrogen, more distant from fluorine atoms), which favours ED contragradience (clash) to occur further away from the carbon atom, although the bond strength polarization still remains in the direction of the carbon. It is worthy to note that, from the above three examples, the most polarizing effect is observed for the C-O bond (polarized by at most 0.238), the less sensitive bond being C-H (polarized by at most 0.068).

VII. Future prospects and outlook

- Accelerating the code is one of our next concerns. This will be addressed in two ways. First, in addition to the OpenMP current possibility, an MPI version of IGMPlot is under progress, which will extend the parallel execution on several nodes on a computational center. This work is made in collaboration with the Irish Centre for High-End Computing (https://www.ichec.ie). In addition, the algorithm will be improved such to perform a pre-filtering of the grid points to reduce the number of calculations.

- Deep-learning possibilities will be leveraged to improve the IGM energy scoring function

- Performing IGM calculations for excited has still to be tested

- Output files in the ParaView format will be generated in order to make use of the ParaView application to visualize isosurfaces from extremely large datasets and for cylindrical grid as well (bonding feature).

- Using libraries of Extremely Localized Molecular Orbitals (ELMO)¹⁰ will allow going beyond the promolecular approximation at lower cost than QM calculations for large biomolecular systems.

3 Miguel Ponce-Vargas, Corentin Lefebvre, Jean-Charles Boisson, Eric Hénon, J. Chem. Inf. Model. 2020, 60, 1, 268 doi: https://doi.org/10.1021/acs.jicim.9b00106
7 AIMAll (Version 19.10.12), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2019 (aim.tkgristmill.com)