Introductory session Basic Formats

- Opening *Dynamo*
- Getting help
- Manipulating particles and metadata

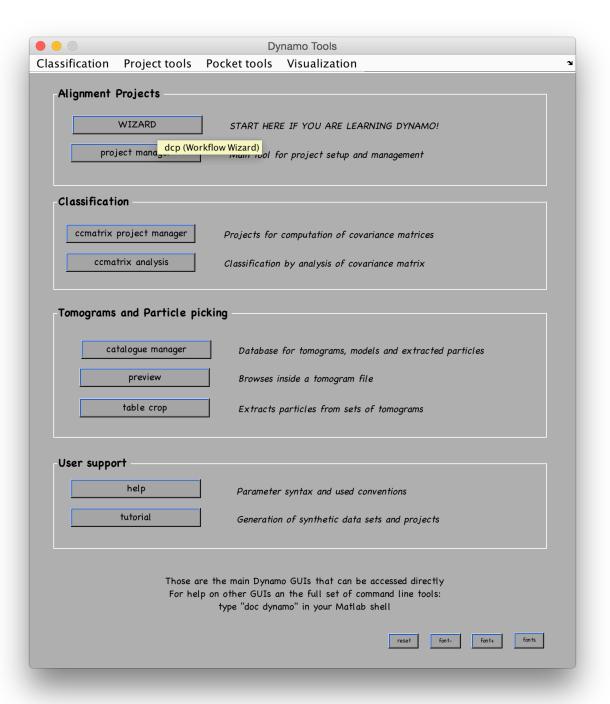
General GUI

In matlab:
>>dynamo
Or
>>dqui

In the standalone
console:
Dynamo >> qui

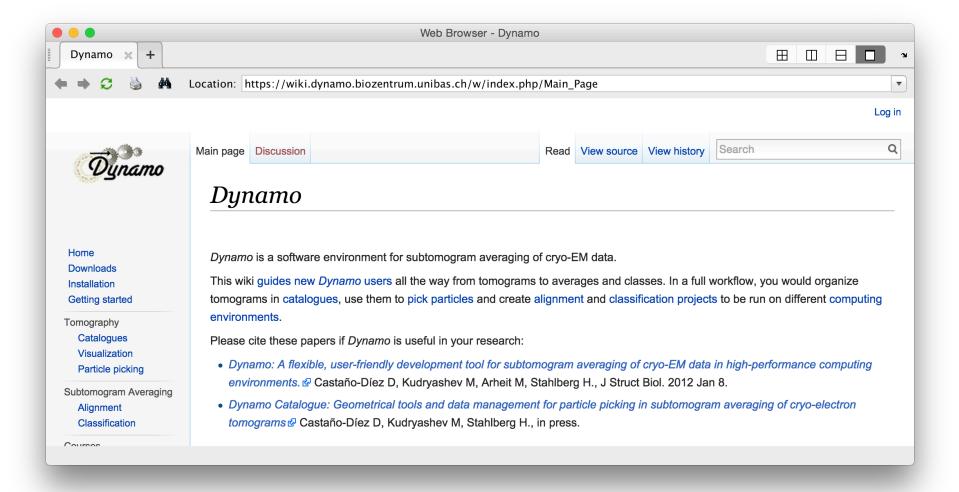
All the modules are independently accessible through the command line.

Keep the cursor on the corresponding button to find out which command Invokes the corresponding functionality.



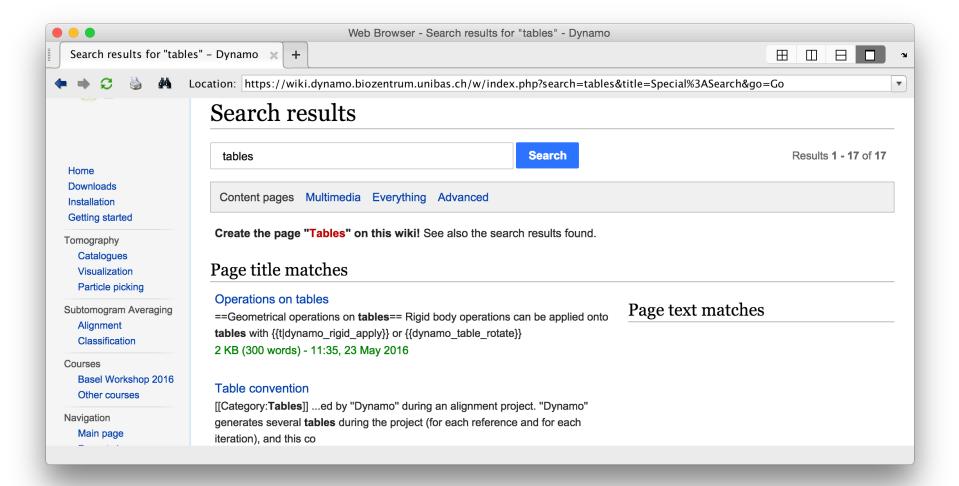
Online help is available through the wiki:

