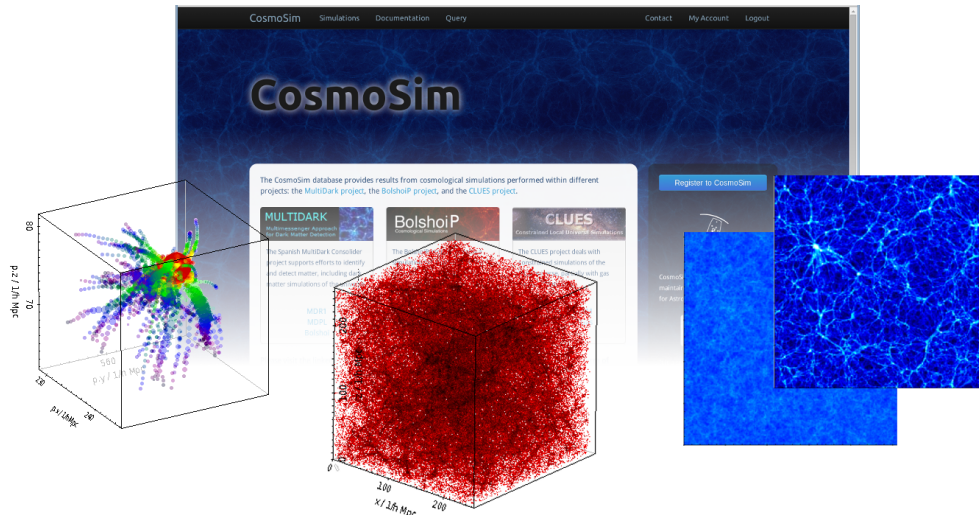# Introduction to simulation databases: Density fields and dark matter halos

Kristin Riebe,[*] GAVO

March 13, 2015

Simulation databases like the Millennium database or CosmoSim contain data sets from computer simulations of the universe. Via a web interface, users can access these tables. The data contain properties of dark matter halos forming in such simulations, their merging history, density information as well as the simulation particles themselves.

The simulations we use in this tutorial do not contain any gas dynamics, only gravitational forces are at work. Such simulations can be used to study the formation of the large scale structure in the universe: where are clumps forming, where are filaments? Are there big voids (empty regions) in between? Where are the dark matter halos located, in which galaxies may have formed, and how did they evolve?

In this tutorial, we will extract slices of the density field from a simulation, get dark matter halos for the same region, and view the particle distribution and merger history for one halo.

Requirements:

- website CosmoSim (`http://www.cosmosim.org/`)

- Topcat (`http://www.star.bris.ac.uk/~mbt/topcat/`)
  alternatively: Gnuplot (`http://www.gnuplot.info/`)

Further material:

- CosmoSim documentation
  `http://www.cosmosim.org/cms/documentation/`

- Demo video and more extensive simulations-tutorial at
  `http://www.cosmosim.org/cms/documentation/demos-and-tutorials/`

[*]Leibniz-Insitut für Astrophysik Potsdam (AIP); Email: kriebe@aip.de

1   Density field

We will start with extracting a slice of the density field in the x-y-plane from the database. See Appendix A for more background information.

1.1   Get the data

- Open the CosmoSim website www.cosmosim.org and login with your username and password.
  If you do not have an account yet, please register. Registration is free and open to anyone.

- Go to the **Query Form** in the top menu (http://www.cosmosim.org/query).
  Queries for the data are written in SQL, directly into the white text area, and then send to the database.

- Click on **Example Queries** below the text area. Choose Slice of density field at redshift 0. The SQL expression for getting the data from the database is automatically inserted into the text area:

```
SELECT ix, iy, iz, dens
FROM Bolshoi.Dens256
WHERE snapnum=416 AND iz = 94
ORDER BY ix, iy
```
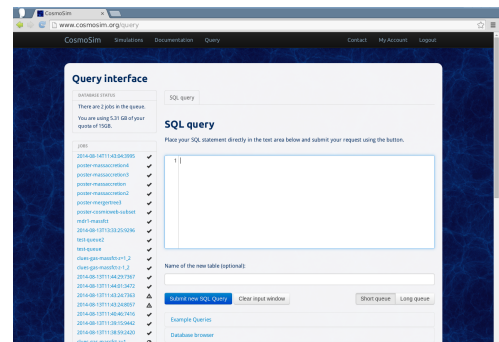
- Enter a name for your result table in the line below the main text area, e.g. density_416.

