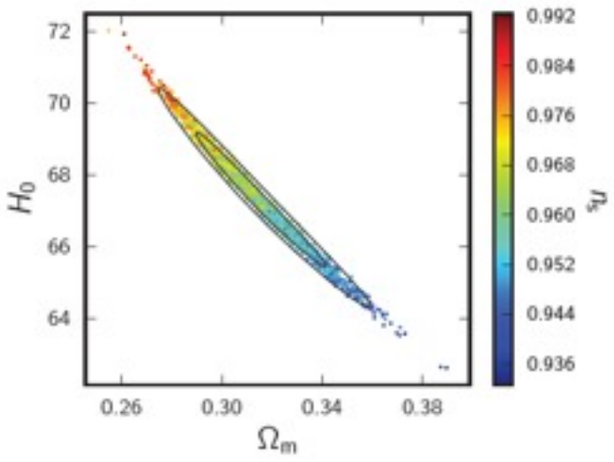
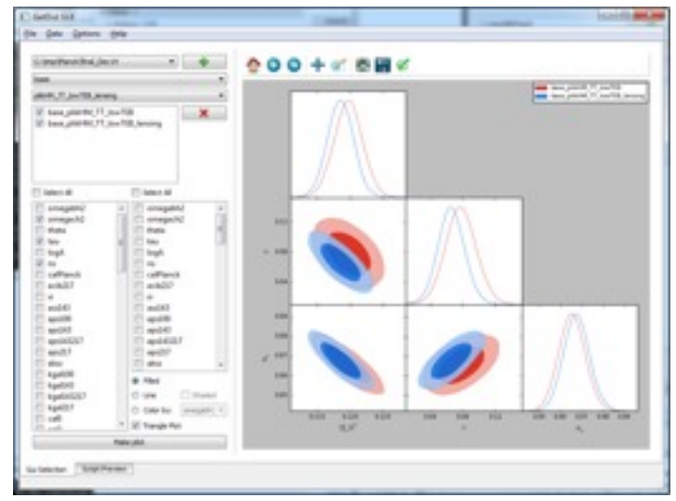




Introduction to CosmoMC

Part III: Analysis - GetDist & GetDist GUI

Antonio J. Cuesta
Institut de Ciències del Cosmos - Universitat de Barcelona

Institut de Ciències del Cosmos

EXCELENCIA MARÍA DE MAEZTU

UNIVERSITAT DE BARCELONA

Dept. de Física Teórica y del Cosmos, Universidad de Granada, 1-3 Marzo 2016

Outline of Part III

- The analysis code GetDist
- GetDist products: Margestats vs likestats
- How to generate plots with GetDist
- Using Planck public chains
- Interactive plots with GetDist GUI

GetDist: Analysis of MCMC chains

- When your chains are **mostly converged**, you can now retrieve your cosmological **bounds** (there is no need to wait until the code stops running)

- A typical execution of **GetDist** will look like this:

```
[acuesta@nova0 cosmomc2015jul]$ ./getdist distparams.ini chains/lcdmbaosn
 skipped unused params: omegak mnu nnu yhe Alens nrun r r02
 reading chains/lcdmbaosn_1.txt
 reading chains/lcdmbaosn_2.txt
 reading chains/lcdmbaosn_3.txt
 reading chains/lcdmbaosn_4.txt
 reading chains/lcdmbaosn_5.txt
 reading chains/lcdmbaosn_6.txt
 reading chains/lcdmbaosn_7.txt
 reading chains/lcdmbaosn_8.txt
Number of chains used =             8
var(mean)/mean(var), remaining chains, worst e-value: R-1 =       0.00666
RL: Thin for Markov:          30
RL: Thin for indep samples:   31
RL: Estimated burn in steps: 120 (50 rows)
mean input multiplicity =    2.39603247549020
using        65280  rows, processing           92  parameters
Approx indep samples:          5046
Best fit sample -log(Like) =    6822.39800000000
mean(-Ln(like)) =    6835.30173439549
-Ln(mean like)  =    6830.18592659466
Warning: sharp edge in parameter chi2_JLA - check limits[chi2_JLA] or limits86
Warning: sharp edge in parameter chi2_6DF - check limits[chi2_6DF] or limits87
```

Syntax: ./getdist distparams.ini chains/root

parameters defined in distparams.ini but not in the chains

number of chains read (check that these are all your chains)

MCMC Convergence status

number of parameters (cosmo+nuisance+der)

warnings about non-converged parameters

GetDist products

- The following files are generated when you run GetDist (the first time you need to create the folder **plot_data**, not created by default !!)

- **root.likestats** and **root.margestats**: parameter constraints (see next slide)

- **root.corr** and **root.covmat**: these are the correlation matrix R_{ij} and the covariance matrix C_{ij} of the chain parameters, where $R_{ij}=C_{ij}/\mathrm{sqrt}(C_{ii}C_{jj})$. Note the file root.covmat can be used in params.ini to help in faster convergence

- **root.converge**: contains information on the convergence of your chains (e.g. R-1 for each parameter, chain auto-correlations, etc)

- all the information to make 1D, 2D, 3D, and triangle plots will be there. Also **root.py** and **root_2D.py** will be created to make matplotlib plots (cf.later)

Margestats

- This is probably the **most valuable** product of your chain. It contains the **mean** value of each parameter (cosmological, nuisance, and derived ‘*’), its **standard deviation**, its **lower and upper limits** for the **68%, 95%, 99%** confidence levels Also if limits are **1-tail** (‘<‘, ‘>’) or **2-tail**, and even **chi2** from each likelihood

- It is useful to write a **script** to parse this file to make a **LaTeX table** automatically

