

Introduction to simulation databases using CosmoSim

Kristin Riebe*, GAVO

July 05, 2016

Introduction

Simulation databases like the Millennium Database or CosmoSim contain data sets from cosmological simulations. Each user has access to catalogues of dark matter halos and their history (merger trees), supplemented by (mock) galaxy catalogues and information about the density field or tables with the simulation particles themselves and other data sets. Access to these data is possible via a web interface directly with the browser; for larger data it is more convenient to access the data via [TOPCAT](#), wget and using their TAP/UWS interface, if available. This tutorial gives a quick overview on the CosmoSim web interface and shows an example use cases for extracting the mass function of a simulation, particles for a halo, the mass growth history and progenitors of a halo as well as a slice of the density field.

This document is a short version which contains parts of the [simulations.pdf](#) and [simulation-teachers.pdf](#) available at the GAVO tutorials webpage^a. More demos, examples and tutorials can be found at the [CosmoSim](#) webpage.

^a<http://www.g-vo.org/pmwiki/Documents/Tutorials>

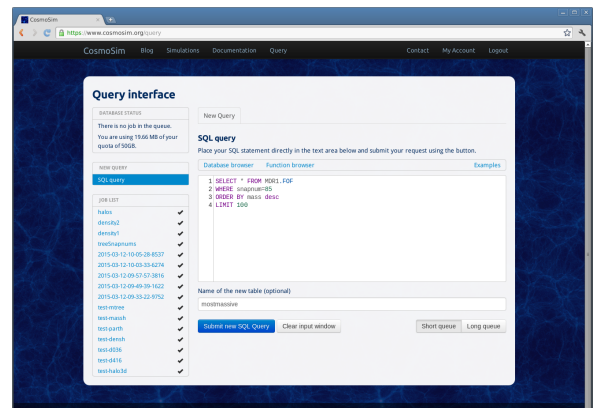
Software: [CosmoSim](#) web interface, [TOPCAT](#)

1 The web interface

1. Open <https://www.cosmosim.org> in a web browser. This webpage provides access to simulations from different projects, with data sets like dark matter halo catalogues and merger trees. The available simulations and data are described in the **Simulations** section of the documentation. There are also many documentation pages on the database and table structures, example queries and a demo video¹. You don't need to look at this now, just be aware that there is extensive documentation (and it may even be more up-to-date than this tutorial).
2. **Query interface:**
 - Login with your username and password (or as the demo user/guest) and navigate to the **Query interface** (<https://www.cosmosim.org/query>).
 - SQL queries are written directly into the main text area. The dialect used here is MySQL. Write the following text into the form:

```
SELECT * FROM MDR1.FOF
WHERE snapnum=85
ORDER BY mass DESC LIMIT 100
```

This query will retrieve the 100 most massive clusters (ordered by mass) for redshift 0 (= snapshot number 85) from the MDR1-simulation.



*Leibniz-Institute for Astrophysics Potsdam (AIP); email: kriebe@aip.de

¹<https://www.cosmosim.org/cms/documentation/demos-and-tutorials/first-steps-with-cosmosim/>

- Enter a new **name** for the result table below the SQL area. This name must be unique, if you leave the field empty, the web application will generate a unique name for you.
 - Select **Short queue** and click on **Submit new SQL query**.
 - Each query returns results in a (private) result table, the names of these tables are listed in the **Jobs list** on the left side. Click on your job to see its details.
 - Once your job is finished (check-mark), click the job name again, then choose the tab **Results Table**. Browse the table to view the properties of the most massive halos.
 - Above the form on the right side is the **Examples** link. Click on it to view the available example queries that you may also wish to try for yourself. Just double-click on one of them and the corresponding SQL query will be inserted into the SQL form.
 - Browse the available simulations and tables by clicking at **Database Browser** on the left side above the query form.
 - If you have long running queries, be sure to select **Long queue** before submitting, since the short queue times out after 30 seconds. It is also wise to restrict your query to use only the columns you are really interested in. This saves query time and disk space.
3. Consult the **documentation** (Database Structure → Tables) or the database browser to get to know which table provides which data.

2 Mass function for FOF groups

In this section we plot the mass function for dark matter clusters of a simulation and learn about the **SAMP** interface to plot the results with **Topcat**.

1. Get the mass function

- Open <https://www.cosmosim.org> in a web browser. Login and go to the **Query** form.
- A mass function gives the number of objects for a selected mass bin. Let's try this for the MultiDark Run 1 (MDR1) simulation, using the FOF table. It contains dark matter clusters found with a Friends-of-Friends algorithm. Type the following query into the Query form:

```
SELECT 0.5+FLOOR(LOG10(mass)) AS logmass,
COUNT(*) AS num
FROM MDR1.FOF
WHERE snapnum=85
GROUP BY FLOOR(LOG10(mass))
ORDER BY logmass
```



This will sort the FOF clusters into logarithmic mass bins using the FLOOR-function. The combination of COUNT and GROUP BY allows to count the number of clusters per bin.