- Click the **Submit new SQL Query** button and confirm all additional messages with Submit/OK.

- Your job now appears in the **Jobs list** on the left side. The small icon next to the result name indicates if your job is running (triangle) or has already finished (check mark). Click on the name to see the details.

- Once your job is finished (which may take a while), click the job name again, then choose the tab **Results Table**. This shows the data you retrieved from the database.

1.2   View the density field with Topcat

- **Start Topcat**, a tool for viewing database tables and (simple) visualisations. If you don't have it on your computer yet, make sure that Java is installed and download the latest version from the Topcat website. You can also start it directly by clicking this link:
  http://www.star.bris.ac.uk/~mbt/topcat/topcat-full.jnlp.

- Switch back to the **web browser** with the CosmoSim query page. When your query is finished, click on the job name and switch to the **Results Table** tab. Scroll down below the table, click **Register with SAMP** and confirm the security exception.

- Enter your **password** in the text field and click **Send Table to topcat**. This sends your result table directly to Topcat.

- If this does not work for some reason, then switch to the Download-tab, choose Comma-Separated Values and click on Download. The file will be prepared on the server. Click on the new link to download and save the file. In Topcat, open the file: File → Load table, select Format: CSV and use System Browser to select your file.

- In **Topcat**, choose **Graphics** → **Plane Plot** . This opens a new window.
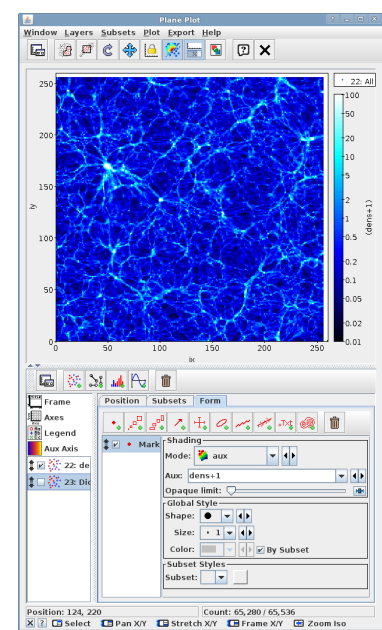
- In the **Position** tab, choose following axes: `X: ix`, `Y: iy`.

- Select the `Form` tab next to `Position` and `Subsets`. Increase the size of your window (click and drag with left mouse button) or switch to full screen mode in order to see all the possible options.

- At `Shading` choose the `Mode: aux`; and for `Aux: dens+1`. This adds density coloring to your dots, but currently your image looks nearly black.

- Select the `Aux Axis` from the left menu to adjust its properties. Change here the `Scaling` to `log` instead of linear. This will apply a logarithmic scaling for the colors. Since the logarithm is undefined for negative values, we had added +1 to `dens` for shifting the values above 0 (i.e. from overdensities to relative densities, see Appendix A). The web like structure should be visible now.

- At `Aux Shader` choose a color table you like, e.g. `Cold` for blue shades.

- Switch to the `Range` tab to fix the range of your density values: `Minimum Aux: 0.01`, `Maximum Aux: 100`
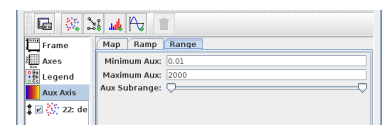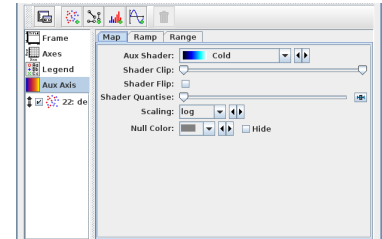
1.3   More density fields

- Go back to the CosmoSim website, **Query form**, and select again the same example query. Change now the snapnum inside the query from 416 to **snapnum=36**. This changes the snapshot number, i.e. the timestep for which the density field is retrieved. 416 corresponds to today; snapnum=36 means $\approx$ 500 million years after the Big Bang for this simulation. Choose a different name for the results, e.g. `density_36` and **submit your query**.