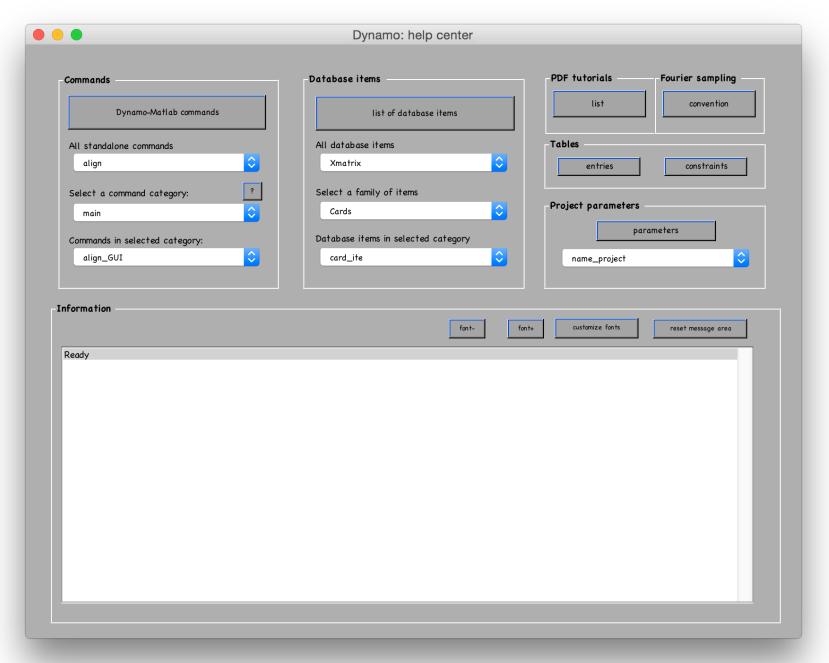
>> dwiki



... or with arguments

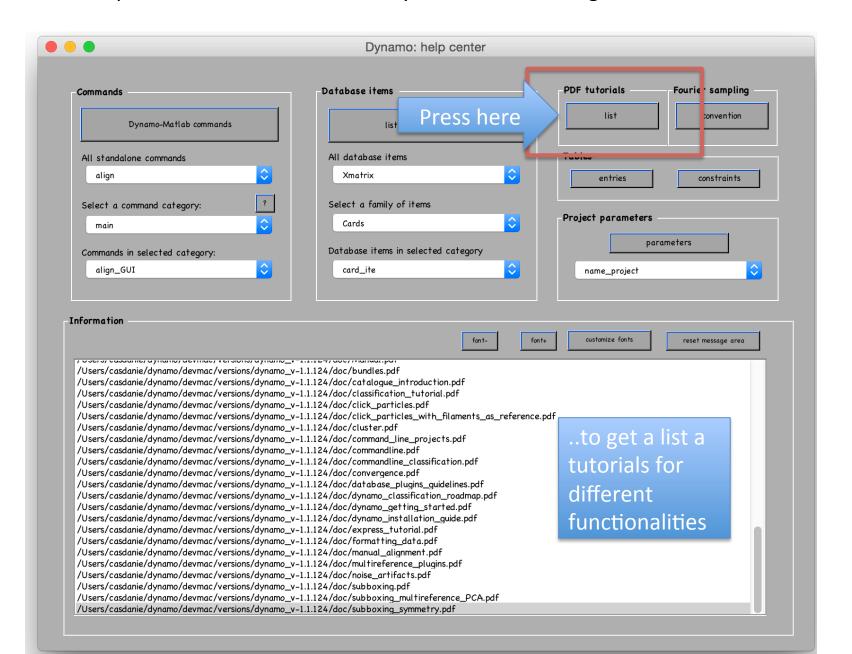
>> dwiki table





For instance the help center an also be invoked by the command line with dhelp

The help center offers access to many tutorials, including this one:

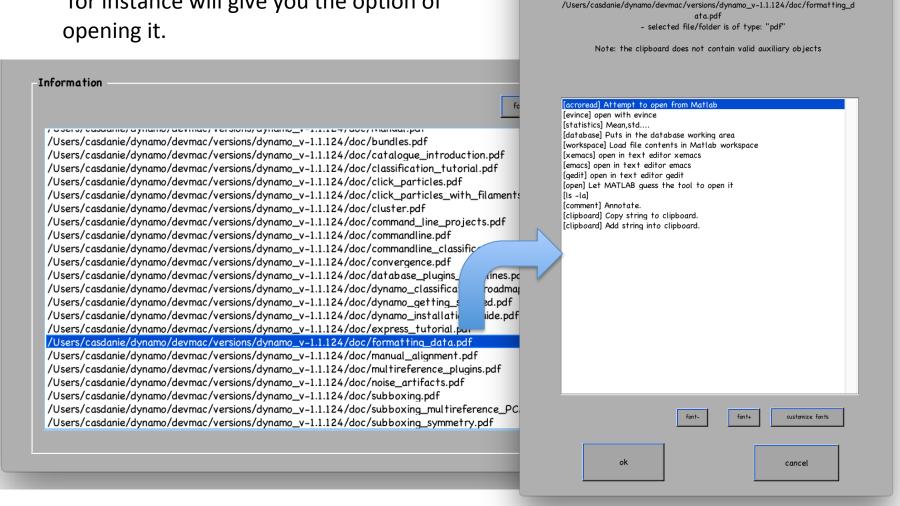


In any Dynamo GUI, file and folder names are selectable [secondary click or ctrl+main click on a Mac] Selecting an object will get you a menu of options that will vary for the different objects.