- **Submit** the query and view the **Results Table** after it finished.
- Refine the bins by using steps of 0.25:


```
SELECT 0.25*(0.5+FLOOR(LOG10(mass)/0.25)) AS logmass,
COUNT(*) AS num
FROM MDR1.FOF
WHERE snapnum=85
GROUP BY FLOOR(LOG10(mass)/0.25)
ORDER BY logmass
```

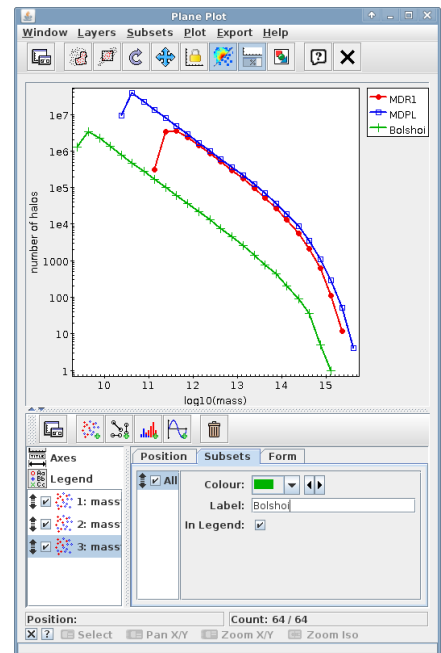
Submit again and check the **Results Table**: There are more rows now (factor 4).

2. View the result

- Start [TOPCAT](#), e.g. using `java -jar topcat-*.jar` from the command line in Linux.
- Go back to the browser window, choose the **SAMP** tab of your latest result and click **Register with a local SAMP Hub**. You may need to add a security exception in your browser first (see instructions on web page). Authorize the SAMP connection with "Yes" in the pop-up window. If you cannot see the SAMP-link, be sure that you are logged in. Guest users cannot use the SAMP interface. If you don't have an account or SAMP does not work, you can still download the table via the [Download](#) tab (choose e.g. ASCII VO-table) and open the file in Topcat manually.
- Click **Send table to topcat** in the Topcat-row. This sends the result table to Topcat; the table name should appear there at [Tables](#).
- In **Topcat**, choose the **Plane Plot** . This opens a new window. In the [Position](#) tab, choose following axes: **X: logmass**, **Y: num**.
- Select the [Form](#) tab next to [Position](#) and [Subsets](#) and click the button **Add new Line form** . This connects the points.
- Choose [Axes](#) in the left side to get access to the axes-properties. At tab [Coords](#) (that's the default) tick **Y Log**. This makes the y-axis logarithmic. This is your mass function of FOF dark matter clusters for the MDR1 simulation!

3. Experiment with mass functions for other simulations: How does the MDR1-mass function compare to MDPL or Bolshoi at the same redshift? How do mass functions change with the halo catalogue (use BDMV instead of FOF with Mvir for the mass) and redshift? Here are some tips:

- Take care to use the correct snapnum for each simulation to get the same redshift for each of them. The mapping between snapnum and redshifts is given in the [Redshifts](#) table for each simulation. For example, following snapnums correspond to redshift $z_{red}=0$: MDR1: 85, Bolshoi: 416, MDPL2: 125.
- In order to plot two tables in one [Plane Plot](#) window in Topcat, click the button **Add new positional plot control** . Then select a table and enter x- and y-axis.
- For adjusting the main color and label for each table, select a table at the left side, then choose the [Subsets](#)-tab. Here you can enter your label text at [Label](#) and choose the [Color](#).



3 Particles of a dark matter halo


For a few simulations, CosmoSim also provides the snapshots of selected timesteps. Thus it is possible to extract particles for a given position or a halo. In this section, the particles of an object from the FOF-catalogue of the MDR1-simulation will be extracted and plotted with Topcat. Please note, that we use here the term *halo*, though actually, the FOF catalogue returns only groups of connected particles and makes no statement if they are truly gravitationally bound or not. Thus it would be more correct to speak of *FOF groups* instead, but let's keep it simple.