- Make sure that **Topcat is still open** and follow the same steps as before to **send your results to Topcat**.

- In Topcat, the new table should appear in the main window. Go to the **Plane Plot** window we prepared in the previous section. Add the new data to the same plot window by clicking `Layers` → `Add Positional Control`.

- Choose your new table, and again enter `ix` and `iy` for the positions. Make the same adjustments as in the previous section, i.e. at `Form` tab add **Aux**, `dens+1`.

- Tick the checkmark next to the table name on/off to switch between the different density fields. Can you see how the early density fluctuations have evolved into more condensed structures?

- [**Extra**] Repeat these steps again for more snapnums, see the table at http://www.cosmosim.org/cms/news/density-fields-for-bolshoi-updated/ for an overview of all available values. Experiment also with different slices, e.g. choose a different value for `iz` or fix `ix` instead.

- [**Extra**] You can also compare with the density field calculated using a 512^3 grid. Just change the table name in your SQL query to `Bolshoi.Dens512`.

- [**Extra**] Compare with the density field of a simulation with a larger volume: MDR1 with a 1000 Mpc/h box. For this simulation, we only have the density field at redshift z=0, i.e. today, at the end of the simulation, available. Here's an examle query for a slice of this simulation:
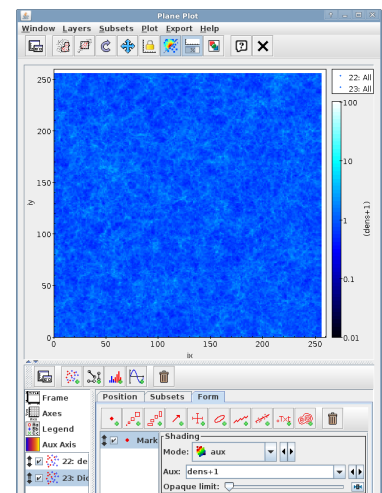
```
SELECT ix, iy, iz, dens FROM MDR1.Dens512 WHERE iz = 447
ORDER BY ix, iy
```

2   Dark matter halos

In this section we retrieve dark matter halos from the Friends-Of-Friends (FOF) halo catalogue. In such halos, galaxies like the Milky Way and galaxy clusters can form. See the documentation at CosmoSim on Simulations and Halofinders for more information.
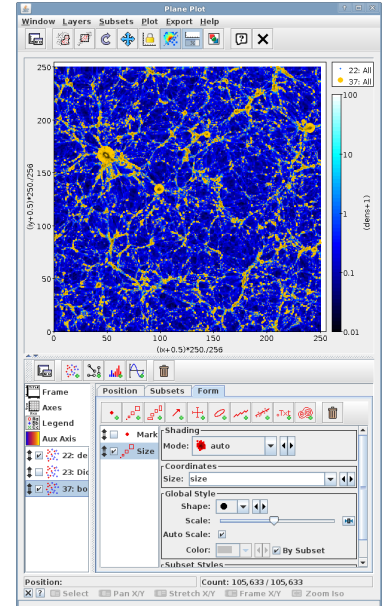
2.1   Dark matter halos for the slice

- Go to the **Query Form** at the CosmoSim webpage and enter following query:

```
SELECT fofId, x, y, z, mass, size FROM Bolshoi.FOF
WHERE snapnum=416 AND floor(iz/4.) = 94
```

  We need to divide `iz` by 4, since the grid index for the FOF catalogue is based on a 1024^3 grid.

- Enter a **result name** and **Submit**. When the job is finished, click on it in the **Jobs list** and switch to the **Results Table** tab.