Dynamo Linker:

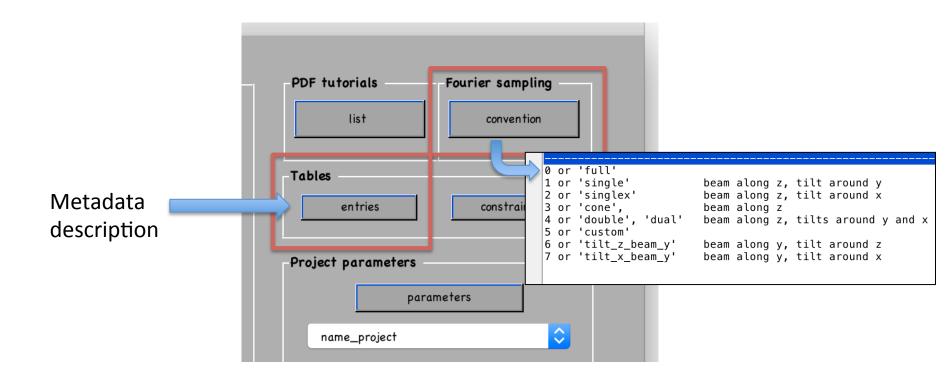
Selection:

Secondary-clicking on a selected pdf, for instance will give you the option of



Dynamo uses several conventions: metadata, data, fourier sampling

In the Help Center is the place to look when you need a quick reminder.

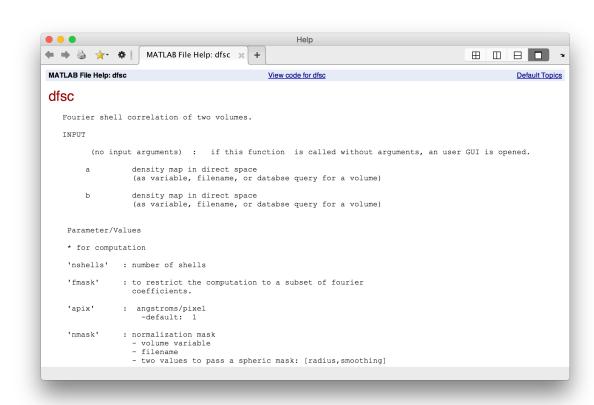


Besides the tutorials and the Help Center, every Dynamo command provides help about its syntax.

You'll need this when you are not operating through the GUI.

>>doc dfsc Shows the help in an interactive window.

>>help dfsc Prints the output directly on the screen.



But... how to know which commands are available in Dynamo?

In general, the different tutorials will introduce you to the most important commands in each functionality area.

To find them on a more systematical way, a selection of commands is classified through areas of functionality:

>>dapropos
Shows the names of the areas:

You can select then one area to show a selection of commands related to one functionality:

>>dapropos data
Shows the main classification related commands:

```
>> dapropos
Available families of Dynamo commands:
  * main
                    GUI utilities
                    displaying volumes and images
  * display
  * project
                    managing Dynamo projects
  * bundle
                    managing groups of Dynamo projects
  * compaction
                    Using sets of computed averages as particle sets.
  * sequence
                    Sequential alignment of particles.
  * subboxing
                    Extraction of subboxes out of subtomograms.
  * data
                    manipulation of particles and data sets
  * classification
                    tools for classification of aligned particles
  * seeds
                    generic manipulation of "seed" files (masks, templates,...)
  * table
                    creation and manipulation of table files
```

Basic formats

There are three basic object types that will appear continuously while performing subtomogram averaging with Dynamo:

metadata: tables

• data: data folders

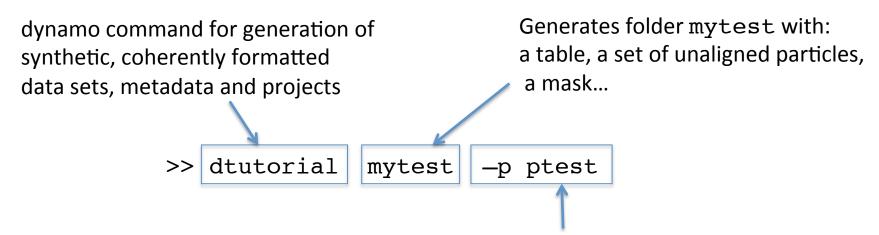
computation: projects

(we are assuming in this part that 3D particles are already cropped from out of the tomograms, and that we are interested in aligning and averaging a set of subtomograms)

We are going to get familiar with them running an example on a synthetic data set.

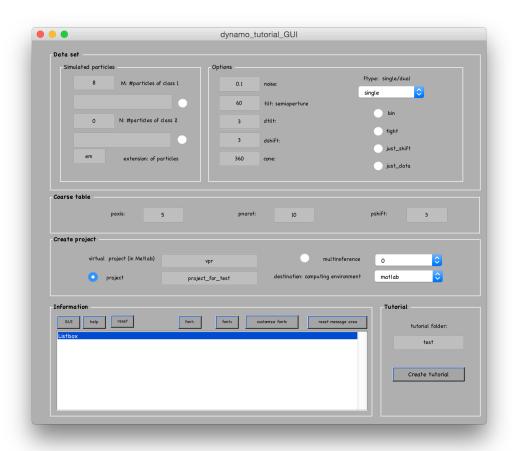
We will generate synthetically a folder that contains all elements that would appear during a real work session.