- Open <https://www.cosmosim.org> in a web browser. Login and go to the **Query** interface.
- 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
```

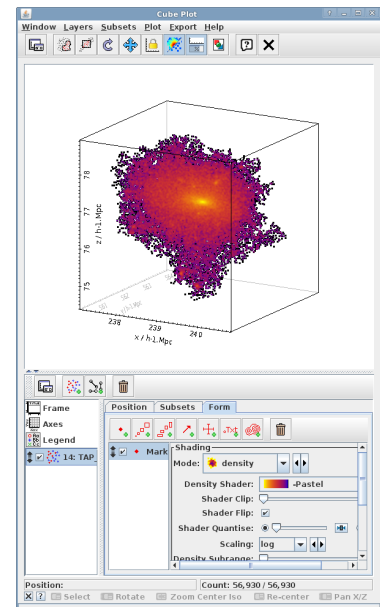
This query first selects the ids of all particles for the FOF-halo 85000001014 and then uses them to extract their positions from the snapshot stored in table MDR1.Particles85.

- **Submit** the query.
- When the query is finished, **click on the job** in the **Job list** and go to the **Plot-tab**. Choose **p.x** and **p.y** as X and Y axis and click **Create Plot**. This plots the particle positions in x-y-plane and is a nice simple feature for getting a first impression of the results.
- Make sure that **Topcat** is still running.

- In the browser, switch to the **Results Table-tab**. Scroll down below the table. You should still be registered with SAMP – if not, then click again [Register with a local SAMP Hub](#). Click **Send table to Topcat**.
- In Topcat, select the new table, then choose **Graphics** → **Cube Plot** . This opens a new window for **3D plotting**.

- Within the **Cube Plot** window of Topcat, choose following axes: **X: p.x**, **Y: p.y** and **Z: p.z**

There are so many particles, that the true shape is somewhat obscured. Fortunately, since Topcat version 4, the dots can be colored by projected density: choose the **Form-tab** (next to **Positions** and **Subsets**) and select at **Shading** the **Mode: density**. Select a colormap to your liking as well.



- You can navigate round the 3D plot using the mouse: drag to rotate, use the mouse wheel to zoom in and out, and right-click (or CTRL-click) to re-center. More help on navigation is available by clicking the small question mark at the bottom left of the window.
- If you look close enough, you can even detect small lumps of higher density that could represent small subhalos. You could check if halo finders have detected them as well by extracting the positions and sizes of dark matter halos in this region from the CosmoSim database and overplotting them here (see more extensive tutorials at the CosmoSim webpage or GAVO Tutorials page).


4 Mass growth of a halo

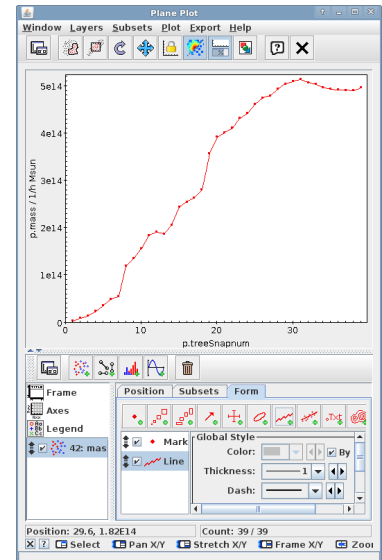
Dark matter halos grow with time, because they accrete loose material ("smooth accretion") as well as other dark matter halos. This can be detected when plotting the mass growth of a halo with time.


- Go to the **Query interface** 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. Step increases in the mass from one timestep to another indicate the accretion of a lot of mass at once, i.e. merging events with other dark matter halos.

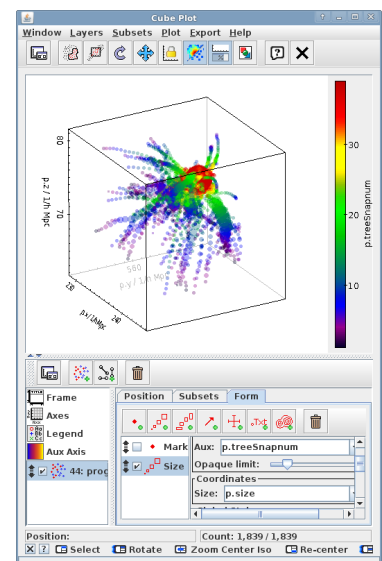
5 Merger tree: progenitors of a halo

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 interface** 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.

6 The density field with TOPCAT

- Go to the **Query interface** at the CosmoSim webpage and enter following query to extract a slice of the density from the Bolshoi simulation at the location of the most massive BDM halo:

```
SELECT ix,iy,iz,dens FROM Bolshoi.Dens256
WHERE snapnum=416 AND iz = 94
ORDER BY ix, iy
```

We use here the table Dens256, which contains a 256^3 grid, and for each grid cell the density value. Since we only fix the iz -value, we will get a 256×256 slice in x - y plane returned.

- **Submit** the query and after it finished switch to the **SAMP** tab. Make sure that **Topcat** is still open before clicking **Send table to topcat**.
- In **Topcat**, select your table and 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). 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** This will make the structures more prominent.
- Inspect the same slice region at different redshifts, e.g. for snapnum 243 ($z=1$), 132 ($z=2.6$) or 36 ($z=10.9$). Just replace the snapnum in the query above, submit the query and view the results with Topcat. Can you see how the large scale structures formed from small overdensities in the much more homogenous early universe?