- **Send the table to Topcat** via SAMP. If necessary, click again `Register with SAMP` beforehand and enter your CosmoSim password. If Topcat is not listed, make sure that it is still open.

- Go to Topcat's **Plane Plot** from the previous section and show only the density plot for snapnum=416 (remove checkmark for all other plots). **Add the new data set** to the same plot window by clicking `Layers → Add Positional Control`.

- Select the most recent table and adjust the **coordinates**: X: `x`, Y: `y`. This plots the positions of the dark matter halos on top of the density field.

- Switch to the **Subsets** tab and choose a good contrasting color, e.g. yellow.

- Switch to the **Form** tab, click the icon **Add new Size Form**. Enter at `Size: mass`. This scales the dot's size proportional to their mass. Uncheck `Mark` in order to have only the sized markers visible.

- Move the **Scale** slider for the sized markers to adjust their size.

- The halos should lie on top of the bright regions in the density field, but they look slightly shifted and scaled. The reason lies in the density field positions: There we plotted only the ix, iy grid cell index, which corresponds to the lower left corner of each grid cell. In order to get the position at the center, we need to add +0.5 to each coordinate and to scale them down from 256 grid cells to 250 Mpc/h. Thus, select the **density** data in the left menu and **adjust the positions** to: X: `(ix+0.5)*250./256`, Y: `(iy+0.5)*250./256`.

- [**Extra**] Scale the dots by their `size` (radius) instead of `mass` by changing the column at `Size`. The differences in size are not as strong as differences in mass (mass-factor $\approx$ size-factor^3).

2.2   3D distribution of halos

- Go to the **Query Form** at the CosmoSim webpage and enter following query:

```
SELECT x, y, z, mass, size FROM Bolshoi.FOF
WHERE snapnum=416 AND mass >= 1.e10
AND sprng_dbl(111) < 0.01
ORDER BY mass DESC
```

This randomly selects 1% of all halos above the given mass threshold and sorts them by mass. The query retrieves $\sim$ 3 MB of data. The results will look even nicer, if you change the fraction from 0.01 to `0.1`; but this may slow down Topcat's 3D view substantially. If you already have performance problems with Topcat or intend to view the results with Gnuplot only, add `LIMIT 2000` at the end of the query for limiting your results to 2000 rows.

- Enter a **result name** and **Submit**. When the job is finished, send the table to **Topcat**.

- In Topcat's main window, select the new table, then choose **Graphics → Cube Plot** ⬡. This opens a new window for 3D plotting.

- Choose for **positions**: `X: x`, `Y: y`, `Z: z`.

- Switch to the **Form** tab. Increase the window size for better visibility. Select `Shading, Mode: density`.

- Scroll down to `Global Style` and set the dot **Size** to `0` instead of the default 1. Use the left mouse button to rotate the 3D plot. More help on navigation is available by clicking the small question mark at the bottom left of the window. One can see that the large scale structure is reflected in the halo distribution.
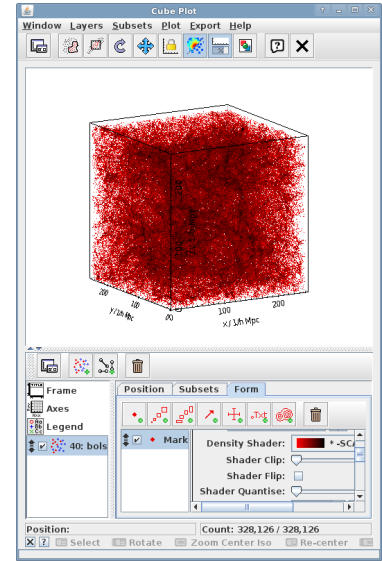
2.3   Particles of a halo

- Go to the **Query Form** at the CosmoSim webpage and enter following query:

```
SELECT p.x,p.y,p.z
FROM
  MDR1.Particles85 AS p,
  (SELECT particleId from MDR1.FOFParticles
   WHERE fofId = 85000001014) AS fp
WHERE fp.particleId = p.particleId
```

This query selects all particles and their positions for the halo with `fofId = 85000001014` from the MDR1 simulation, at redshift 0 (snapnum=85 for this simulation). You can change the digits at the end of this `fofId` to select a different halo.