So, go to an empty folder in your file system and generate a tutorial with the command:



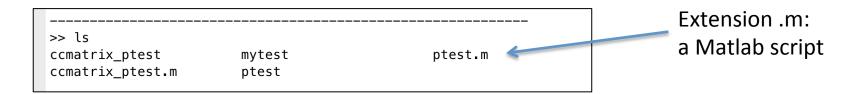
Creates a project using the elements generared in mytest.

Side note: notice that dtutorial can be invoked without arguments. This generates a GUI that allows entering the main parameters. Command line use normally provides more functionality.



Many of the main commands in Dynamo work this way.

After some moments, *Dynamo* will have create some contents in the folder. We can explore it with the system/Matlab command Is



Also notice the *Dynamo* command dsummary which reports the *Dynamo*-relevant contents found in a folder:

Let's see what is inside what Dynamo marks as a tutorial folder

```
>> dsummary . on
Gathering a summary of Dynamo contents in folder "."
projects
                     : 2
                       ./ccmatrix_ptest
                       ./ptest
multireference
                     : 0
 ccmatrix
                       ./ccmatrix_ptest
data folders
subboxing folders
bundles
compaction folders
compace folders
tutorial folders
                     : 1
                       ./mytest
  [ok] summary
```

```
>> ls mytest
coarse tbl
data
fmask.em
folder_multireference
info_tutorial_parameters.doc
initial.tbl
mask.em
real.tbl
real.tbl
template.em
```

data folder is a directory where the subtomograms have been stored as individual files

.tbl files are tables (metadata objects) a real.tbl that actually algn the particles, provided for demonstration purposes

.em files are 3d volumes

Files suggested to be used as initial reference, mask, fourier mask, etc...

DATA folders

```
ok
>> ls mytest/data
particle_00004.em particle_00020.em particle_00034.em particle_00077.em

particle_00014.em particle_00024.em particle_00043.em particle_00077.em
```

Dynamo needs the particles named after the convention:

```
<folder_name>/particle_XXXXX.em.
```

The integer that identifies a particle is called **tag** in the Dynamo documentation.

- Any number of integers, left padded with zeros.
- Extensions .mrc, .spi, .vol and .em are accepted
- For very large data sets (>100k subtomograms), other formats can be used: dData

The formatting data tutorial provides more exhaustive information about ways to ease the generation of this data folders

dinfo will print a quick check on the contenst of a data folder

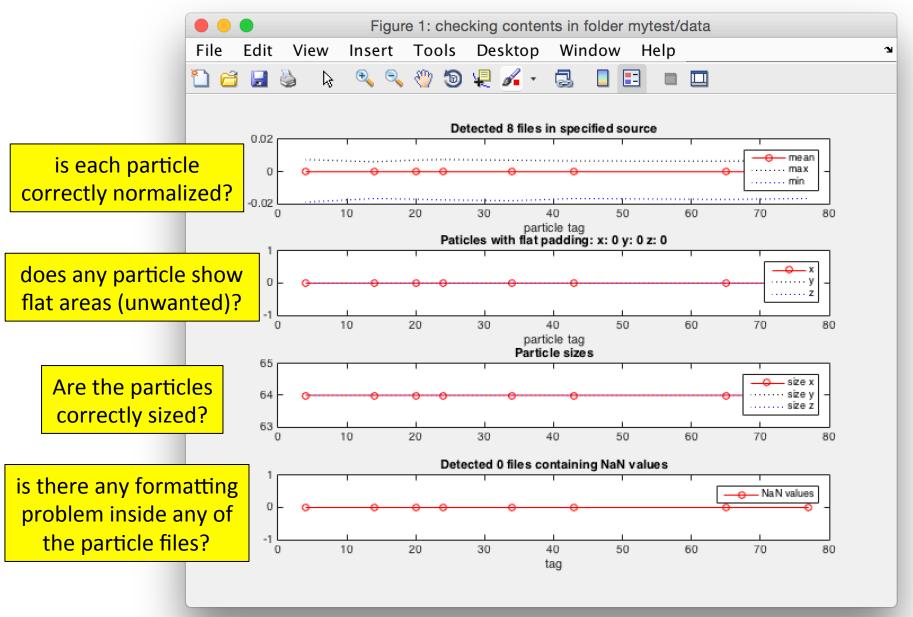
```
>> dinfo mytest/data
"mytest/data" is of type "data folder"
               asproject: 0
                                              Number
             folder_name: 'mytest/data'
                                            of particles
                 correct: 1
                    tags [4 14 20 24 34 43 65 77]
                      Mb: 2
                       1: 64
             source_type: 'dynamo_folder'
               extension: 'em'
                 padding = 5
                                            Size
                    size: [64 64 64]
                                        of particles
                   exist: 1
                  Mb_all: 16
    N_accompanying_fmask: 0
   [ok] data info
```

dinfo is a general command that will work on other types of objects: tables, projects, etc.

