CS732 : Data Visualization
Technical Report for Datathon-1

Neetha Reddy
International Institute of Information Technology Bangalore
IMT2018050
Neetha.Reddy@iiitb.ac.in

***Abstract*—This technical report consists of a brief overview of methodologies employed in the visualisation of Ionization front instability simulation data set. Further, this report attempts to draw inferences from the results of the above mentioned visualizations.**

I. Introduction

The ionization front instability simulation data set submitted by Mike Norman and Daniel Whalen to the 2008 IEEE Visualization Design Contest is being used for the scope of this paper. The dataset contains *10 scalar fields*:

1) Total particle density (# of particles/cm^3)
2) Gas temperature (degrees Kelvin)
3) H mass abundance
4) H+ mass abundance
5) He mass abundance
6) He+ mass abundance
7) He++ mass abundance
8) H- mass abundance
9) H_2 mass abundance
10) H_2+ mass abundance

and a *velocity field*.

II. Methods

We have visualized all the data using *numpy*, *matplotlib* and additional packages like MinMaxScaler from *sklearn* for normalising and animation from *matplotlib* for creating animations of the visualized values over different timestamps. For all variables chosen for scalar field visualization as well as vector field visualization, the data is visualized over 11 timestamps across the 200 timestamps in the dataset with a gap of 20 timestamps, i.e. 0000, 0019, 0039,..., 0199. This was done to ensure that necessary information is communicated over the course of the dataset at a much faster speed (sampling of the dataset).

A. Scalar Field Visualization

The contourf() function was used to generate the plots. The colour mapping for each variable was chosen to ensure that the interpretation of the visualized data is easy. The maximum and the minimum values for the colour maps of each variable have been fixed after studying the upper limits and lower limits of the given data.

1) Log of total particle density: : Here, the log of the total particle density data is displayed, thereby making it easier to observe the changes in the particle density. The colormap chosen for this plot was the Diverging bwr colormap. This was chosen since it captured the information about the change in density the best among others. We have chosen to study this field to understand the distribution of mass due to chemical reactions. An example is shown in Fig. 1.

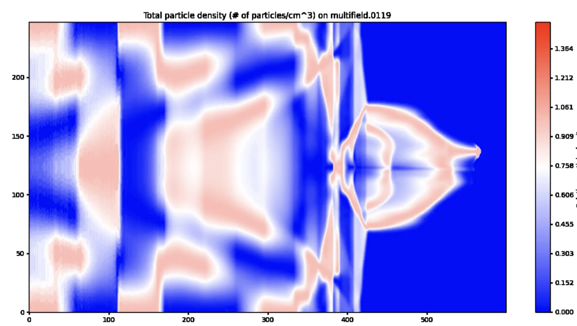

Fig. 1. Log of total particle density.

2) Gas temperature: : The colormap chosen for this plot was the Sequential YlOrRd colormap. This colormap is intuitive because it works akin to thermal sensors detecting heat from an object. An example is shown in Fig. 2. Here the light yellow region is for extremely cold temperatures. We observe that completely ionised particles are at much higher temperatures (about 2000K) than cool, neutral gas (72K) and shocked gas (2000-3000K). We have chosen to study this field to understand how the temperature of gases might cause several ionisation reactions, giving rise to volatile by-products.

3) He+ mass abundance: : The colormap chosen for this plot was the Sequential (2) summer colormap. This scalar field was chosen for study since the helium being a small, noble gas has the highest ionisation energy. Due to the high temperatures of the UV photons, despite its high ionization energy, the amount of He+ ions increases. An example is shown in Fig. 3.

4) H+ mass abundance: : The colormap chosen for this plot was the Sequential (2) winter colormap. This scalar field was chosen for study since the H+ ions resulting due to the high temperatures of the UV photons, will form react with

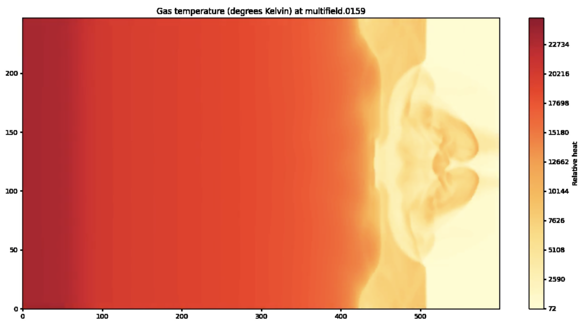

Fig. 2. Gas temperatures (in Kelvin)