- Enter a **result name** and **Submit**. When the query is finished, send the table to **Topcat**.

- In Topcat's main window choose **Graphics → Cube Plot** ⬡.

- Choose following **axes**: `X: p.x`, `Y: p.y` and `Z: p.z`

- Color the particles by projected density: choose the **Form**-tab, select at `Shading` the `Mode: density`. Select a nice colormap, e.g. `Pastel`. Choose `Shader Flip` and `Scaling: Log` to get pretty results.

- Use the left mouse button to rotate the plot. If you look close enough, you can even detect small lumps of higher density. These small subhalos are remnants of halos that were trapped in the gravitational potential of this halo during the simulation.
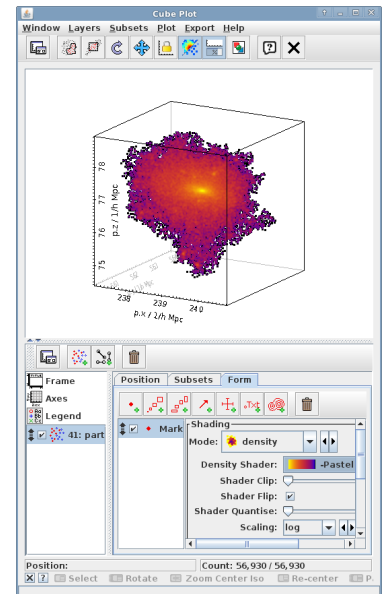
2.4 Mass growth of a halo

Dark matter halos grow with time, because they accrete loose material ("smooth accretion") as well as other dark matter halos. This can be detected when plotting the mass growth of a halo with time.

- Go to the **Query Form** at the CosmoSim webpage and enter following query:

```
SELECT p.fofTreeId, p.treeSnapnum, p.mass, p.size
FROM MDR1.FOFMtree AS p,
(SELECT fofTreeId, mainLeafId FROM MDR1.FOFMtree
  WHERE fofId = 85000001014) AS m
WHERE p.fofTreeId BETWEEN m.fofTreeId AND m.mainLeafId
ORDER BY p.treeSnapnum
```

This query selects all direct progenitors of the given halo (so-called *main branch*). Adjust the `fofId`, if you want to get the mass growth of a different halo.

- Enter a **result name** and **Submit**. Wait until it is finished and send the table to **Topcat**.

- In **Topcat**, choose **Graphics → Plane Plot** .

- In the **Position** tab, choose following axes: X: `p.treeSnapnum`, Y: `p.mass`. The `treeSnapnum` is the time axis here: 0 is a snapshot closer to the beginning and 39 at the end of the simulation.
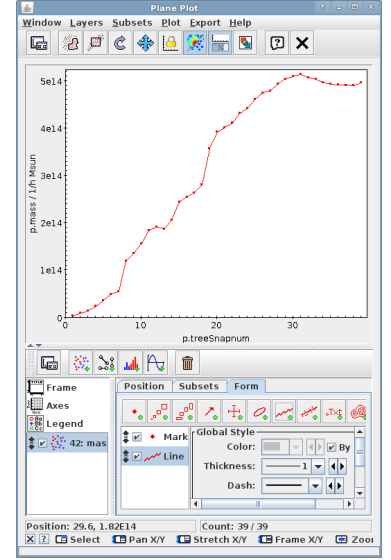
- Switch to the **Form**-tab and add a new line form for connecting the points. Steep increases in the mass from one timestep to another indicate the accretion of a lot of mass at once, i.e. a merging events with other dark matter halos.

2.5 Merger tree: progenitors of a halo

Let us now find all the halos which have merged into one dark matter halo at the end of the simulation.