You can check the different commands related to management of data folders

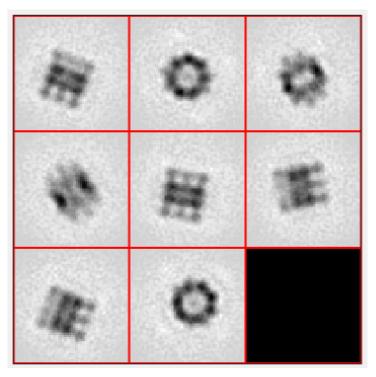
```
>> dapropos data
  data utilities:
                              : Manual alignement of data to produce initial coarse orientations
      align manual
      data
                              : Lists valid data folders in the current directory
      data_browse
                              : Light GUI to browse large sets of particles
                              : Thorough check on all the volumes present in a data folder
      data check
     data_format
                              : formats the names of a set of files to use them as Dynamo data source
                              : GUI version of data format
      data format GUI
     data import
                              : Some utilities for data importation from other formats.
     data_info
                              : gets the basic info on a data folder or generic source of data
                              : Merges two data sets, possibly also with tables.
      data merge
                              : Normalizes the particles in a data folder.
      data_normalize
      data slider
                              : Light GUI to browse large sets of particles one by one.
                              : Extracts subboxes from the data set for region focused alignment.
      data subboxing
     gallery
                              : Interactive visualization of large sets of volumes
     modellist crop
                              : Crops particles from a set of tomograms with the geometry defined in a model
                              : Crops subparticles from a tomogram.
      table crop
                              : Retrieves a particle from a data set.
     tag2particle
 Some related pdf tutorials [available through docpdf]:
      classification_tutorial
      formatting_data
```

dynamo_data_check (shortened to ddcheck) for instance, scans for the contents of e particle inside the data folder.



A data folder name is used by several *Dynamo* commands (as dslices in this example) to refer collectively to all the 3d particles in a data folder:

"project [j] all slices [*] along the z direction"



In these commands, the flag 'tags' is normally Used to refer to a subset of particles

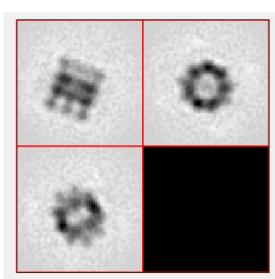


Table files

Table files have several uses.

In their simple form, they are related to a data folder.

A table is simply a matrix that contains

- N rows [for N particles]
- P columns [for P described properties]

The convention that maps column numbers to properties is predefined, and can be consulted in the data center (or with the dthelp command)

```
>> dthelp
                    tag of particle file in data folder
    : tag
2 : aligned
                    value 1: marks the particle for alignment
    : averaged
                    value 1: the particle was included in the average
4
    : dx
                    x shift from center (in pixels)
    : dy
                    y shift from center (in pixels)
    : dz
                    z shift from center (in pixels)
    : tdrot
                    euler angle (rotation around z, in degrees)
                    euler angle (rotation around new x, in degrees)
    : tilt
```

Column 1 is the "tag" that identifies individual particle numbers.

This is the link between a row in a table and a file in a table folder

A table is normally stored as a text file, So it can be inspected with any usual tool

```
type mytest/initial.tbl
or
!cat mytest/initial.tbl
In the standalone version
```

```
>> type mytest/initial.tbl

34 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

4 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

65 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

14 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

24 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

77 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0

20 1 0 0 0 0 0 0 0 0 0 1 -60 60 0 0
```

This row in the table will store the metadata for a particle with tag 4' in a data folder

```
>>
>> ls mytest/data
particle_00004.em particle_00020.em particle_00034.em particle_00065.em
particle_00014.em particle_00024.em particle_00043.em particle_00077.em
```

Which information is coded inside a table?

Most important are alignment parameters:

Columns 4 to 6: shifts

```
4 : dx x shift from center (in pixels)
5 : dy y shift from center (in pixels)
6 : dz z shift from center (in pixels)
```

Columns 7 to 9: **rotation** (Euler angles)

```
7 : tdrot euler angle (rotation around z, in degrees)
8 : tilt euler angle (rotation around new x, in degrees)
9 : narot euler angle (rotation around new z, in degrees)
```

When you start a project for alignment in Dynamo, you need to provide a table with initial guesses for these parameters.

Leave them to zero if you don't have any a priori informatio

The final result of a *Dynamo* alignment project will be a table that contains the alignment parameters for each particle.

The next most important information that you need to provide is the "fourier type"

Column 13 defines the "ftype" (fourier type): a description of the geometry. Type 1 is the most usual one: electron beam comes along z, tilt axis is in y. For this type, columns 14 and 15 describe the extent of the missing wedge.

```
>> type mytest/initial.tbl

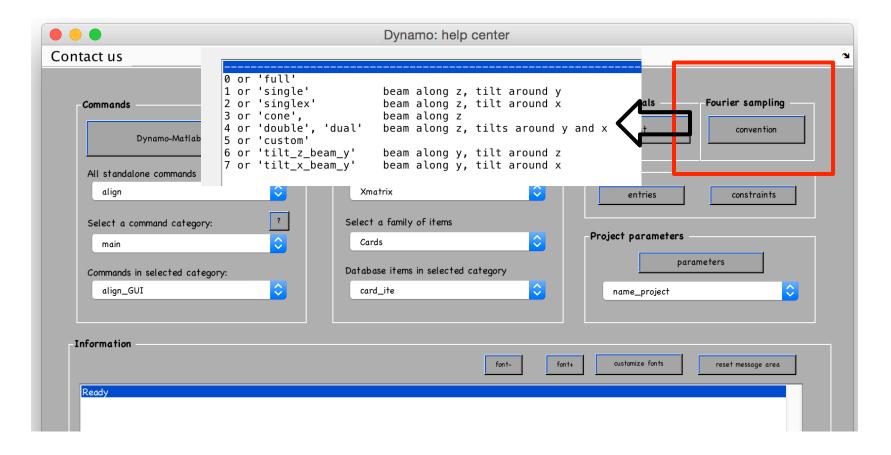
34 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0 0
4 1 0 0 0 0 0 0 0 0 0 0 1 -60 60 0 0
65 1 0 0 0 0 0 0 0 0 0 1 -60 60 0 0
43 1 0 0 0 0 0 0 0 0 0 1 -60 60 0 0
14 1 0 0 0 0 0 0 0 0 0 1 -60 60 0 0
```