Marginalized limits: 0.68; 0.95; 0.99

parameter	mean	sddev	lower1	upper1	limit1	lower2	upper2	limit2	lower3	upper
omegabh2	0.2227594E-01	0.1409804E-03	0.2213470E-01	0.2241652E-01	two	0.2200101E-01	0.2255144E-01	two	0.2191544E-01	0.226
omegach2	0.1192956E+00	0.1068506E-02	0.1182545E+00	0.1203491E+00	two	0.1171635E+00	0.1213484E+00	two	0.1164438E+00	0.122
theta	0.1040831E+01	0.2995374E-03	0.1040534E+01	0.1041130E+01	two	0.1040247E+01	0.1041420E+01	two	0.1040057E+01	0.104
tau	0.8200741E-01	0.1643758E-01	0.6557802E-01	0.9862151E-01	two	0.4956239E-01	0.1138723E+00	two	0.4020595E-01	0.124
logA	0.3097935E+01	0.3246072E-01	0.3065291E+01	0.3130707E+01	two	0.3034228E+01	0.3160888E+01	two	0.3013767E+01	0.318
ns	0.9658500E+00	0.4063676E-02	0.9618321E+00	0.9698272E+00	two	0.9579955E+00	0.9740912E+00	two	0.9555469E+00	0.976
calPlanck	0.1000465E+01	0.2516048E-02	0.9979804E+00	0.1002947E+01	two	0.9954956E+00	0.1005439E+01	two	0.9939398E+00	0.100
acib217	0.6359261E+02	0.6577819E+01	0.5697960E+02	0.7011801E+02	two	0.5088221E+02	0.7649609E+02	two	0.4702681E+02	0.807
xi	0.5220534E+00	0.2840235E+00	0.0000000E+00	0.1000000E+01	none	0.0000000E+00	0.1000000E+01	none	0.0000000E+00	0.100
asz143	0.5399273E+01	0.1871846E+01	0.3556131E+01	0.7442447E+01	two	0.1661532E+01	0.8952850E+01	two	0.7504497E+00	0.965
aps100	0.2590954E+03	0.2746669E+02	0.2316349E+03	0.2864320E+03	two	0.2053967E+03	0.3134392E+03	two	0.1890985E+03	0.331
aps143	0.4325286E+02	0.7687977E+01	0.3552913E+02	0.5095578E+02	two	0.2797599E+02	0.5791646E+02	two	0.2349064E+02	0.619
aps143217	0.4026165E+02	0.9948208E+01	0.3006222E+02	0.5063956E+02	two	0.2158706E+02	0.5957445E+02	two	0.1655117E+02	0.641
aps217	0.9815534E+02	0.1100181E+02	0.8722243E+02	0.1092326E+03	two	0.7626975E+02	0.1192058E+03	two	0.6946944E+02	0.125
aksz	0.3160785E+01	0.2332050E+01	0.0000000E+00	0.4034485E+01	>	0.0000000E+00	0.7749367E+01	>	0.0000000E+00	0.100

Remember that the difference between the likelihood and the marginalized likelihood is an integral over all the other parameters

Likestats

- This contains less relevant, but still sometimes useful, information.
  It is NOT what you usually would quote in a paper

- As opposed to the margestats file, which contains the mean and marginalized bounds for each parameter, this contains the **bestfit** value and the **bounds** of the **mean likelihood** distribution (not the posterior!) for each parameter

- Remember the bestfit value is better computed by **action=2** anyway

```
 Best fit sample -log(Like) =    6473.21100000000
 mean(-Ln(like)) =    6485.79215678916
 -Ln(mean like)  =    6480.82803795574

param  bestfit        lower1         upper1         lower2         upper2
    1  0.2217489E-01  0.2174304E-01  0.2277782E-01  0.2171054E-01  0.2282913E-01   \Omega_b h^2
    2  0.1188333E+00  0.1156735E+00  0.1232585E+00  0.1153518E+00  0.1233322E+00   \Omega_c h^2
    3  0.1040975E+01  0.1039816E+01  0.1041871E+01  0.1039671E+01  0.1042056E+01   100\theta_{MC}
    4  0.7869452E-01  0.2493042E-01  0.1376300E+00  0.1574288E-01  0.1376300E+00   \tau
    5  0.3093001E+01  0.2979962E+01  0.3210592E+01  0.2963312E+01  0.3212904E+01   {\rm{ln}}(10^{10} A_s)
    6  0.9666794E+00  0.9525799E+00  0.9811147E+00  0.9509012E+00  0.9853592E+00   n_s
    7  0.1001416E+01  0.9911159E+00  0.1009053E+01  0.9911159E+00  0.1010508E+01   y_{\rm cal}
    8  0.6287429E+02  0.4097516E+02  0.8724035E+02  0.3840288E+02  0.9034937E+02   A^{CIB}_{217}
    9  0.5315644E+00  0.6881551E-04  0.9999398E+00  0.4266870E-04  0.9999699E+00   \xi^{tSZ-CIB}
```

Covariance matrices

- These are important for two reasons:
- First, because they contain the **variances and covariances** of all parameters,
- and second, because they can be **included** in your **params.ini** file in future runs, especially if you think they are a **good proposal** for the expected covariance in that run. In that case it will help the code converge **much faster**
- Also they allow for computation involving covariances, e.g.

$$\mathrm{FoM}=1/\sqrt{\mathrm{Det}(w_0,w_a)}$$

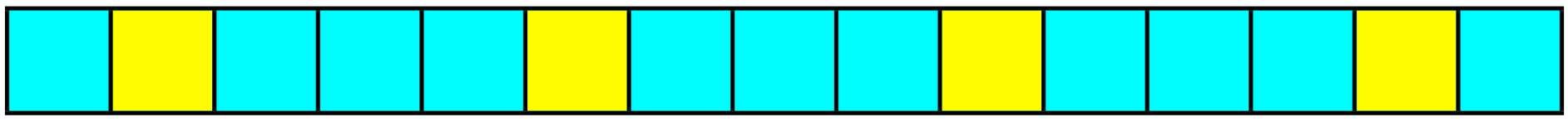
```
# omegabh2 omegach2 theta tau omegak w wa logA ns calPlanck alpha_JLA beta_JLA acib217 xi asz143 aps100 aps143 aps143217
galfTE143 galfTE143217 galfEE217 galfTE100 galfTE100143 galfTE100217 galfTE143 galfTE143217 galfTE217 cal0 cal2
   0.2651846E-07   -0.1573555E-06    0.2262243E-07    0.7493495E-06   -0.2178083E-06    0.2494428E-05   -0.1282951E-04
  -0.1276817E-03    0.3048007E-05    0.3182245E-04   -0.5163862E-03   -0.2623848E-03   -0.2672275E-04    0.5491586E-04
   0.6566699E-07    0.6987950E-07   -0.4491755E-08    0.7394760E-07    0.7627618E-07    0.1544016E-06   -0.7835759E-07
   0.5267135E-08   -0.1162376E-07
  -0.1573555E-06    0.2309257E-05   -0.2034277E-06   -0.8635566E-05    0.2342118E-05   -0.1762828E-04    0.5916942E-04
   0.6098790E-03   -0.1536866E-04   -0.1868635E-03    0.3260676E-02    0.1666958E-02    0.3891061E-03    0.5674583E-05
  -0.7797215E-06   -0.8867264E-06   -0.2319531E-06   -0.9924366E-06   -0.5761369E-07   -0.5707465E-06    0.4339609E-06
  -0.1512159E-07    0.8016534E-07
   0.2262243E-07   -0.2034277E-06    0.1079311E-06    0.8600283E-06   -0.2854636E-06    0.2191388E-05   -0.9364706E-05
  -0.4922939E-04    0.7175632E-06    0.1838311E-04   -0.3280211E-03   -0.1661522E-03   -0.5051060E-04   -0.1266215E-04
   0.5806126E-07    0.6715962E-07   -0.4900255E-07    0.8440468E-07   -0.1549223E-07   -0.1767368E-06   -0.2624164E-07
  -0.3238776E-09   -0.6995240E-08
   0.7493495E-06   -0.8635566E-05    0.8600283E-06    0.3067628E-03    0.6601728E-05   -0.3160861E-03    0.2664801E-02
  -0.1101247E-01    0.2663647E-03    0.1958146E-02   -0.4638594E-01   -0.1679490E-01    0.8509775E-03    0.8120948E-02
   0.2453571E-05    0.3160169E-05   -0.3837760E-05    0.3061417E-05   -0.2408161E-05    0.6865577E-06   -0.5999997E-05
   0.6884432E-07   -0.1136366E-05
  -0.2178083E-06    0.2342118E-05   -0.2854636E-06    0.6601728E-05    0.1239649E-04   -0.2453361E-03    0.1771642E-02
   0.1352245E-02   -0.3317391E-04   -0.2555516E-03    0.6191716E-02    0.2827934E-02    0.5207134E-03   -0.4245678E-03
  -0.7657512E-06   -0.9966254E-06   -0.3259406E-06   -0.9760108E-06    0.9960048E-06   -0.2267493E-05   -0.5780844E-06
   0.8867090E-09    0.1802837E-06
   0.2494428E-05   -0.1762828E-04    0.2191388E-05   -0.3168061E-03   -0.2453361E-03    0.1802248E-01   -0.8222754E-01
  -0.2777562E-01    0.6669691E-03    0.3438096E-02   -0.1048167E+00   -0.3650931E-01   -0.1747068E-02    0.1570753E-01
   0.2373831E-05    0.9164784E-05    0.1955910E-04    0.8339281E-05   -0.3171642E-04    0.3161242E-04    0.5126581E-04
  -0.5532359E-06   -0.3937042E-05
  -0.1282951E-04    0.5916942E-04   -0.9364706E-05    0.2664801E-02    0.1771642E-02   -0.8222754E-01    0.4693822E+00
   0.1704310E+00   -0.3840870E-02   -0.1874376E-01    0.6177674E+00    0.2201816E+00    0.4514734E-02   -0.1098969E+00
  -0.6929873E-05   -0.3774352E-04   -0.2495280E-04   -0.2653066E-04    0.1760635E-03   -0.4047583E-03   -0.2110406E-03
   0.3140856E-05    0.2634835E-04
```