- Go to the **Query Form** at the CosmoSim webpage and enter following query (again adjust the `fofId`):

```
SELECT p.fofTreeId, p.treeSnapnum, p.x, p.y, p.z, p.size
FROM MDR1.FOFMtree AS p,
(SELECT fofTreeId, lastProgId FROM MDR1.FOFMtree
  WHERE fofId = 85000001014) AS m
WHERE p.fofTreeId BETWEEN m.fofTreeId AND m.lastProgId
ORDER BY p.treeSnapnum
```

- Enter a **result name** and **Submit**. When the query is finished, send the table to **Topcat**.

- In Topcat, go to **Graphics → Cube Plot** .

- Select following **axes**: X: `p.x`, Y: `p.y` and Z: `p.z`.

- Switch to the **Form**-tab and add a new size form . Set the `Size` to `p.size` for scaling the dots with the average halo radius.

- Apply more color by selecting at `Shading`, the `Mode`: `aux` and entering `Aux`: `treeSnapnum`. Adjust the `Opaque limit` to your liking.

- Now halos from the beginning of the simulation are colored black-purple, more recent progenitors have yellow-red colors. This nicely shows how the progenitor halos have merged with each other at earlier times to form the final halo (red) at the center.
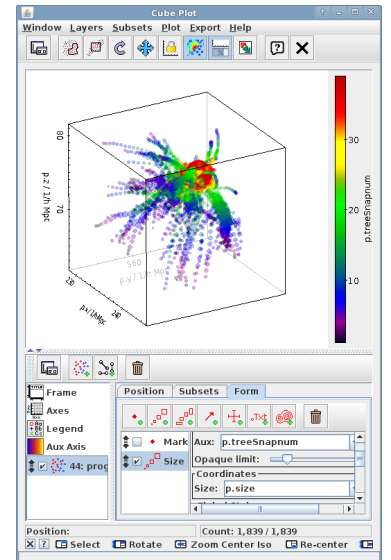
A   Density field and grid cells

The density was calculated by dividing the simulation box in grid cells of equal size. For each cell, the number of simulation particles was counted and used to calculate the density: the more particles there are inside, the higher is the density[1].
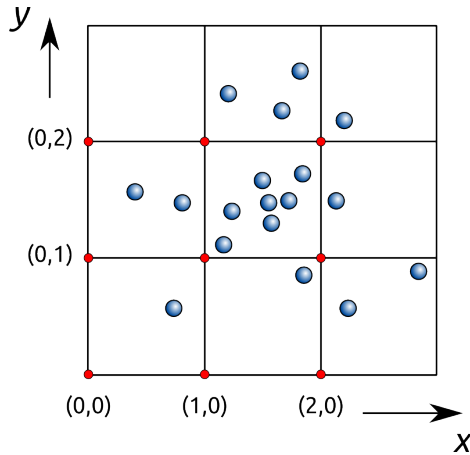
ix	iy	particles	relative density	overdensity
0	0	1	0.5	-0.5
1	0	1	0.5	-0.5
2	0	2	1.0	0.0
0	1	2	1.0	0.0
1	1	7	3.5	2.5
2	1	1	0.5	-0.5
0	2	0	0.0	-1.0
1	2	3	1.5	0.5
2	2	1	0.5	-0.5

Simplified density estimation: count the particles in each cell, determine the relative density by dividing through the mean number of particles per cell (18 particles/9 cells = 2) and subtract 1 to get the overdensity (above the mean). Such overdensities along with the corresponding indizes (ix, iy, iz) can be stored in database tables.

For the Bolshoi simulation at CosmoSim, the density was calculated on a grid with 256^3 cells. The simulation box has a side length[2] of 250 Mpc/h, which gives for each of our grid cells a side length of $250/256$ Mpc/$h \approx 0.977$ Mpc/h.

B   Plotting with Gnuplot

In case there is no Java/Topcat available on your computer, you can also visualize the data using Gnuplot (http://www.gnuplot.info/). Download the necessary executable (Windows) or tar-ball (Linux) and install it on your computer. Gnuplot can be used via a command line or by loading a file which contains the commands, a so-called *script*. Here we provide the commands to plot the data from the exercises in this tutorial.