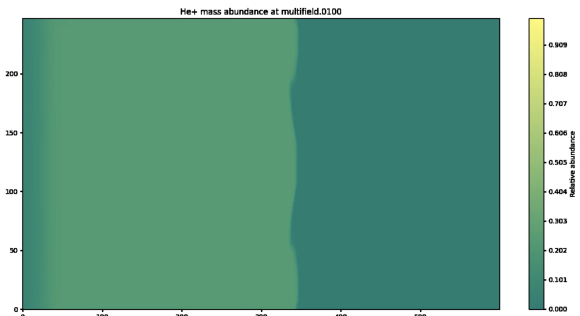

Fig. 3. He+ mass abundance

helium to form helium hydride ion. This is believed to be one of the first compound formed in the Universe after the Big Bang and is extremely volatile. An example is shown in Fig. 4.

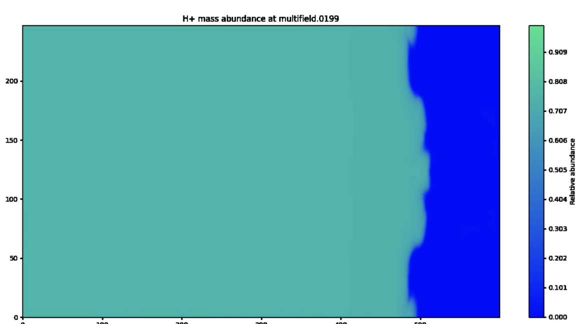

Fig. 4. H+ mass abundance

B. Vector Field Visualization

We use the quiver() function to generate the plots. Here, we use the Curl X and Curl Y data of the velocity field(since we chose to visualize an X-Y plane). These curl values are calculated as follows:

$$curl_x(i,j,k)$$
$$= \frac{vz(i,j+1,k) - vz(i,j,k) - vy(i,j,k+1) + vy(i,j,k)}{0.001}$$

$$curl_y(i,j,k)$$
$$= \frac{vx(i,j,k+1) - vx(i,j,k) - vz(i+1,j,k) + vz(i,j,k)}{0.001}$$

Curl magnitudes of the velocity field are used as an estimator of turbulence. After reading and calculating the curl values, they have been plotted using quiver plots in the following ways:

1) **Basic quiver plot:** Here the curl vector at each index is visualized using adjusted arrows from the quiver plot. This turbulence is shown in Fig. 6.
2) **Quiver+Contour plot:** Here the curl vector at each index is visualized using adjusted arrows from the quiver plot. This is plotted alongside the log of density function which in turn is plotted using the contourf() function. This plot helps us visualize the relation between turbulence and density at each point. This plot is shown in Fig. 7. Here the quiver plot is represented by white arrows while the scalar field log of density, uses the Sequential (2) summer colormap.

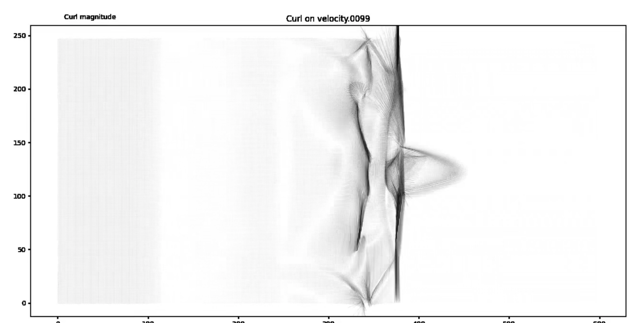

Fig. 5. Curl of velocity field (XY plane)

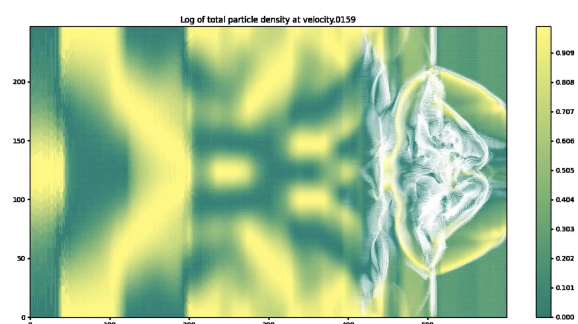

Fig. 6. Curl of velocity field with log of density(XY plane)

Note:- These visualizations were done in Kaggle notebooks. Due to memory limit exceeded error, the computations for quiver plots were animated in batches of 3 and then merged.