Thinning chains

- Sometimes you need to share your chains with a collaborator but they occupy **too much space** (especially when convergence criterion is set to a strict value), but you want to "**compress**" your chains in a quasi-loseless way

- Or you want that the points in your chain have **very little correlation** between step *N* and *N+1*, e.g. to compute weighted means $\langle x\rangle=\int xP(x)dx=(1/N_{mcmc})\sum x_i$

- If that is the case, you can create it automatically by setting in **params.ini** file the option **`indep_sample`** to some large number (e.g. 100). This also creates **.data** files with all the information (C_ℓ, P(k)) for importance sampling Or just do it with **GetDist** with a non-zero value in the **`thin_factor`** option



The file Distparams.ini

- **Distparams.ini** is the parameter file for **GetDist**. The syntax is:

```
./getdist distparams.ini chains/root
```

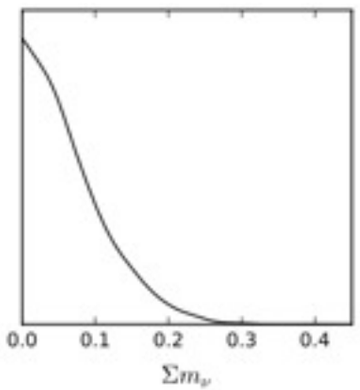
(note that you don't need to specify root_1.txt, this will **automatically** detect how many chains you have) and the most useful settings are:

- To create (or not) python scripts for **plotting** just set `no_plots` = (T or F)

```
#if no_plots =F, produce plot script files for specific plots producing ./plot_data/ density files
no_plots = T
```

- **1D plots** (marginalized probability distribution functions): automatically these will be generated for all parameters, but you can **choose** which ones

```
#Parameters to use. If not specified use all parameters which have lables.
#plot_params = omegabh2 omegach2 tau ns
```

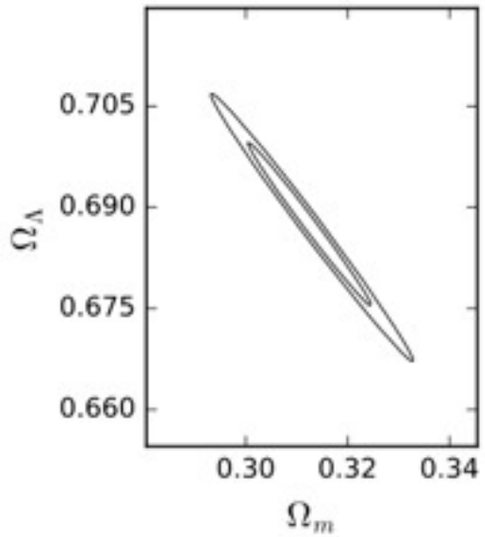


The file Distparams.ini

- for **2D plots** (contour plots):

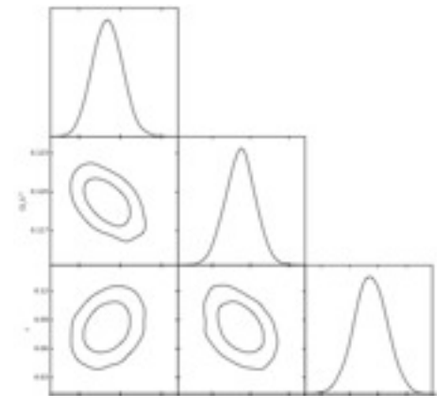
```
#if we only want 2D plots agains a particular variable
plot_2D_param = 0

#if above zero, instead plot just these combinations:
#if both zero it will plot most correlated variables
plot_2D_num = 0
plot1 = ns omegabh2
plot2 =
```



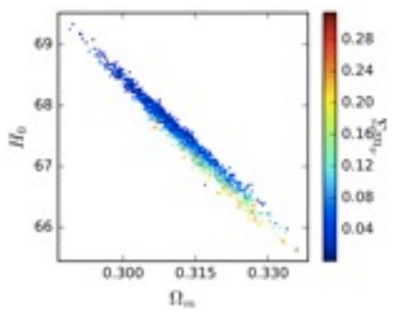
- for **triangle plots** (combination of 2D plots):

```
#Output 2D plots for param combos with 1D marginalized plots along the diagonal
triangle_plot = T
triangle_params = omegabh2 omegach2 tau omegak mnu nnu yhe Alens ns nrun logA r H0 omegam omegal sigma8 r02
```



- for **3D plots** (2D plots color-coded using a 3rd parameter):

```
#number of sample plots, colored by third parameter
#if last parameter is 0 or -1 colored by the parameter most correlated
#with one of the eigenvector directions (e.g. parallel or orthogonal to degeneracy)
num_3D_plots = 1
3D_plot1 = H0 omegam tau
```



The file Distparams.ini

- Distparams.ini is more than just about plotting:

- **Thinning** of existing chains

```
# if non-zero, output _thin file, thinned by thin_factor
thin_factor = 0
#Do probabilistic importance sampling to single samples
make_single_samples = F
single_thin = 4
```

- **Principal Component Analysis** of chain parameters

```
#PCA - analysis output in file file_root.PCA
#number of parameter to do PCA for
PCA_num = 0
PCA_normparam = omegam
#The parameters to use
PCA_params = omegam H0 tau
#L for log(x), M for log(-x), N for no log
PCA_func   = LLL
```

The file Distparams.ini

- Note that this configuration file also supports `INCLUDE` and `DEFAULT`
  The most technical details of distparams.ini are actually set here:
  python/getdist/analysis_defaults.ini
  (or alternatively **batch2/getdist_common.ini** if you use disttest.ini)

- Set a different **confidence level** of your contours (default 68% 95% 99%),
  or change the **binning** and **smoothing** of your 1D or 2D plots.