- Download the result table of your SQL query via the **Download**-tab at the CosmoSim website. Choose CSV format (columns separated with comma), generate the file and click on the link. Save the data to a directory of your choice.

- Put the Gnuplot commands of one of the sections below into a pure text-file (on Windows use e.g. Notepad for this). Replace the file name at the plot/splot command by your own data file name and save your script-file as a text file, e.g. `density.plt`.

- Start the Gnuplot command line: in Linux type `gnuplot` in your terminal; in Windows double-click the corresponding exe-file. Change your local directory to the script-directory (if necessary). Then load the script with `load 'density.plt'` and view the result.

B.1   Density plots

- Density slice, section 1, `density.plt`:

---

[1]This is only half of the truth: in fact, a smoothing function was applied: each particle also contributes to adjacent cells.

[2]h is the Hubble constant divided by 100. It can be ignored here and just considered as part of the unit.

```
set datafile separator ","
set palette defined (0 "black", 1 "blue", 2 "white")
set size square
set logscale cb
set xlabel 'x␣[Mpc/h]'
set ylabel 'y␣[Mpc/h]'
set cblabel 'density'
set xrange [0:256]
set yrange [0:256]
set zrange [0.01:100]

set view map
splot 'density-416.csv' u ($2):($3):($5+1) with points pt 5 ps 0.1 palette notitle

pause -1
```

- Density slice with dark matter halos, section 2.1, `densityhalos.plt`:

```
set datafile separator ","
set palette defined (0 "black", 1 "blue", 2 "white")
set size square
set logscale cb
set xlabel 'x␣[Mpc/h]'
set ylabel 'y␣[Mpc/h]'
set cblabel 'density'
set xrange [0:256]
set yrange [0:256]
set zrange [0.01:100]

set view map
splot 'density-416.csv' u ($2):($3):($5+1) with points pt 5 ps 0.1 palette notitle

pause -1
```

B.2 Halo plots

- Plotting the result from the query of section 2.2, random selection of halos, won't look as nice as with Topcat, since Gnuplot does not calculate projected densities on the fly. You can still look a selection of the most massive halos in 3D, `halos3d.plt`:

```
set datafile separator ","
set palette defined (0 "black", 1 "red", 2 "white")
set size square
set logscale cb
set xlabel 'x␣[Mpc/h]'
set ylabel 'y␣[Mpc/h]'
set zlabel 'z␣[Mpc/h]'
set ticslevel 0
set xrange [0:250]
set yrange [0:250]
set zrange [0:250]

splot \
'halos-3d.csv' every ::::2000 u ($2):($3):($4):($5) with points pt 7 ps 0.5 palette
    notitle

pause -1
```

- Halo particles, section 2.3, `haloparticles.plt`:

```
set datafile separator ","
set xlabel 'x␣[Mpc/h]'
set ylabel 'y␣[Mpc/h]'
set zlabel 'z␣[Mpc/h]'
set ticslevel 0

splot \
'halo-particles.csv' u ($2):($3):($4) with points pt 0

pause -1
```

- Mass growth, section 2.4, `massgrowth.plt`:

```
set datafile separator ","
set xlabel 'time'
set ylabel 'mass␣[M_sun/h]'

plot \
'halo-massgrowth.csv' u ($3):($4) with linespoints lt 1 lw 2

pause -1
```

- Merger tree, section 2.5, `halomergertree.plt`:

```
set datafile separator ","
set palette defined (0 "black", 1 "blue", 2 "green", 3 "yellow", 4 "orange", 5 "red")
set xlabel 'x␣[Mpc/h]'
set ylabel 'y␣[Mpc/h]'
set zlabel 'z␣[Mpc/h]'
set cblabel 'time'
set ticslevel 0
set pointsize 2

splot \
'halo-mergertree.csv' u ($4):($5):($6):($7):($3) with points pt 7 ps variable palette
    notitle
pause -1
```