Here, each particle is noted to be of ftype 1, with a missing wedge between -60 and 60 deg.

For advanced users:

it is possible to define your own fourier type (ftype =5), as long as you pass a file in the data folder for each particle.

The coding of the fourier type is one of the conventions listed in the dhelp Help Center of Dynamo



Other entries are provided by *Dynamo* during the alignment (use dhelp)

1	0 :	CC	Cross	correlation	coefficient	
1						

... related to cropping particles from tomograms...

```
20: tomo tomogram number
24: x x coordinate in original volume
25: y y coordinate in original volume
26: z z coordinate in original volume
```

... related to classification tasks....

```
41 : eig1 "eigencoefficient" #1
42 : eig2 "eigencoefficient" #2
```

... or foreseen for user annotations....

```
21: reg for arbitrary region assignations
23: annotation arbitrary annotations
```

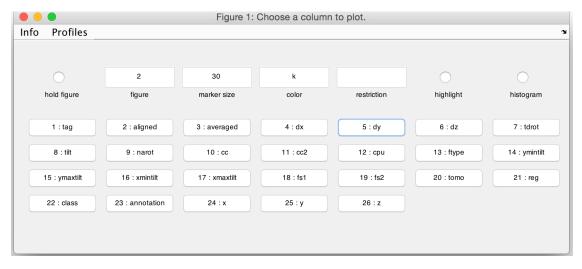
Check on dapropos table to see the commands to manage tables. Probably, the most useful ones will be:

dtgrep selects subsets of particles inside a table using different criteria

dinfo

prints summaries of the content of a table for each column

dtshow To quickly plot contents inside a table (dtview for a more complex browser)



Combining TABLE and DATA

Tables are interesting as they express the alignment of a data set.

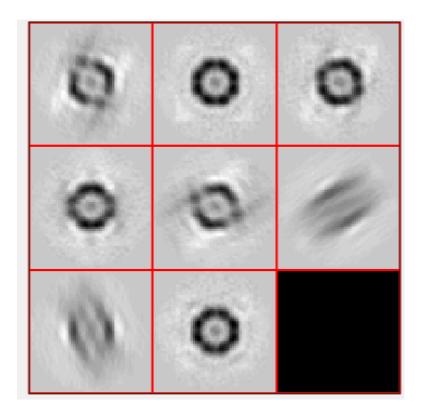
The dtutorial command provided use with a real.tbl table, that contains the actual alignment parameters of the provided synthetic data set.

```
>> dinfo mytest/real.tbl
"mytest/real.tbl" is of type "table"
 table file: mytest/real.tbl
       size
                            8 26
       NaNs
                                          Shifts and alignments are initialized
 COLUMN
                                                     for all the particles
     ] marked for alignment: 8
       included in average:
                             initialized: 8
                                                   mean: -0.94
                                                               std: 2.18
                             min: -3.18 max: 3.30
                             min: -4.63 max: 1.24 mean: -2.30 std: 2.29
                             min: -4.51 max: 4.16 mean: -0.95 std: 2.76
                             initialized: 8
                          : min: -169.36 max: 152.91 mean: -14.72 std: 107.26
                             min: 14.42 max: 168.00 mean: 73.92 std: 49.38
         * narot
                             min: -81.29 max: 150.23 mean: 17.01 std: 82.37
```

So, let's see how the data set looks when the table is used to align the particles

We use the previously shown command dslices, adding: a table (-t mytest/real.tbl) and the order of use it to align the particles (-align on)

dslices mytest/data -jz * -t mytest/real.tbl -align on



Now Dynamo shows all [*] the sections of each particle along the [z] direction after alignment

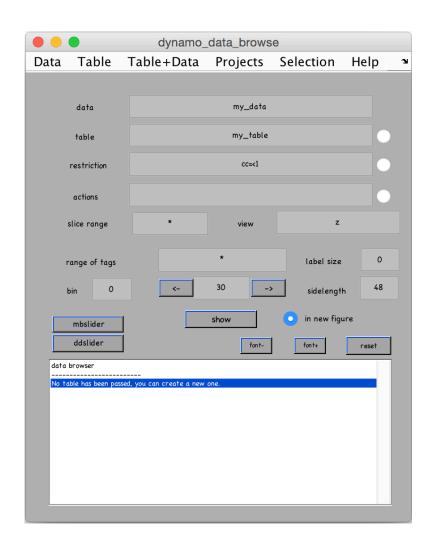
Notice how the missing wedge Works differently for different particles depending on their original orientation The most often way of looking at aligned particles is with the command ddbrowse.

ddbrowse is a lightweight GUI for exploration of subsets of particles in a data set.