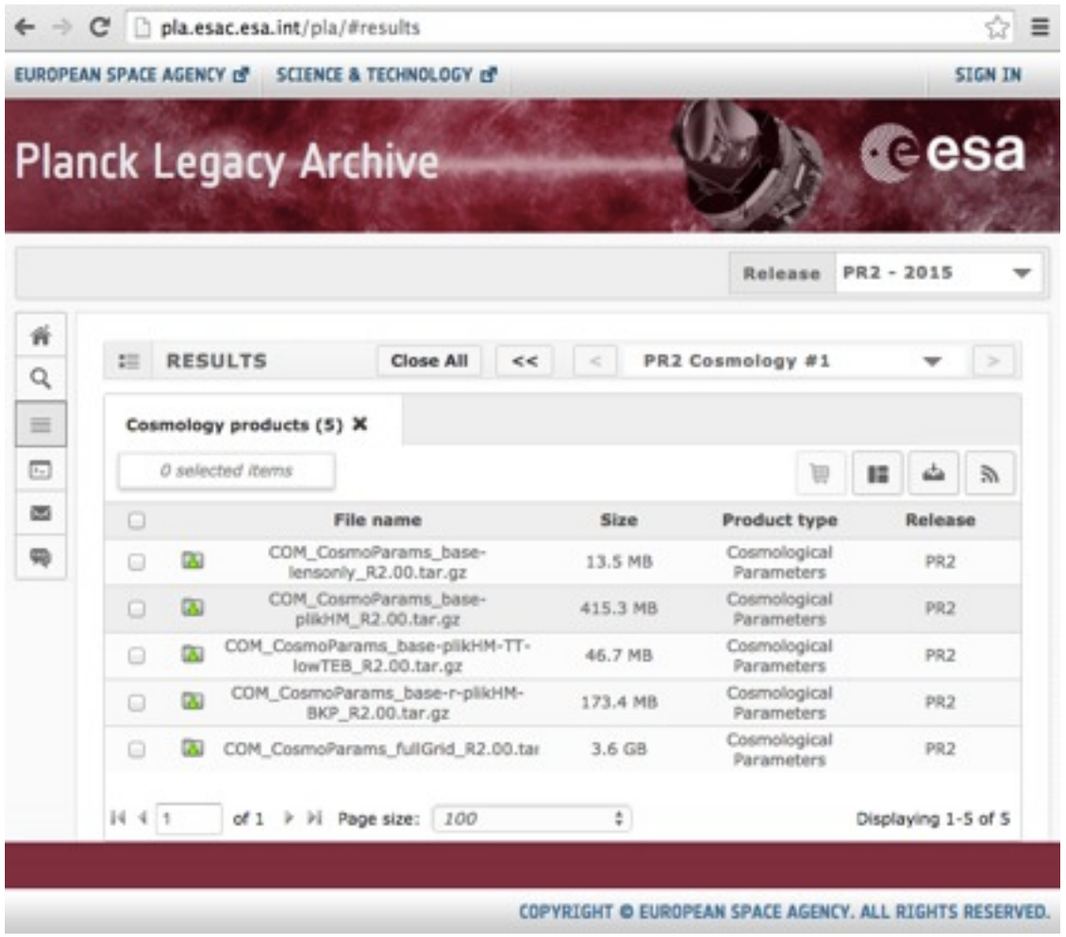
- Due to a **bug** in the current version, **you need to modify** the parameter file
  python/getdist/analysis_defaults.ini and set **`smooth_scale_2D`** to **2.0**
  (otherwise if you leave it set to -1 all contour plots will be empty)

Planck public chains

- The **Planck collaboration** has made their CosmoMC chains **publicly** available (used in Paper XIII, “Cosmological Parameters”)

- The chains can be **downloaded** from the ESA/Planck webpage: *http://pla.esac.esa.int/pla*

- Go to “**Cosmology**” and then click on “**Cosmological Parameters**”



Planck public chains

- The **full grid** is `COM_CosmoParams_base-plikHM_R2.00.tar.gz` (3.6GB) Remember to read the wiki for details! e.g. "*Note that the baseline model includes one massive neutrino (0.06eV). Grid outputs include WMAP 9 results for consistent assumptions*." These contain a large number of cosmological **models** and **dataset** combinations (all combinations are Planck+others)

- If you are interested in constraints on inflation, you might want this file too (constraints on primordial B-mode polarization by **Bicep-Keck-Planck**) `COM_CosmoParams_base-r-plikHM-BKP_R2.00.tar.gz`

- The **Documentation** can be found on the ESA/Planck wiki: *http://wiki.cosmos.esa.int/planckpla2015/index.php/Cosmological_Parameters*