III. TASKS

How did you arrive at the timesteps? Answered previously in II METHODS

How did you arrive at which plane you are going to explore?
Each data file has 600*248*248 number of lines and each of these lines represent a particular 3D point. They are ordered as (0, 0, 0), (1, 0, 0)....(599, 0, 0), (0, 1, 0), (1, 1, 0), (2, 1, 0)....(599, 1, 0), (0, 2, 0)....(599, 247, 0), (0, 0, 1), (1, 0, 1)....(599, 0, 1), (0, 1, 1), (1, 1, 1), (2, 1, 1)....(599, 247, 247). This means that X indices value most rapidly, followed by Y indices and then Z indices. Therefore, XY plane was chosen with Z=40 due to its ease of indexing and large plane size which allows us to draw meaningful inferences.

What is your rationale for the selection of variables? Answered previously under respective variable subheadings.

For contours, will you use the same contour values for all time steps?
Yes, we used the same contour values for a particular variable over all time steps. This was done to maintain consistency of resulting plots and allow easy inference drawing from the animations.

Did any color palette outperform the others? How would you rationalize the performance?
While deciding on a color palette, I have explored sequential, diverging, qualitative color palettes to find a suitable match for a contour plot. I have observed that sequential colormaps(Fig. 7) despite being effective, struggle to provide easy distinction within the spectrum (mid-range values). Despite being aesthetically pleasing, qualitative colormaps made it difficult to draw inferences(Fig. 8). I found Sequential (2) colormaps(Fig. 2) and Diverging colormaps (Fig. 1) the most for visualising variables within the scope of this paper.

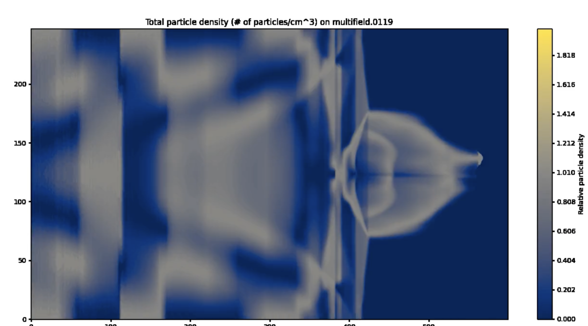

Fig. 7. Curl of velocity field with log of density(XY plane)

Did combining 2 visualization techniques enable you to make joint inferences of different fields?
Yes, combining density and curl as shown in Fig. 6, helped me realise that the density distribution which occurs due to chemical reactions follows the pattern of the turbulence which in turn is detected using the curl magnitude of the velocity field.

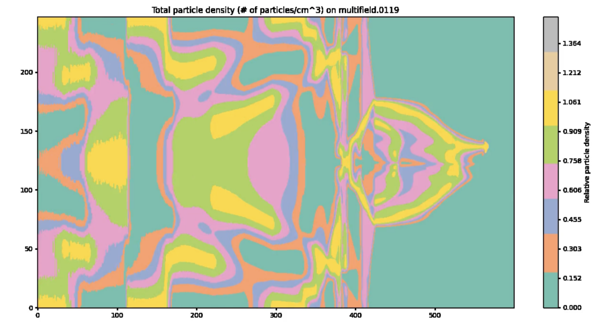

Fig. 8. Curl of velocity field with log of density(XY plane)

Do your visualizations help you infer the shadow instability as shown in the image in the contest website? What in your choice of the data for visualization allowed you to see the moving front?
Yes, my visualizations helped me the infer shadow instability as depicted in the contest website in various seemingly similar plots. While different visualizations like log of particle density and gas temperature allowed me to see the moving front, my choice of the data for the visualization is the curl magnitude of the velocity field. This is because the quiver plot of the velocity curl made it easier for me to visualize the moving front and is the most similar to the actual image in the context website(Fig. 5).

REFERENCES

[1] IEEE Visualization 2008 Design Contest.