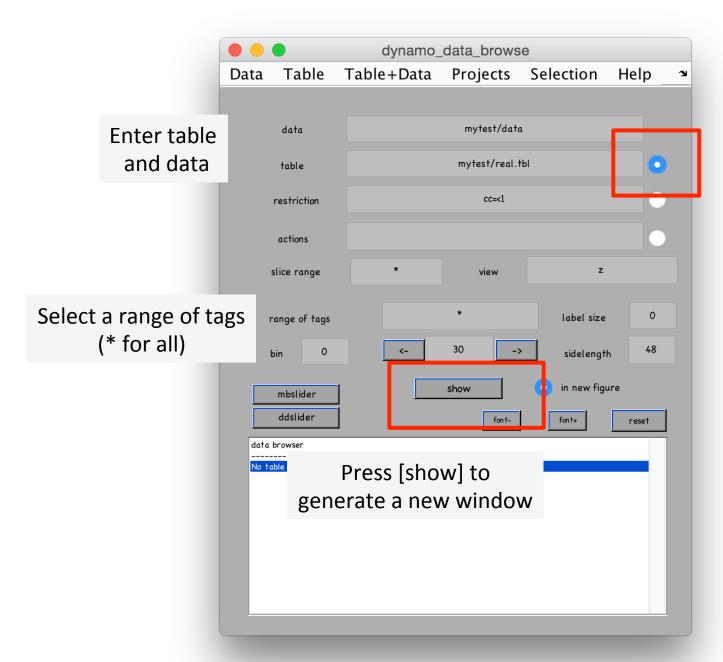
Many Dynamo GUIs will launch instances of this command when you are running preset workflows, but you can also launch it independently.

>> ddbrowse

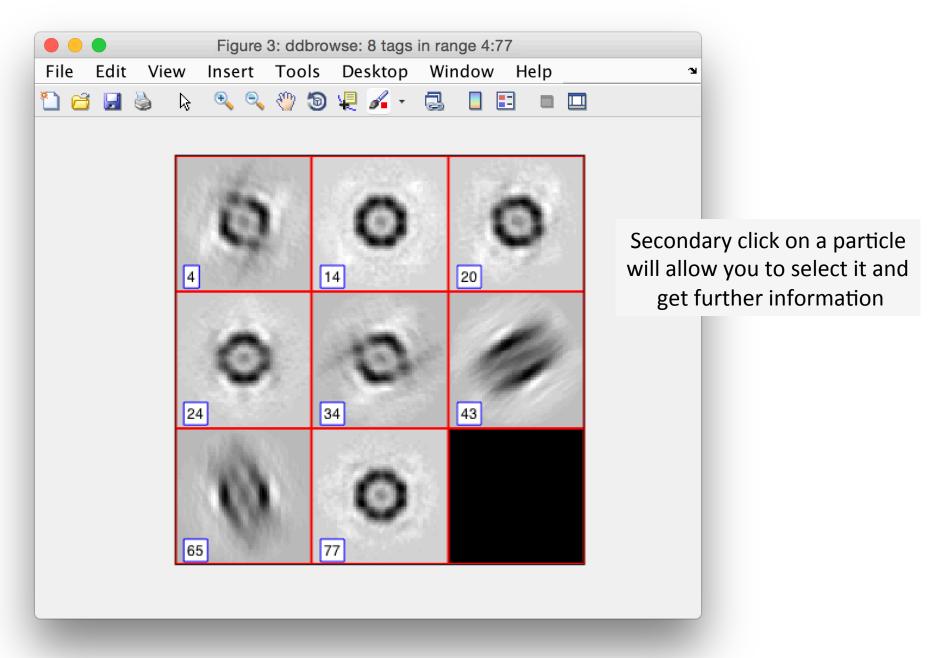
or you can launch it



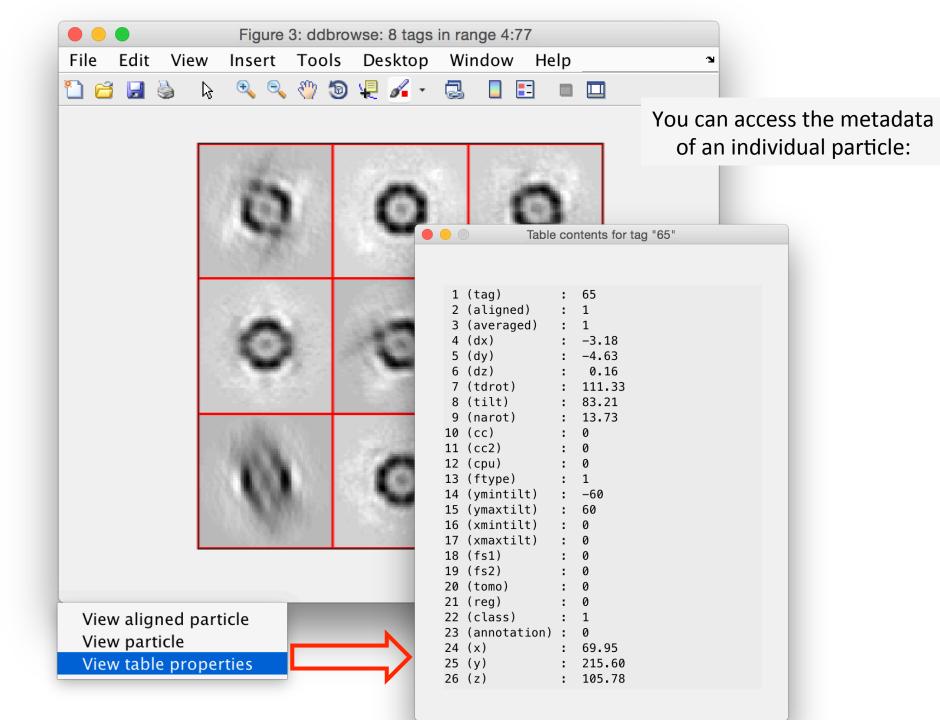
>> ddbrowse -d mytest/data -t mytest/real.tbl