Using GetDist GUI

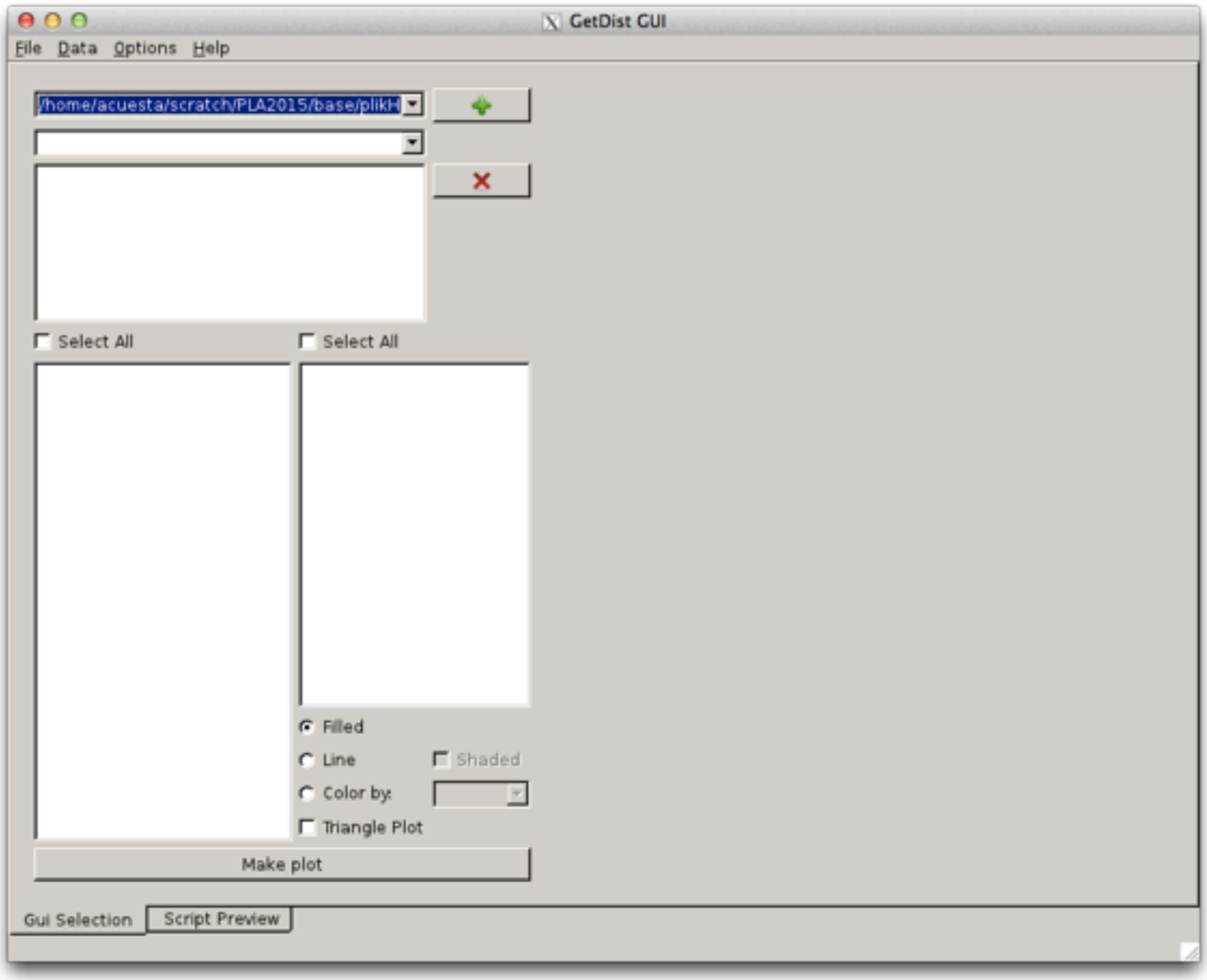
- **GetDist** has also a python version with a **graphical** user interface (GUI), which makes really easy to **generate plots** similar to those in Planck papers

- To **start the GUI** (remember you need to access the remote computer with graphical mode, i.e. `ssh -X user@ftaecluster.ugr.es`):
  `python python/GetDistGUI.py`

- Make sure you have the **python environment variables** set up so that you can use the code from any directory:
  `export PYTHONPATH=/home/ajcuesta/cosmomc/python:$PYTHONPATH`

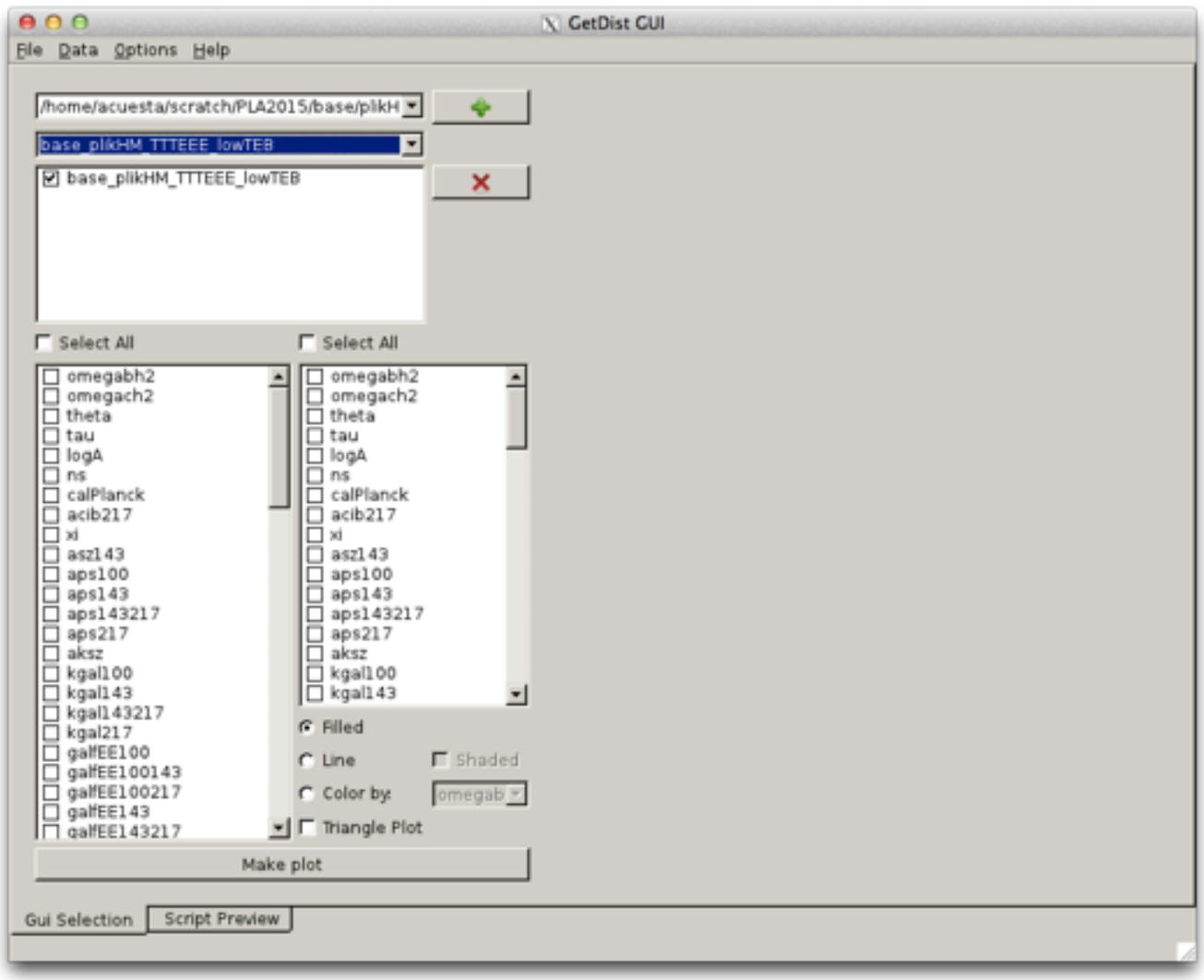
- The **Planck2015 public chains** are already installed in ftaecluster here:
  `/home/ajcuesta/PLA2015/`

