FILTERING & SAMPLING in 2D

1) You have to collect data:
   `JKCS4_collect XTB`
   File `collectionXTB.txt` is formed containing 3 columns: XYZ-file-name Rg El.E.

2) You can visualize the collected data:
   ```
   gnuplot
   plot 'collectionXTB.txt' u 2:3
   ```
   (zoom in by mouse might be required)

3) Redundant files are automatically removed (uniqueness filtering), see file `resultsXTB.dat`
   UNIQUENNESS: (you can skip)
   Two files with energy difference less than 0.001 hartree and Rg difference 0.01 are assumed to be the same -> one of the files is removed
   If you are not satisfied with filtering thresholds, you can remake resultsXTB.dat by:
   ```
   JKCS7_filter resultsXTB.dat –u –u1 2 –u2 4
   ```
   // this mean thresholds for Rg - 0.01 and for El.E.- 0.0001

4) **FILTERING:**
   If you need to remove out-lying structures use
   ```
   JKCS7_filter resultsXTB.dat –rgm 3 –d 8
   ```
   // this filter out structure with relative energy higher than 8 kcal/mol
   // please use also -rgm 3 if use plan to use sampling/selection (see -help)

5) **SAMPLING/SELECTION:**
   There will probably still remain a lot of structures in `resultsXTB_FILTERED.dat`.
   Thus, you can use uniform sampling from them to cover PES as best as possible for lower comp. cost
   (Oh, yes! You might lose global minimum structure, but you should not be so far from it)
   Join filtering and sampling:
   ```
   JKCS7_filter resultsXTB.dat –rgm 3 –d 8 –s 100
   ```

6) Visualize the result:
   ```
   gnuplot
   p 'resultsXTB.dat' u 2:3, 'resultsXTB_FILTERED.dat' u 2:3 pt 5 ps 2
   ```
   (zoom in by mouse might be required)
   // green points indicate selected points

7) `JKCS5_runDFT` will now take structures saved in `resultsXTB_FILTERED.dat`

FILTERING & SAMPLING in 3D

1) You have to collect data also with dipoles:
   `JKCS4_collect XTB -dip`
   File `collectionXTB.txt` is formed containing 4 columns: XYZ-file-name Rg El.E. dipoles

2) You can visualize the collected data:
   `gnuplot`
   `splot 'collectionXTB.txt' u 2:3:4`
   (zoom in might be required … set yrange [0,10] …)

3) Redundant files are automatically removed (uniqueness filtering), see file `resultsXTB.dat`
   UNIQUENNESS: (DO NOT skip)
   Two files with energy difference less than 0.001 hartree, Rg difference 0.01 and dipole difference 0.001 Debye are assumed to be the same -> one of the files is removed
   We are not satisfied with filtering thresholds, so let us remake resultsXTB.dat by:
   `JKCS7_filter resultsXTB.dat -u -u1 2 -u2 3 -u3 1`
   // this mean thresholds for Rg - 0.01 and for El.E.- 0.001 and dip - 0.1

4) **FILTERING:**
   If you need to remove out-lying structures use
   `JKCS7_filter resultsXTB.dat -rgm 3 -d 8`
   // this filter out structure with relative energy higher than 8 kcal/mol
   // please use also -rgm 3 if use plan to use sampling/selection (see -help)

5) **SAMPLING/SELECTION:**
   There will probably still remain a lot of structures in `resultsXTB_FILTERED.dat`.
   Thus, you can use uniform sampling from them to cover PES as best as possible for lower comp. cost
   (Oh, yes! You might lose global minimum structure, but you should not be so far from it)
   Join filtering and sampling: (to sample 100 structures)
   `JKCS7_filter resultsXTB.dat -rgm 3 -d 8 -c3 4 -s 100`

6) Visualize the result:
   `gnuplot`
   `splot 'resultsXTB.dat' u 2:3:4, 'resultsXTB_FILTERED.dat' u 2:3:4 pt 5 ps 2`
   (zoom in might be required … set yrange [0,10] …)
   // green points indicate selected points

7) `JKCS5_runDFT` will now take structures saved in `resultsXTB_FILTERED.dat`

ALL TOGETHER IS ALSO POSSIBLE:
`JKCS7_filter resultsXTB.dat -u -u3 1 -rgm 3 -d 8 -c3 4 -s 100`