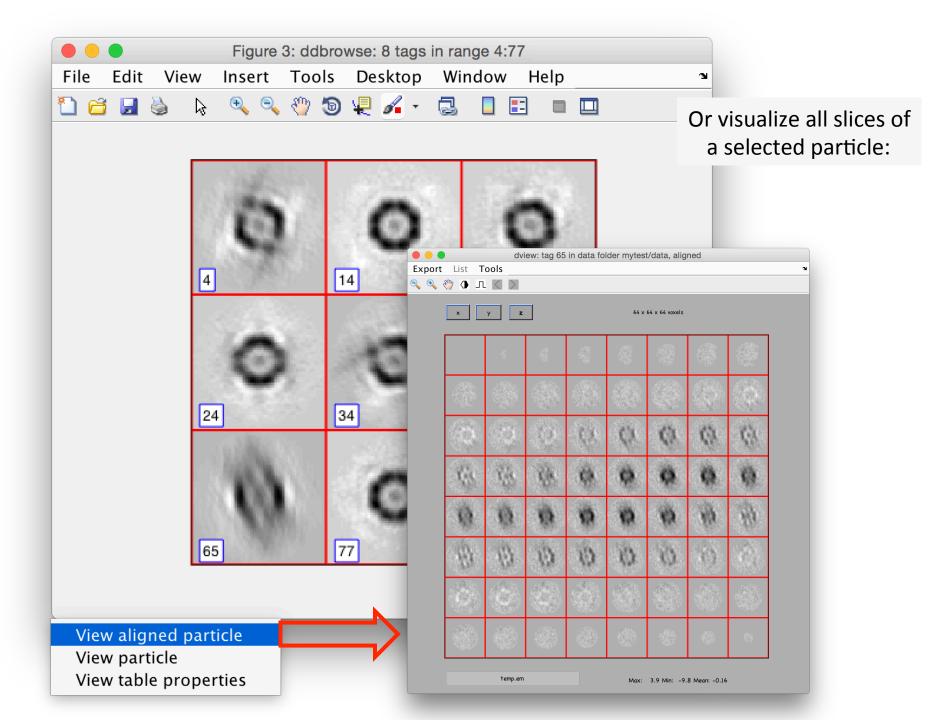


Switch on the table if you want to see the aligned particles (otherwise they will be shown as contained in the raw files).

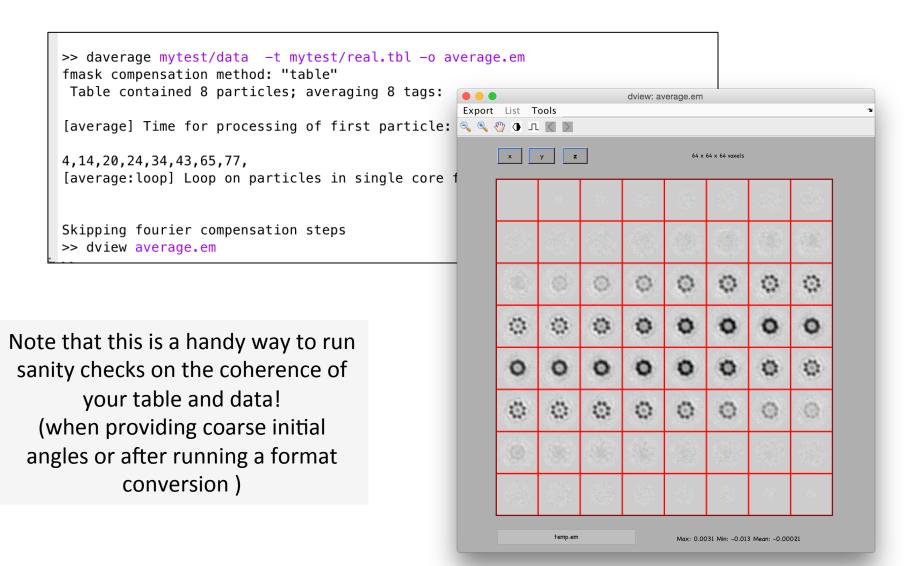


Note: you can get labels with the tag number with the [label size] entry





Obviously the final goal of an alignment is creating an average from the aligned data set. Our last command is daverage.



Suggested further material

Deepening into using tables and data folders
PDF tutorial: commandline.pdf
(available through dynamo_docpdf)

Creating alignment projects:

PDF tutorial: introduction2AlignmentProjects.pdf