GetDist GUI



First select your **model+data**

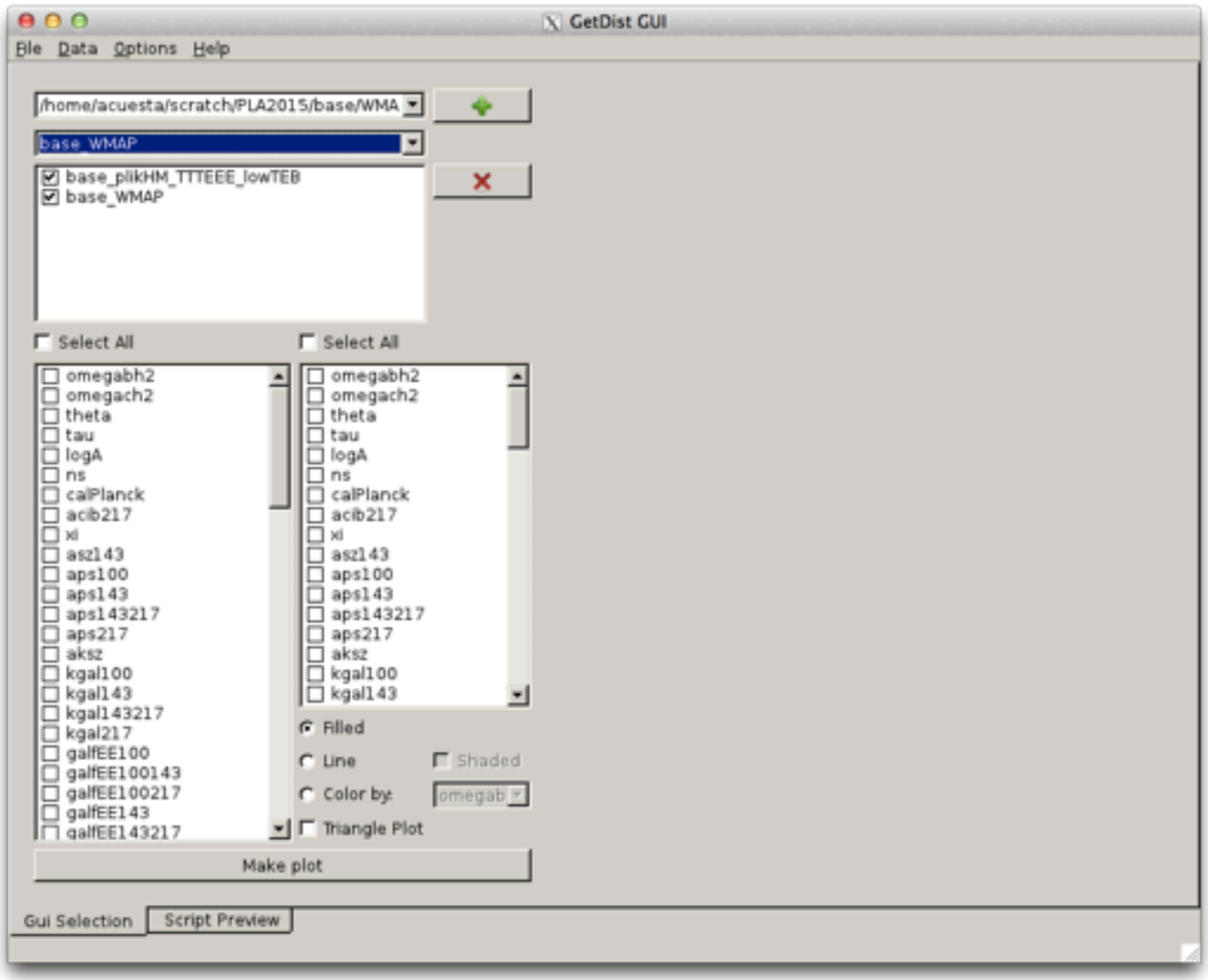
GetDist GUI



Then select the **plain** version
or a **post-processed** version
(e.g. with BAO)

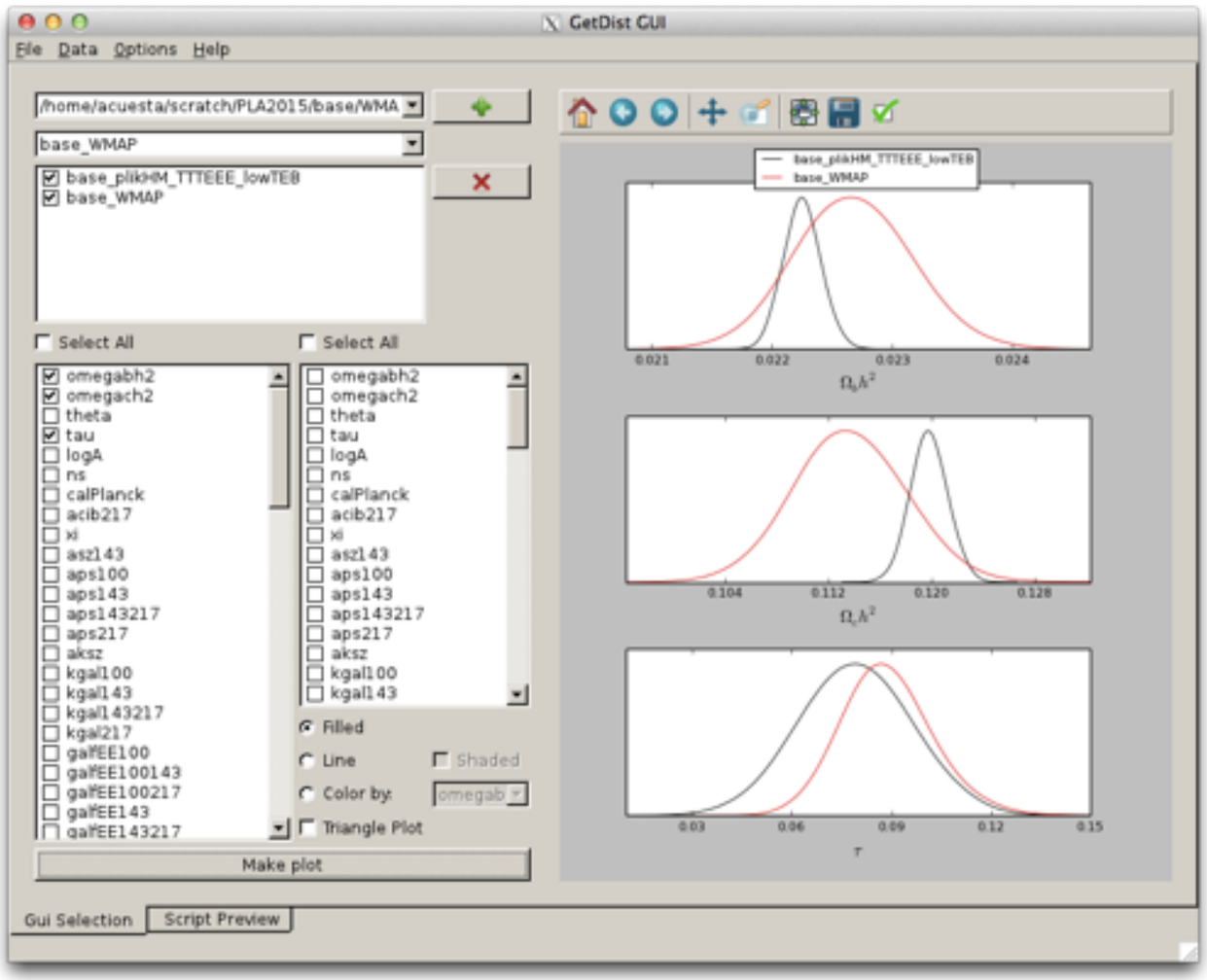
If you don't want any plots
you can instead click
on the **Data** menu
and check the **margestats**,
the **likestats**, and even create
a **LaTeX table** of your parameters

GetDist GUI



If you want to **compare** different models and/or dataset combinations click the **green +** button and **add** them

GetDist GUI

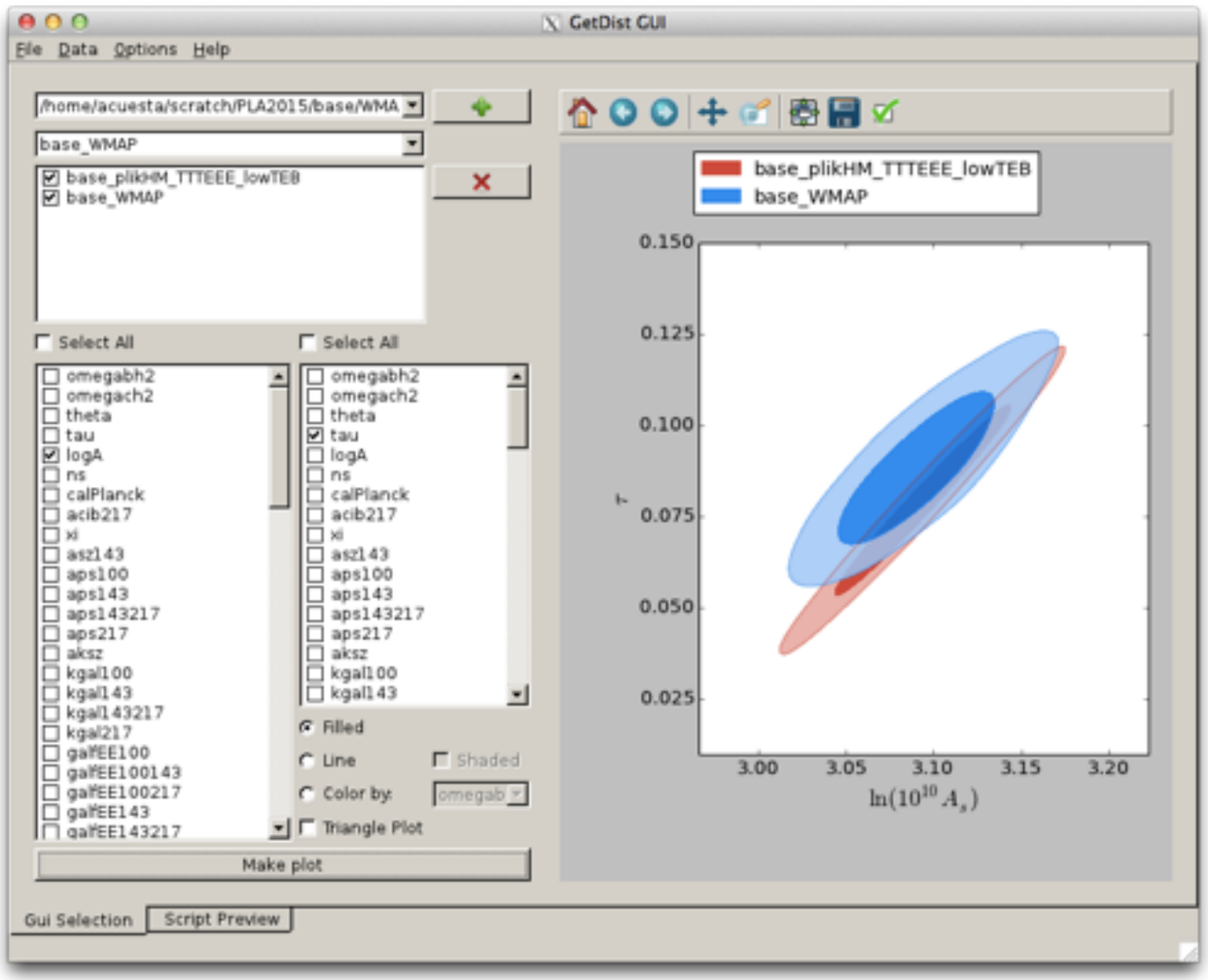


For **1D plots**, select the parameters you want from the **left column** and click **“Make plot”**

You can **configure** the plot (colors, shading, etc - matplotlib style) using the “**Options**” menu, then “**Plot settings**”

Save the figure by clicking on the **floppy disk** icon (many formats available: png, pdf...)

GetDist GUI

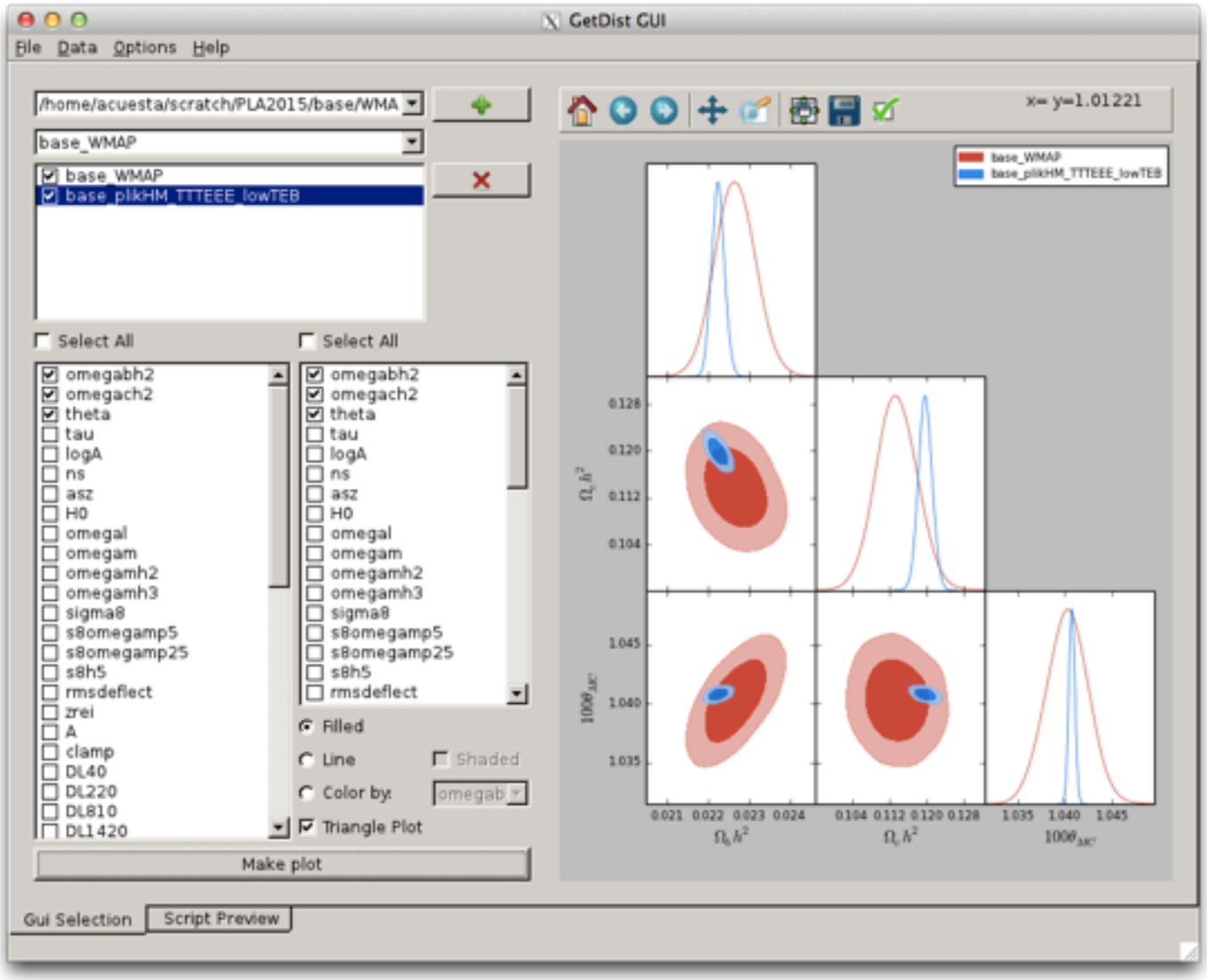


For **2D plots**, select one parameter from the **left** column and another from the **right** column and click "**Make plot**"

You can plot only **empty contours** if you click on "**Line**", or **filled contours** if you click on "**Filled**"

If the plot looks very noisy change **smooth_scale_2D** to **2.0** or play with that value

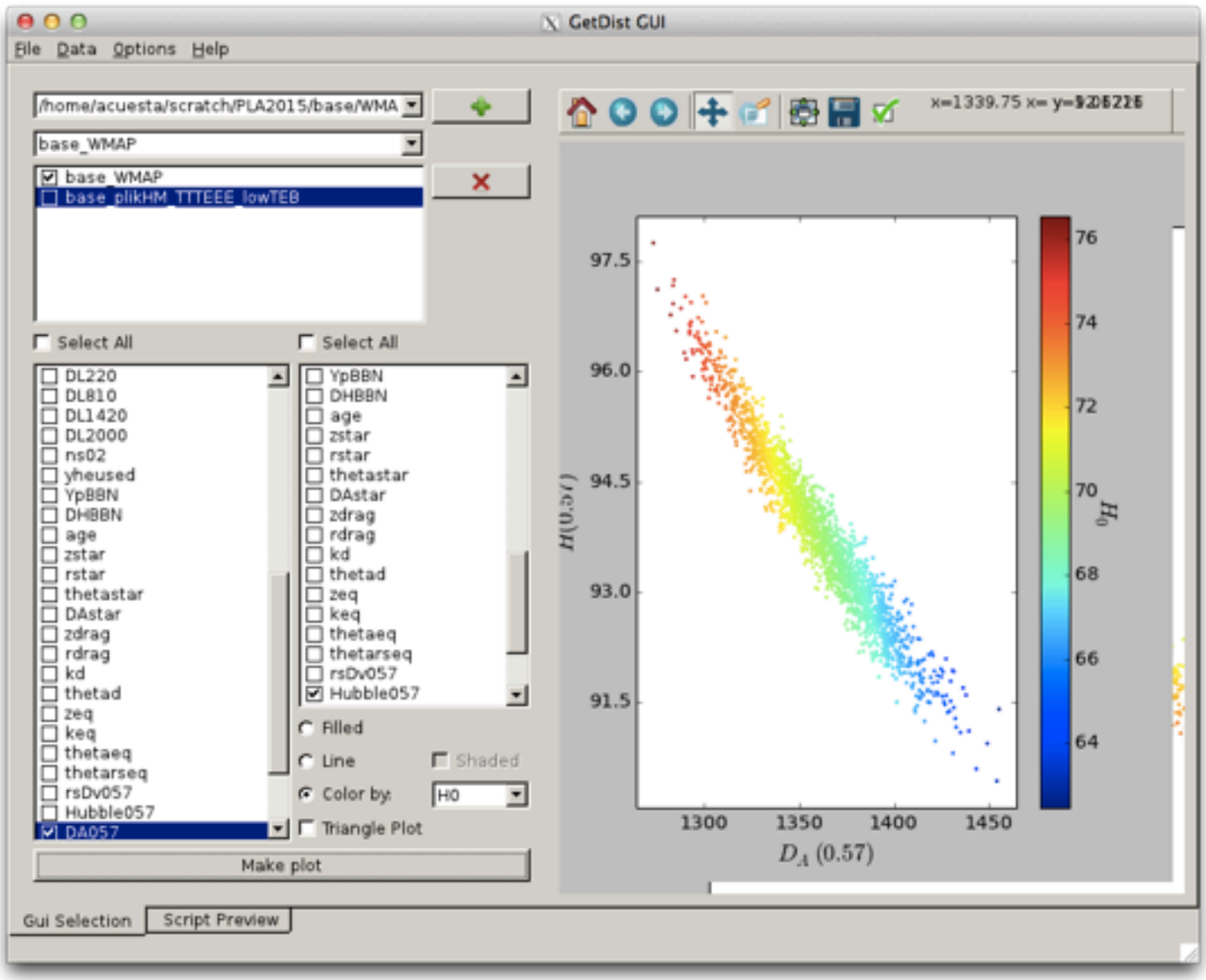
GetDist GUI



If you select **N** parameters
from the **left** column and **M** parameters
from the **right** column, you will get
an **array of N*M plots**

In that case, if the parameters
are **identical** in **both** columns,
you can also make it a
"**Triangle Plot**" by clicking on that option

GetDist GUI



Finally you can do also **3D plots**.
In that case it makes sense only
to select **1 single model+dataset**.

These plots are useful to get some
intuition of **degeneracies** and
parameter **ranges** in each model
or dataset combination!

more examples at http://getdist.readthedocs.org/en/latest/plot_gallery.html

More questions?

- **Help resources**:
  http://cosmologist.info/cosmomc/readme.html (CosmoMC Readme)
  http://cosmologist.info/cosmomc/readme_planck.html (Planck Readme)
  http://cosmologist.info/cosmomc/readme_python.html (Python Readme)
  http://cosmologist.info/cosmomc/readme_gui.html (GetDist GUI Readme)
  http://cosmologist.info/cosmomc/readme_grids.html (Grids Readme)

- **Support** (CosmoCoffee wiki):
  http://cosmocoffee.info/

- **Mail** me!
  ajcuesta@icc.ub.edu

Thank you

ajcuesta@icc.ub.edu